

D.E.A. Acoustique, Traitement du Signal et Informatique Appliqués à la musique
Université Paris VI, ENST, Université d'Aix-Marseille II, UJF Grenoble I

Reproduction de champ sonore sur une zone étendue de l'espace

Rapport de stage de D.E.A.

Mathieu GUILLAUME

Sous la direction de Yves Grenier

30 juin 2003



Laboratoire de Traitement de Signal et des Images
École Nationale Supérieure des Télécommunications
Année 2002-2003

Table des matières

1	Etude bibliographique	2
A	Préliminaires	2
A.1	Rappels sur la théorie des champs	3
A.2	Résumé de la théorie de l’acoustique linéaire	7
A.3	Fonctions de Green	9
A.4	Equations intégrales	14
A.5	Conclusion	18
B	Enregistrement du champ sonore	19
B.1	Préliminaires	20
B.2	Des différentes méthodologies pour l’enregistrement du champ sonore	22
B.3	Prospectives d’applications à la prise de son	27
C	Reconstruction du champ sonore	29
C.1	Panorama des différentes méthodes de spatialisation des sons	29
C.2	Des différentes méthodologies pour la synthèse d’un champ sonore	31
C.3	Prospectives d’applications à la restitution du son	34
2	Conception d’un nouveau système de restitution sonore	37
A	Présentation du projet Demix-Remix	37
A.1	Séparation de sources	38
A.2	Indexation multimédia	39
B	Paramètres synthétiques des dispositifs de restitution sonore	39
B.1	Taille de la zone d’écoute	39
B.2	Modélisation des sources acoustiques	40
B.3	Implantation d’un effet de salle artificiel	41
B.4	Égalisation du système de reproduction	42
B.5	Conclusion	42
C	Présentation de la solution retenue	43
C.1	Description générale	43
C.2	Présentation du problème direct	44
C.3	Formulation du problème inverse	45
D	Méthodes de calcul des filtres inverses	48
D.1	Préliminaires : Complexité de l’inversion d’un canal acoustique	48
D.2	Méthode directe d’inversion du canal acoustique	51
D.3	Approche du filtrage inverse par des méthodes adaptatives : la solution optimale?	56
3	Conclusion	59
A	Ce qui a été fait.	59
B	... Et ce qu’il reste encore à faire	60

A	Des différentes méthodes de mesure de réponses impulsionnelles en acoustique	61
A	Les méthodes existantes	61
A.1	L'enregistreur de niveau	62
A.2	Time-Delay Spectrometry (TDS)	62
A.3	Dual-Channel FFT Analysis	64
A.4	Excitation en régime harmonique	65
A.5	Impulsions	65
A.6	Maximum-Length Sequences (MLS)	65
A.7	Excitation par des signaux périodiques de longueur 2^N	66
A.8	Excitation du système par des signaux Chirp	67
B	Présentation de la méthode retenue	68
B.1	Critères de choix	68
B.2	Description du protocole expérimental	69

Remerciements

Je tiens à remercier particulièrement mon encadrant Yves Grenier, ainsi que Gaël Richard, Bertrand David, et Denis Matignon pour tous les conseils avisés qu'ils ont su me donner, et le temps qu'ils ont pu me consacrer.

Je tiens aussi à remercier toute l'équipe de thésards (Thomas Fillon, Miguel Alonso) et de stagiaires présents à l'ENST (Chloé Clavel, Sylvain Paturel, et Alexandre Tauvy) au moment de mon stage, qui m'ont permis de passer d'agréables moments dans ces locaux, particulièrement au sous-sol. Mais je n'oublie pas non plus celles et ceux qui errent dans les bureaux du premier étage.

Résumé

Le travail présenté marque le début d'une collaboration de trois centres de recherche sur le thème de la spatialisation sonore : L'École Nationale Supérieure des Télécommunications, le Laboratoire de Mécanique et d'Acoustique de Marseille, et l'Institut de Recherche et Coordination Acoustique et Musique (IRCAM). Il constitue le commencement d'un projet sur la reproduction d'un champ sonore sur une zone étendue de l'espace, pour ce qui concerne l'ENST, qui possédait peu d'expérience dans le domaine au début de ce stage. Ce projet sera poursuivi en thèse à la rentrée prochaine.

Dans le passé, les aspects temporels et fréquentiels du champ sonore retenaient beaucoup plus notre attention que l'organisation spatiale. Cela a permis le développement de systèmes d'enregistrement et de restitution de qualité au regard de ces critères. Pourtant, l'intérêt pour l'organisation spatiale du champ sonore n'est pas nouveau, étant donné que les scientifiques s'y sont intéressés depuis les années 30. Le développement de l'enregistrement et de la restitution stéréophonique, pendant les années 60, en est le plus bel exemple. A cette époque, le recours à un nombre plus élevé de capteurs et de transducteurs était déjà dans les esprits, mais ce sont plus des difficultés d'ordre économique, et liées à la technologie, qui ont préféré un système n'utilisant que deux microphones pour la prise de sons, et deux enceintes pour la restitution. En effet, les systèmes utilisant plus de deux voies n'étaient pas jugés convaincants, et n'apportaient pas suffisamment de valeur ajoutée pour justifier un tel investissement. Ce n'est que récemment, avec l'apparition des nouveaux systèmes de restitution à cinq voies utilisés principalement par l'industrie cinématographique que la question de l'organisation spatiale du champ sonore est revenue à l'ordre du jour. Elle connaît même un réel essor.

Malheureusement, ces systèmes de restitution souffrent d'un inconvénient majeur : ils ne fournissent une bonne image du relief sonore que pour une seule position dans l'espace. Pour pallier cet inconvénient, A.J. Berkhout fut le premier à imaginer un système de restitution de champ sonore sur une zone étendue de l'espace, dont le développement a débuté il y a une quinzaine d'années maintenant, comme en témoigne le premier papier sur le contrôle du champ sonore par une approche holographique.

Toutes les approches actuelles permettant de reproduire le champ sonore sur une zone étendue de l'espace sont basées, de manière ostentatoire ou cachée, sur l'équation intégrale de Kirchhoff, qui n'est qu'une généralisation du principe de Huygens énoncé dans le domaine de l'optique, mais valable pour toutes ondes, dans le cas où les sources secondaires permettant de reconstruire le champ initial ne sont plus forcément situées sur un front d'onde de la source primaire. Parmi ces techniques, nous pouvons citer la Wave Field Synthesis, développée à l'université technologique de Delft, ainsi qu'à l'IRCAM, l'holophonie et les méthodes ambisoniques, développées au Centre National d'Etudes en Télécommunications (CNET) de Lannion, mais aussi les techniques de contrôle des conditions aux limites, qui n'ont pas encore montré leur potentiel dans la pratique, et dont le système développé à l'ENST essaiera de s'inspirer.

Le travail effectué au cours de ce stage comporte une composante majoritairement bibliographique et théorique. Il a permis de faire une synthèse sur les différents procédés de restitution sonore existants, de connaître leurs avantages et leurs défauts. A partir de ce constat, une nouvelle piste pour la restitution sonore sur une zone étendue est envisagée, essayant de se démarquer des approches actuellement développées dans les autres centres de recherche, afin d'étudier une vision complémentaire du domaine. Les problèmes sous-jacents à cette nouvelle technique sont mis en évidence, et des pistes de solution sont fournies, même si leur mise en œuvre dans la pratique n'a pas encore été réalisée pour le moment.

Introduction

Cette partie introductive a pour finalité de répondre à la question : Comment lire ce rapport ? En effet, il peut sembler assez volumineux au premier abord, et il n'est pas nécessaire que tout lecteur le lise en totalité pour en apprécier le contenu. Ainsi, si l'objectif du lecteur est de se faire une idée sur le travail qui a été effectué, sans vouloir trop y consacrer de temps, je lui conseillerais de commencer directement la lecture de ce rapport au deuxième chapitre, qui marque la partie la plus personnelle de ce travail, et qui concerne la mise en place de nouvelles idées pour la reproduction fidèle d'un champ sonore sur une zone étendue de l'espace. Normalement, tous les résultats utilisés renvoient à leur démonstration effectuée dans la première partie au moment où ils apparaissent dans le texte. Par contre, si l'objectif du lecteur est d'approfondir ses connaissances dans le domaine de la restitution du champ sonore sur une zone étendue de l'espace, il lui est vivement conseillé de le lire en totalité.

Quoiqu'il en soit, voici un descriptif sommaire des principales parties de ce rapport, et il est conseillé au lecteur de se reporter à la table des matières afin de localiser rapidement l'information désirée s'il désire de plus amples renseignements :

- Le premier chapitre constitue une étude bibliographique ayant rapport au sujet de la reproduction de champ sonore sur une zone étendue de l'espace. Tous les rappels théoriques permettant de comprendre les difficultés d'un tel problème y sont effectués. Puis, nous essayons de décrire les différentes techniques possibles permettant de caractériser le champ sonore dans un certain volume de l'espace, leur éventuelle mise en pratique, et les différentes perspectives d'utilisation de ces techniques dans le domaine de la prise de son. Enfin, nous détaillons les méthodes possibles de reconstruction d'un champ sonore dans un certain volume de l'espace, les procédés à mettre en œuvre pour y parvenir en pratique, puis différentes applications de ces systèmes de restitution sonore sur zone étendue.
- Le deuxième chapitre commence par décrire le contexte dans lequel s'inscrit le stage effectué. Puis il est suivi d'une réflexion personnelle permettant de décrire quelques-uns des paramètres essentiels de tout système de restitution sonore, ce qui aboutit à la description d'une nouvelle possibilité d'étude dans le domaine de la restitution sonore sur une zone étendue de l'espace. Ce nouvel axe d'étude est ensuite largement décrit, et notamment les problèmes qu'il faudra résoudre pour donner naissance à un prototype en pratique. Nous donnerons finalement des pistes de solution permettant de surmonter ces difficultés.
- Le dernier chapitre vient conclure cette étude en effectuant le bilan du travail qui a été effectué, et le bilan des choses qu'il reste à faire.

Chapitre 1

Etude bibliographique

Cette étude bibliographique est une synthèse de toutes les connaissances nécessaires pour aborder le domaine de l'enregistrement et de la reproduction du champ sonore sur une zone étendue de l'espace avec sérénité. Beaucoup des concepts de bases, autant sur le plan des mathématiques que sur le plan de la physique, et notamment l'acoustique, sont rappelés, de sorte que le lecteur non avisé n'aura pas besoin d'effectuer une multitude de recherches complémentaires dans d'autres ouvrages pour pouvoir cerner le sujet. Cette étude se compose de trois parties :

- La première partie détaille tous les outils théoriques nécessaires à la compréhension des deux parties suivantes. Elle aborde notamment les outils de base utiles pour l'étude des champs, détaille les équations de l'acoustique linéaire, et donne des éléments concernant la résolution de ces équations en adoptant le point de vue des fonctions de Green, puis des équations intégrales associées.
- La deuxième partie envisage le problème de l'enregistrement du champ sonore sur une zone étendue. Elle évoque en premier lieu les caractéristiques requises pour le dispositif d'enregistrement afin d'obtenir un enregistrement parfait du champ, puis elle expose les limites associées aux dispositifs réellement utilisés dans la pratique. Pour finir, des réflexions sont faites, dont le but est d'envisager des nouvelles méthodes dans le domaine de la prise de son.
- La dernière partie concerne le domaine de la reproduction du champ sonore sur une zone étendue de l'espace. Un bilan des différentes méthodes de spatialisation actuelles y est effectué, puis différentes stratégies sont exposées pour restituer fidèlement le champ sonore sur une zone étendue. Finalement, une prospective d'utilisation des dispositifs exposés tout au long de cette partie, dans le domaine de la reproduction sonore, est effectuée.

A Préliminaires

L'étude théorique de l'acoustique s'effectue par la connaissance de ses grandeurs caractéristiques —la pression acoustique, et la vitesse particulaire du fluide, pour ne citer que les deux plus importantes— et des équations régissant l'évolution de ces grandeurs. Le formalisme de la théorie des champs, scalaires ou vectoriels, en physique, s'applique à ces grandeurs caractéristiques, c'est pourquoi de larges rappels concernant les opérateurs élémentaires appliqués sur les champs, ainsi que les théorèmes permettant une manipulation de ces opérateurs, sont effectués. Puis on effectue une sorte de memento sur les équations de l'acoustique linéaire. Pour continuer, l'étude de ces équations à partir du formalisme des fonctions de Green est faite, ce qui aboutira dans la dernière partie par l'énoncé des équations intégrales satisfaites par les différentes grandeurs caractéristiques du champ acoustique.

A.1 Rappels sur la théorie des champs

Le but de ce stage est de reproduire à l'identique un champ sonore initialement émis par une source sonore quelconque dans une certaine zone bien définie de l'espace. Avant de reproduire le champ, il est nécessaire au préalable de disposer de sa connaissance, en tout point de l'espace dans lequel on souhaite le restituer, et quel que soit l'instant considéré. Notre organe de perception des événements sonores est l'oreille, et plus particulièrement le tympan, qui est uniquement sensible aux variations de pression d'un fluide, c'est pourquoi il est nécessaire de connaître le champ de pression acoustique $p(\vec{r}, t)$, en tout point de la zone de restitution et à tout instant. Comme nous le verrons plus tard dans la partie sur les équations de l'acoustique linéaire, le champ de pression acoustique $p(\vec{r}, t)$ est couplé à celui de la vitesse particulaire du fluide $\vec{v}(\vec{r}, t)$, c'est pourquoi nous effectuons des rappels de la théorie des champs, qu'ils soient scalaires ou vectoriels, tout au long de cette partie.

Dans l'expression $p(\vec{r}, t)$, \vec{r} désigne le vecteur position d'un point M , dans un référentiel d'étude \mathcal{R} , supposé galiléen, et dont l'origine est O . La notation $\vec{r} = \overrightarrow{OM}$ est synthétique, c'est-à-dire qu'elle peut désigner tout aussi bien les coordonnées cartésiennes $\vec{r} = (x, y, z)$, que les coordonnées polaires $\vec{r} = (r, \theta, z)$ ou les coordonnées sphériques $\vec{r} = (r, \theta, \phi)$.

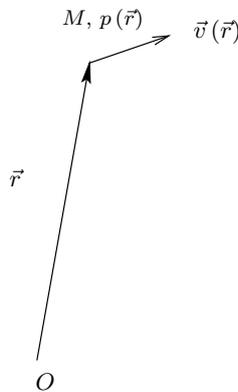


FIG. 1.1 – Représentation des champs de vecteurs

Les développements effectués dans cette partie viennent principalement de [1] et de [2].

A.1.1 Gradient, divergence et rotationnel

Voici quelques rappels concernant la définition et l'interprétation des principaux opérateurs agissant soit sur des champs scalaires, soit sur des champs vectoriels.

A.1.1.1 Gradient d'un champ scalaire

Le gradient¹ d'un champ scalaire est défini par la relation générale :

$$dp = \text{grad } p \cdot d\vec{r} \quad (1.1)$$

Ainsi, il décrit le taux de variation spatiale du champ $p(\vec{r})$ entre deux points infiniment proches : $p(\vec{r} + d\vec{r}) = p(\vec{r}) + dp = p(\vec{r}) + \text{grad } p \cdot d\vec{r}$. Il est immédiat de voir que la variation spatiale du champ est maximale dans la direction du gradient, car c'est dans cette direction que le produit scalaire est maximal —les deux vecteurs $\text{grad } p$ et $d\vec{r}$ sont dans ce cas colinéaires.

¹Tout au long de ce texte, nous n'emploierons jamais la notation ∇ , pour noter le gradient, la divergence ou le rotationnel, car la ressemblance entre ces notations a vite tendance à induire le lecteur en erreur.

Deux propriétés fondamentales découlent de cette remarque :

$$\int_A^B \text{grad } p \cdot d\vec{\mathcal{L}} = \int_A^B dp = p_B - p_A \quad (1.2)$$

La circulation du gradient entre les points A et B est indépendante du trajet effectué entre A et B . On dit que le gradient d'un champ scalaire est à *circulation conservative*, ce qui se traduit par la relation suivante : $\oint_C \text{grad } p \cdot d\vec{\mathcal{L}} = 0$.

Le gradient d'un champ scalaire est orthogonal à toute surface équi- p . En effet, pour deux points infiniment proches appartenant à une telle surface, nous avons $dp = 0 = \text{grad } p \cdot d\vec{r}$, et $\text{grad } p$ est bien orthogonal à $d\vec{r}$.

A.1.1.2 Divergence d'un champ vectoriel

La divergence d'un champ vectoriel $\vec{v}(\vec{r})$ est définie comme la limite du rapport entre le flux de $\vec{v}(\vec{r})$ à travers une surface fermée \mathcal{S} et le volume intérieur τ délimité par la surface fermée, lorsque ce volume tend vers zéro :

$$\lim_{\tau \rightarrow 0} \frac{\oiint_{\mathcal{S}} \vec{v}(\vec{r}) \cdot d\vec{S}}{\tau} = \text{div } \vec{v}(\vec{r}) \quad , \quad d\vec{S} \text{ orientée vers l'extérieur} \quad (1.3)$$

La divergence est un opérateur vectoriel dont le résultat est un champ scalaire.

Interprétation physique : considérons le cas où le champ $\vec{v}(\vec{r})$ désigne celui de la vitesse particulaire d'un fluide. Dans ce cas, si la divergence est positive, cela implique que le flux de $\vec{v}(\vec{r})$ à travers \mathcal{S} est positif, donc les particules de fluide ont tendance à être expulsées du volume intérieur τ . Le mouvement associé est une *détente* locale du fluide. Inversement, si la divergence est négative, le flux de $\vec{v}(\vec{r})$ à travers \mathcal{S} est négatif, donc les particules de fluide ont tendance à rentrer dans le volume intérieur τ . Le mouvement correspondant est une *compression* locale du fluide.

A.1.1.3 Rotationnel d'un champ vectoriel

D'un point de vue physique, le rotationnel d'un champ vectoriel peut être défini comme la limite du rapport entre la circulation d'un champ vectoriel $\vec{v}(\vec{r})$ le long d'un contour fermé \mathcal{C} et une surface fermée S s'appuyant sur le contour \mathcal{C} —le vecteur surface \vec{S} est orienté selon la règle du tire-bouchon— lorsque la surface tend vers zéro :

$$\lim_{S \rightarrow 0} \frac{\oint_{\mathcal{C}} \vec{v}(\vec{r}) \cdot d\vec{\mathcal{L}}}{S} = \text{rot } \vec{v}(\vec{r}) \cdot \vec{n} \quad , \quad \vec{n} = \frac{\vec{S}}{S} \text{ vecteur unitaire normal à la surface } S \quad (1.4)$$

Le rotationnel est un opérateur vectoriel dont le résultat est un champ vectoriel.

Interprétation physique : le rotationnel est un peu plus complexe à interpréter que les deux précédents opérateurs. Pour y parvenir, considérons un contour fermé \mathcal{C} positionné de manière orthogonale à la direction du rotationnel. Ainsi, si le rotationnel n'est pas nul, le membre de droite de l'équation 1.4 n'est pas nul — $\text{rot } \vec{v}(\vec{r})$ et \vec{n} sont colinéaires— ce qui implique que la circulation du champ $\vec{v}(\vec{r})$ sur le contour fermé \mathcal{C} n'est pas nulle : les particules ont tendance à tourner autour d'un axe dans le sens trigonométrique, la direction et le sens de cet axe étant donnés par ceux du rotationnel du champ au point considéré.

Remarque : les mouvements de fluide donnant naissance à des sensations auditives au niveau de nos tympans sont seulement ceux qui mettent en jeu des oscillations de particules autour d'une position d'équilibre. Ainsi, des mouvements globaux de particules, qu'il s'agisse d'une translation ou d'une rotation —existence d'un vecteur tourbillon en mécanique des fluides par exemple—, ne sont pas perçus par l'oreille.

A.1.2 Théorème de Stokes

Énoncé : la circulation d'un champ vectoriel $\vec{v}(\vec{r})$ sur un contour fermé \mathcal{C} est égale au flux du rotationnel de ce même champ à travers une surface fermée \mathcal{S} s'appuyant sur le contour \mathcal{C} :

$$\oint_{\mathcal{C}} \vec{v}(\vec{r}) \cdot d\vec{\mathcal{L}} = \iint_{\mathcal{S}} r\vec{\otimes}t \vec{v}(\vec{r}) \cdot d\vec{\mathcal{S}} \quad , \quad d\vec{\mathcal{S}} \text{ orientée par la règle du tire-bouchon} \quad (1.5)$$

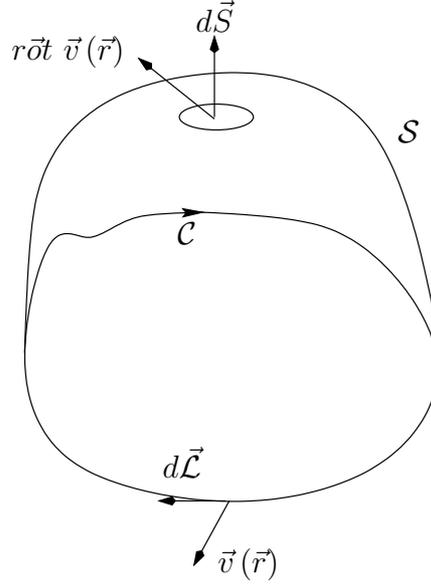


FIG. 1.2 – Théorème de Stokes

Les conséquences du théorème de *Stokes* sont importantes dans le domaine de l'acoustique. En effet, nous avons dit dans le paragraphe précédent que seuls les mouvements d'oscillation autour d'une position d'équilibre donnaient naissance à des phénomènes acoustiques, donc nous pouvons faire l'hypothèse que $r\vec{\otimes}t \vec{v}(\vec{r}) = 0$ si nous ne considérons que la composante du mouvement du fluide donnant naissance à des phénomènes acoustiques. Ainsi, le deuxième membre de l'équation 1.5 est nul et la composante acoustique du champ de la vitesse particulière est à circulation conservative. La composante acoustique de la vitesse particulière dérive donc d'un potentiel scalaire :

$$r\vec{\otimes}t \vec{v}_{ac}(\vec{r}) = 0 \Leftrightarrow \vec{v}_{ac}(\vec{r}) = \text{grad } \phi(\vec{r}) \quad (1.6)$$

Le champ $\vec{v}_{ac}(\vec{r})$, qui possédait *a priori* trois composantes indépendantes, peut en fait être calculé par le moyen d'un unique champ scalaire $\phi(\vec{r})$. Il existe donc un lien entre les trois composantes du champ $\vec{v}_{ac}(\vec{r})$. Le recours au potentiel scalaire² $\phi(\vec{r}, t)$ est souvent effectué en acoustique linéaire, car il permet de déduire à la fois la pression acoustique $p_{ac}(\vec{r}, t)$ et la vitesse particulière acoustique $\vec{v}_{ac}(\vec{r}, t)$, comme nous le verrons à la section A.2.

A.1.3 Théorème de Gauss

Énoncé : le flux d'un champ vectoriel $\vec{v}(\vec{r})$ à travers une surface fermée \mathcal{S} est égal à l'intégrale de la divergence de ce champ sur tout le volume intérieur τ délimité par la surface fermée \mathcal{S} :

$$\oiint_{\mathcal{S}} \vec{v}(\vec{r}) \cdot d\vec{\mathcal{S}} = \iiint_{\tau} \text{div } \vec{v}(\vec{r}) dV \quad , \quad d\vec{\mathcal{S}} \text{ étant orientée vers l'extérieur} \quad (1.7)$$

²il est nécessaire de tenir compte aussi de l'évolution temporelle de ce potentiel scalaire

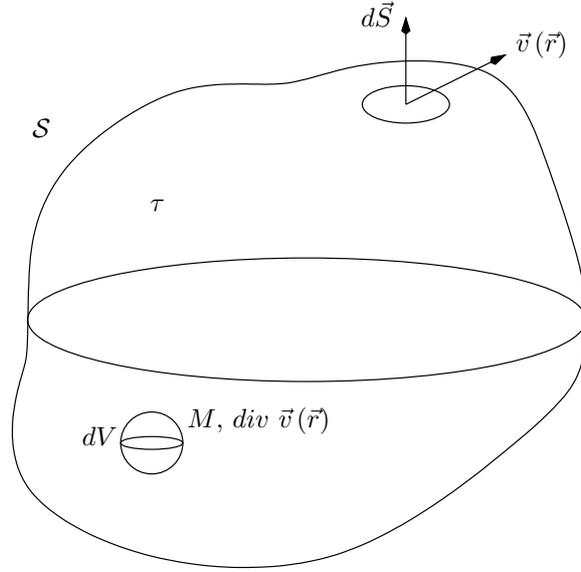


FIG. 1.3 – Théorème de Gauss

Ce théorème met en évidence la relation qui existe entre le comportement du champ sur une surface fermée et son comportement à l'intérieur du volume délimité par cette surface fermée. Cette propriété sera précisée un peu plus tard, car elle nous permettra d'aboutir à l'expression des équations intégrales aux sections A.3 et A.4. C'est aussi cette propriété qui est à l'origine de l'holographie, qu'elle soit optique ou acoustique, car l'objectif de la discipline consiste à enregistrer toute l'information tridimensionnelle d'un champ à partir de sa mesure sur une surface uniquement.

A.1.4 Théorème de Green

Le premier théorème de Green n'est qu'une expression différente du théorème de Gauss, dans le cas où le champ vectoriel est sous la forme : $\vec{I}(\vec{r}) = p(\vec{r}) \text{ grad } \phi(\vec{r})$. Dans ce cas, si nous nous souvenons l'identité $\text{div}(\text{grad}) = \nabla^2$ où ∇^2 désigne le laplacien, l'expression de la divergence de $\vec{I}(\vec{r})$ peut se 'simplifier' comme ceci :

$$\text{div } \vec{I}(\vec{r}) = \text{grad } p(\vec{r}) \cdot \text{grad } \phi(\vec{r}) + p(\vec{r}) \nabla^2 \phi(\vec{r}) \quad (1.8)$$

Ainsi, l'expression du *premier théorème de Green* est la suivante :

$$\oint_S p(\vec{r}) \text{ grad } \phi(\vec{r}) \cdot d\vec{S} = \iiint_{\tau} (\text{grad } p(\vec{r}) \cdot \text{grad } \phi(\vec{r}) + p(\vec{r}) \nabla^2 \phi(\vec{r})) dV \quad (1.9)$$

Appliquons maintenant le premier théorème de Green au champ $\text{grad } p(\vec{r}) \phi(\vec{r})$, nous obtenons :

$$\oint_S \phi(\vec{r}) \text{ grad } p(\vec{r}) \cdot d\vec{S} = \iiint_{\tau} (\text{grad } p(\vec{r}) \cdot \text{grad } \phi(\vec{r}) + \phi(\vec{r}) \nabla^2 p(\vec{r})) dV \quad (1.10)$$

Si nous soustrayons les deux équations 1.9 et 1.10, nous obtenons l'expression du *deuxième théorème de Green* :

$$\oint_S (p(\vec{r}) \text{ grad } \phi(\vec{r}) - \phi(\vec{r}) \text{ grad } p(\vec{r})) \cdot d\vec{S} = \iiint_{\tau} (p(\vec{r}) \cdot \nabla^2 \phi(\vec{r}) - \nabla^2 p(\vec{r}) \cdot \phi(\vec{r})) dV \quad (1.11)$$

Nous verrons un peu plus tard, à la section A.3 que le deuxième théorème de Green permettra d'aboutir à l'équation intégrale de Kirchoff-Helmholtz et à l'équation intégrale de Kirchoff.

A.2 Résumé de la théorie de l'acoustique linéaire

A.2.1 Equations de base [3]

Les inconnues dans le domaine de l'acoustique sont au nombre de trois : la pression acoustique du fluide $p(\vec{r}, t)$, la variation locale de masse volumique $\rho(\vec{r}, t)$ et la composante acoustique de la vitesse particulière $\vec{v}(\vec{r}, t)$, et cette dernière inconnue est vectorielle. Nous avons déjà précisé que ces équations concernent uniquement le comportement dynamique de la pression, de la masse volumique, et de la vitesse particulière, c'est pourquoi les grandeurs mentionnées ne comportent pas de composante statique. Il est nécessaire de pouvoir disposer du nombre suffisant d'équations pour déterminer correctement ces inconnues. Ces équations sont au nombre de trois —deux équations scalaires et une équation vectorielle— et prennent la forme suivante si nous supposons être dans le domaine de linéarité de l'acoustique (variations de pression acoustique très inférieures à la pression statique atmosphérique) :

- *L'équation d'état du fluide*, qui indique la relation entre les différentes grandeurs thermodynamiques du fluide. Elle se résume de la manière suivante pour les grandeurs considérées :

$$p(\vec{r}, t) = c^2 \rho(\vec{r}, t) \quad \text{avec } c = \sqrt{\frac{\gamma RT}{M}} \text{ célérité du son dans le milieu} \quad (1.12)$$

où $\gamma = \frac{C_P}{C_V}$ désigne le rapport des capacités calorifiques du fluide, R désigne la constante des gaz parfaits, T sa température, et M sa masse molaire.

- *L'équation de la dynamique*, connue sous le nom d'équation d'Euler dans le domaine de la mécanique des fluides, et qui prend la forme suivante pour le cas considéré :

$$\text{grad } p(\vec{r}, t) + \rho_0 \frac{\partial \vec{v}(\vec{r}, t)}{\partial t} = \rho_0 \vec{F}(\vec{r}, t) \quad (1.13)$$

où ρ_0 désigne la masse volumique du fluide au repos, et \vec{F} désigne la résultante des forces extérieures par unité de masse à laquelle est soumise le fluide.

- *L'équation de continuité*, qui n'est que l'expression du principe de conservation de la masse au sein du fluide :

$$\frac{\partial p(\vec{r}, t)}{\partial t} + \rho_0 c^2 \text{div } \vec{v}(\vec{r}, t) = \rho_0 c^2 q(\vec{r}, t) \quad (1.14)$$

où $q(\vec{r}, t)$ désigne le débit volumique réduit de sources comprises dans le volume élémentaire $d\tau$ centré sur le point repéré par \vec{r} .

A.2.2 Equations des ondes

A.2.2.1 Point de vue général Les équations 1.13 et 1.14 forment un système d'équations couplées. En prenant la divergence de 1.13 et en soustrayant la dérivée temporelle de 1.14 multipliée par $1/c^2$, nous obtenons l'équation de propagation pour la pression acoustique :

$$\nabla^2 p(\vec{r}, t) - \frac{1}{c^2} \frac{\partial^2 p(\vec{r}, t)}{\partial t^2} = \rho_0 \left(\text{div } \vec{F}(\vec{r}, t) - \frac{\partial q(\vec{r}, t)}{\partial t} \right) \quad (1.15)$$

En prenant le gradient de 1.14, et en soustrayant la dérivée temporelle de 1.13, en utilisant la relation $r \vec{\partial} t \ r \vec{\partial} t = \text{grad } \text{div} - \nabla^2$, et le fait que $r \vec{\partial} t \ \vec{v}(\vec{r}, t) = 0$, nous obtenons l'équation de propagation pour la vitesse particulière du fluide :

$$\nabla^2 \vec{v}(\vec{r}, t) - \frac{1}{c^2} \frac{\partial^2 \vec{v}(\vec{r}, t)}{\partial t^2} = \text{grad } q(\vec{r}, t) - \frac{1}{c^2} \frac{\partial \vec{F}(\vec{r}, t)}{\partial t} \quad (1.16)$$

Nous avons vu à la section A.1.2 que la composante acoustique de la vitesse particulière dérivait d'un potentiel scalaire, c'est-à-dire qu'il existe $\phi(\vec{r}, t)$ tel que $\vec{v}(\vec{r}, t) = -\text{grad } \phi(\vec{r}, t)$. Le potentiel scalaire satisfait lui aussi l'équation des ondes :

$$\nabla^2 \phi(\vec{r}, t) - \frac{1}{c^2} \frac{\partial^2 \phi(\vec{r}, t)}{\partial t^2} = -q(\vec{r}, t) - \frac{1}{c^2} \frac{\partial U(\vec{r}, t)}{\partial t} \quad (1.17)$$

où $U(\vec{r}, t)$ désigne l'énergie potentielle associée à la résultante des forces extérieures, supposées à circulation conservative : $\vec{F}(\vec{r}, t) = -\text{grad } U(\vec{r}, t)$.

Nous avons dit que le potentiel scalaire jouait un rôle important en acoustique, car il permet de déduire aussi bien le champ de pression acoustique que celui de la vitesse particulière. En effet, il est possible de le faire apparaître dans l'équation 1.13, en l'absence de résultante de forces extérieures. Nous en déduisons les relations suivantes :

$$\begin{cases} \vec{v}(\vec{r}, t) &= -\text{grad } \phi(\vec{r}, t) \\ p(\vec{r}, t) &= \rho_0 \frac{\partial \phi(\vec{r}, t)}{\partial t} \end{cases} \quad (1.18)$$

Dans la suite, nous nous concentrerons uniquement sur l'étude de l'équation 1.15 qui concerne la pression acoustique, pour les raisons déjà mentionnées auparavant, c'est-à-dire que c'est la seule grandeur à laquelle le sens de l'ouïe soit sensible. Nous devons nous intéresser au problème de l'existence et de l'unicité des solutions à cette équation aux dérivées partielles. Cette équation appartient à la catégorie des équations aux dérivées partielles hyperboliques — un lecteur avide de plus de détails sur les sujets se référera à [2], pp. 676-692. La question de l'existence de solutions ne soulève pas de difficultés majeures. Néanmoins il est nécessaire d'imposer des conditions supplémentaires pour garantir l'unicité de la solution. Les conditions aux limites du domaine peuvent être de type *Dirichlet* —la pression $p(\vec{r}, t)$ est connue en tout point de la surface—, de type *Neumann* —la composante du gradient de pression normale à la surface du domaine est connue en tout point de cette surface—, ou bien mixtes, c'est-à-dire un mélange des deux précédentes. Les conditions initiales sont de type *Cauchy*³, c'est-à-dire que la pression $p(\vec{r}, t)$ et la dérivée temporelle de la pression $\frac{\partial p(\vec{r}, t)}{\partial t}$ sont connues à l'instant initial en tout point du domaine, excepté sur la frontière— et il faut de plus s'assurer que le comportement des conditions initiales aux limites du domaine soit cohérent avec les conditions imposées sur les frontières du domaine à l'instant initial. Si toutes ces conditions sont réunies, alors, le problème physique initial est *bien-posé* : il possède une et seulement une solution (cf. [2], pp. 834, et [4]).

A.2.2.2 Solutions en régime harmonique On recherche souvent des solutions à l'équation précédente sous la forme : $p(\vec{r}, t) = P(\vec{r}) e^{j\omega t}$, solutions dites harmoniques. Si nous introduisons cette forme de solution dans l'équation initiale, nous obtenons finalement l'équation de Helmholtz :

$$\nabla^2 P(\vec{r}) + k^2 P(\vec{r}) = -4\pi s(\vec{r}) \quad (1.19)$$

où $k = \omega/c$ désigne le nombre d'onde, et $s(\vec{r})$ désigne le terme de source de cette équation. Contrairement à l'équation des ondes, l'équation de Helmholtz appartient à la catégorie des équations aux dérivées partielles elliptiques. Le comportement de ces deux équations est donc différent. Alors que dans le premier cas, des conditions initiales de Cauchy, et des conditions aux limites de Dirichlet ou de Neumann suffisaient à garantir l'existence et l'unicité de la solution,

³Les conditions aux limites et les conditions initiales sont classées de manière générale dans ces trois catégories : *Dirichlet* lorsqu'elles concernent uniquement l'ordre 0, *Neumann* lorsqu'elles concernent uniquement l'ordre 1, et *Cauchy* lorsqu'elles concernent à la fois l'ordre 0 et l'ordre 1

des conditions aux limites de Cauchy peuvent parfois trop contraindre le problème posé par l'équation 1.19, tandis que des conditions de Dirichlet ou de Neumann sont parfois insuffisantes pour garantir l'unicité de la solution — le lecteur se référera toujours à [2], pp. 676-692 s'il souhaite obtenir plus de détails.

Pour s'en convaincre nous allons prendre un exemple simple et très utilisé : celui du parallélépipède rectangle de dimensions L_x , L_y , L_z selon les directions x , y , z . On se propose de résoudre le problème caractérisé par l'équation aux dérivées partielles ci-dessus (sans terme de sources), accompagné de conditions aux limites de type Dirichlet, c'est-à-dire que l'on connaît la pression acoustique $p(x, y, z)$ en tout point de la frontière du domaine. Supposons dans un premier temps que la pression acoustique s'annule en tout point de cette surface, les solutions de ces conditions sont largement connues, ce sont les *modes propres* de la salle parallélépipédique :

$$p_{mnp}(x, y, z) = p_i \sin \frac{m\pi x}{L_x} \sin \frac{n\pi y}{L_y} \sin \frac{p\pi z}{L_z} \quad m, n, p \in \mathbb{Z} \quad (1.20)$$

Supposons que $p_1(x, y, z)$ soit solution du problème précédent pour des conditions aux limites quelconques —différentes de zéro— alors, si $k^2 = \left(\frac{m\pi}{L_x}\right)^2 + \left(\frac{n\pi}{L_y}\right)^2 + \left(\frac{p\pi}{L_z}\right)^2$, la fonction $p_1(x, y, z) + \lambda p_{mnp}(x, y, z)$ est aussi solution du problème, et les conditions de Dirichlet sont insuffisantes pour garantir l'unicité de la solution [4]. Nous pouvons garantir l'unicité de la solution en imposant des conditions de Cauchy au lieu des conditions de Dirichlet, c'est-à-dire en ayant la connaissance supplémentaire de la composante normale du gradient de pression en tout point de la frontière. Cela fonctionnera à merveille pour les valeurs propres du problème de Dirichlet propre au domaine considéré. En revanche, l'utilisation de conditions de Cauchy en dehors des valeurs propres de cette équation surconditionne le problème : il existe une relation **naturelle** entre la pression et sa variation par rapport à la normale en dehors des valeurs propres. Nous verrons plus tard que cela posera des problèmes, à la section B.2.1.1, qui envisagera les différentes méthodes possibles de l'enregistrement du champ sonore sur une zone étendue de l'espace.

A.3 Fonctions de Green

On se reportera largement à [2] pour obtenir plus d'informations sur ce sujet.

A.3.1 Présentation

La section précédente a rappelé les équations aux dérivées partielles gouvernant les grandeurs de l'acoustique, ainsi que les conditions nécessaires à imposer afin de garantir l'existence et l'unicité des solutions obtenues. La technique usuelle de résolution de ce problème est de trouver une solution générale à l'équation aux dérivées partielles, ce qui donne généralement une infinité de candidats. On choisit alors parmi ces candidats celui qui vérifie toutes les conditions aux limites —temporelles ou spatiales— supplémentaires. Cette démarche est laborieuse et il serait intéressant de posséder un outil qui puisse prendre en compte tous ces éléments simultanément, afin de pouvoir calculer la solution générale à un problème donné à partir d'un seul outil. Cet outil existe, et il se nomme *fonction de Green*. Les fonctions de Green permettent aussi bien de prendre en compte les termes de source situés dans l'équation différentielle initiale que les conditions aux limites imposées sur la frontière du domaine. C'est en cela qu'elles sont un portail vers l'expression des équations intégrales correspondantes à l'équation des ondes, ou à l'équation de Helmholtz, comme nous aurons l'occasion de le voir dans la prochaine section.

Avant d'aller plus loin dans la présentation de la théorie des fonctions de Green, il est nécessaire d'éclaircir quelques points sur le vocabulaire employé. On parlera d'équation aux dérivées partielles *homogène* lorsqu'il s'agit d'une équation sans second membre —c'est-à-dire en l'absence

de sources—, et *inhomogène* dans les autres cas. De même, on parlera de conditions de Dirichlet *homogènes* lorsque la pression acoustique doit être nulle sur les frontières du domaine, et *inhomogènes* dans tous les autres cas. Ce choix de vocabulaire n'est pas anodin, il traduit une équivalence entre les conditions aux limites et la présence de termes sources, ce que nous allons démontrer tout de suite.

En effet la fonction de Green est solution de l'équation aux dérivées partielles initiale, *homogène partout, excepté en un seul point*. Si ce point est situé sur la frontière du domaine, la fonction de Green peut être utilisée pour satisfaire des conditions aux limites *inhomogènes*, tandis que si le point est situé à l'intérieur du domaine, la fonction de Green peut être utilisée pour satisfaire une équation *inhomogène*.

En ce qui concerne les utilisations de la fonction de Green, elles permettent de calculer le champ du à une source globale, par la subdivision de cette source globale en sources élémentaires, et la sommation de chacune de ces contributions élémentaires, le tout satisfaisant certaines conditions aux limites. Tout ce qui vient d'être dit de manière générale va s'éclaircir par l'énonciation de quelques exemples sur les deux équations aux dérivées partielles mentionnées jusqu'ici : 1.19 et 1.15.

A.3.2 Exemples de fonctions de Green

A.3.2.1 Fonctions de Green pour l'équation de Kirchoff-Helmholtz

La fonction de Green pour l'équation de Kirchoff-Helmholtz sera notée de la manière suivante : $G_k(\vec{r} | \vec{r}_0)$, où k désigne le nombre d'onde de l'équation 1.19, \vec{r} désigne le point d'observation, et \vec{r}_0 désigne le point source, où la source unitaire se situe. La fonction de Green recherchée est solution de l'équation :

$$\nabla^2 G_k(\vec{r} | \vec{r}_0) + k^2 G_k(\vec{r} | \vec{r}_0) = -4\pi\delta(\vec{r} - \vec{r}_0) \quad (1.21)$$

satisfaisant des conditions aux limites homogènes de Dirichlet (ou de Neumann) sur la frontière \mathcal{S} .

Ainsi, conformément à ce qui a été dit en introduction de ce chapitre, nous souhaitons résoudre l'équation 1.19 :

$$\nabla^2 P(\vec{r}) + k^2 P(\vec{r}) = -4\pi s(\vec{r}) \quad , \forall \vec{r} \in \mathcal{V}$$

avec des conditions aux limites de Dirichlet (ou de Neumann) quelconques, en faisant appel à la fonction de Green $G_k(\vec{r} | \vec{r}_0)$. Dans cette équation, on rappelle que $s(\vec{r})$ est un terme de sources de l'équation. L'équation ci-dessus peut être réécrite pour tout $\vec{r}_0 \in \mathcal{V}$. Voici alors la forme que prend le système d'équation que l'on souhaite résoudre :

$$\begin{cases} \nabla_0^2 G_k(\vec{r}_0 | \vec{r}) + k^2 G_k(\vec{r}_0 | \vec{r}) = -4\pi\delta(\vec{r}_0 - \vec{r}) & , \forall \vec{r}_0 \in \mathcal{V}, \vec{r} \in \mathcal{V} \text{ fixé} \\ \nabla_0^2 P(\vec{r}_0) + k^2 P(\vec{r}_0) = -4\pi s(\vec{r}_0) & , \forall \vec{r}_0 \in \mathcal{V}, \vec{r} \in \mathcal{V} \text{ fixé} \end{cases} \quad (1.22)$$

L'indice 0 sous le Laplacien sert à montrer que l'opérateur agit sur la coordonnée \vec{r}_0 et pas \vec{r} . En multipliant la première équation par $P(\vec{r}_0)$ et la deuxième équation par $G_k(\vec{r}_0 | \vec{r})$, et en les soustrayant terme à terme, nous obtenons l'équation suivante :

$$G_k(\vec{r}_0 | \vec{r}) \nabla_0^2 P(\vec{r}_0) - P(\vec{r}_0) \nabla_0^2 G_k(\vec{r}_0 | \vec{r}) = 4\pi [P(\vec{r}_0) \delta(\vec{r}_0 - \vec{r}) - G_k(\vec{r}_0 | \vec{r}) s(\vec{r}_0)] \quad (1.23)$$

En n'utilisant le fait que la fonction de Green satisfait la relation de réciprocity $G_k(\vec{r}_0 | \vec{r}) = G_k(\vec{r} | \vec{r}_0)$ et en intégrant la relation précédente sur tout le volume \mathcal{V} de restitution, nous obtenons l'intégrale suivante :

$$\begin{aligned}
& \frac{1}{4\pi} \iiint_{\mathcal{V}} [G_k(\vec{r} | \vec{r}_0) \nabla_0^2 P(\vec{r}_0) - P(\vec{r}_0) \nabla_0^2 G_k(\vec{r} | \vec{r}_0)] dv_0 + \iiint_{\mathcal{V}} s(\vec{r}_0) G_k(\vec{r} | \vec{r}_0) dv_0 \\
& = \begin{cases} P(\vec{r}) & \text{si } \vec{r} \text{ est à l'intérieur de } \mathcal{V} \\ 0 & \text{si } \vec{r} \text{ est à l'extérieur de } \mathcal{V} \end{cases} \quad (1.24)
\end{aligned}$$

Il peut être intéressant de connaître la valeur que prend le membre de droite lorsque le point d'observation \vec{r} se situe sur la frontière \mathcal{S} . Il s'agit juste d'une convention, on peut aussi bien prendre la même valeur que celle à l'intérieur du volume \mathcal{V} , prendre une valeur nulle, ou bien prendre la moyenne de la valeur de $P(\vec{r})$ à l'intérieur et à l'extérieur du volume.

Nous n'avons toujours pas répondu à notre problème, qui consistait à résoudre l'équation 1.19 avec des conditions arbitraires de Dirichlet ou de Neumann sur la surface délimitant le volume de restitution \mathcal{V} . Dans l'équation précédente, nous connaissons $G_k(\vec{r} | \vec{r}_0)$, ainsi que $s(\vec{r}_0)$ sur tout le volume de restitution \mathcal{V} . En revanche nous ne connaissons pas $P(\vec{r}_0)$, ni $\nabla_0^2 P(\vec{r}_0)$ dans le volume \mathcal{V} —c'est même justement ce que nous cherchons à déterminer. Nous connaissons uniquement le comportement de $P(\vec{r}_0)$ ou sa dérivée par rapport à la normale sur la surface \mathcal{S} délimitant le volume de restitution. Pour faire intervenir ces grandeurs, il est nécessaire d'utiliser un des théorèmes démontrés à la partie A.1.4, le deuxième théorème de Green. Son application nous conduit à l'équation :

$$\begin{aligned}
& \frac{1}{4\pi} \iint_{\mathcal{S}} [G_k(\vec{r} | \vec{r}_0^s) \text{grad}_0 P(\vec{r}_0^s) - P(\vec{r}_0^s) \text{grad}_0 G_k(\vec{r} | \vec{r}_0^s)] \cdot d\vec{S}_0 + \iiint_{\mathcal{V}} s(\vec{r}_0) G_k(\vec{r} | \vec{r}_0) dv_0 \\
& = \begin{cases} P(\vec{r}) & \text{si } \vec{r} \text{ est à l'intérieur de } \mathcal{V} \\ 0 & \text{si } \vec{r} \text{ est à l'extérieur de } \mathcal{V} \end{cases} \quad (1.25)
\end{aligned}$$

L'équation précédente nous permet de résoudre l'équation de Helmholtz homogène ou inhomogène satisfaisant des conditions aux limites homogènes ou inhomogènes. Elle fait intervenir aussi bien $P(\vec{r}_0^s)$ que $\text{grad}_0 P(\vec{r}_0^s)$. Nous avons vu à la section précédente A.2 que l'unicité de la solution à l'équation de Helmholtz posait quelques problèmes. En effet, des conditions aux limites de Dirichlet sont insuffisantes pour garantir l'unicité de la solution pour des valeurs de k appartenant à l'ensemble des valeurs propres de l'équation. C'est pour cela que l'équation intégrale précédente fait appel aussi bien à la pression qu'au gradient pour garantir l'unicité de la solution. Mais, comme nous l'avons aussi mentionné à la section précédente, il faut aussi se souvenir que la pression et sa dérivée par rapport à la normale sont liées par une relation lorsque k n'appartient pas à l'ensemble des valeurs propres. Nous reviendrons sur ce problème dans la section sur les équations intégrales A.4

A.3.2.2 Fonctions de Green pour l'équation des ondes

La fonction de Green pour l'équation des ondes sera notée $G(\vec{r}, t | \vec{r}_0, t_0)$, où \vec{r} et t désignent les coordonnées spatio-temporelles du point d'observation, et où \vec{r}_0 et t_0 désignent le lieu et l'instant à laquelle l'impulsion unitaire a été émise. Ainsi, la fonction de Green recherchée est solution de l'équation suivante :

$$\begin{cases} \nabla^2 G(\vec{r}, t | \vec{r}_0, t_0) - \frac{1}{c^2} \frac{\partial^2 G(\vec{r}, t | \vec{r}_0, t_0)}{\partial t^2} = -4\pi \delta(\vec{r} - \vec{r}_0) \delta(t - t_0) \\ t_0 \text{ et } \vec{r}_0 \text{ fixés} \\ \forall t, \forall \vec{r} \in \mathcal{V} \end{cases} \quad (1.26)$$

satisfaisant les conditions aux limites homogènes de Dirichlet (ou de Neumann), et ayant pour conditions initiales $G(\vec{r}, t | \vec{r}_0, t_0) = 0$ et $\frac{\partial G(\vec{r}, t | \vec{r}_0, t_0)}{\partial t} = 0$ pour tout $t < t_0$, ce qui semble un

choix raisonnable : les effets d'une impulsion se produisant à t_0 ne peuvent pas se faire sentir à un instant antérieur.

La fonction de Green correspondante sera utile pour résoudre l'équation des ondes 1.15, que l'on réécrira manière équivalente de la façon suivante, afin de simplifier la prise en compte du terme de sources :

$$\nabla^2 p(\vec{r}, t) - \frac{1}{c^2} \frac{\partial^2 p(\vec{r}, t)}{\partial t^2} = -4\pi s(\vec{r}, t) \quad (1.27)$$

Pour l'équation des ondes, la relation de réciprocity prend la forme suivante :

$$G(\vec{r}, t | \vec{r}_0, t_0) = G(\vec{r}_0, -t_0 | \vec{r}, -t) \quad (1.28)$$

Elle spécifie qu'une impulsion émise à l'instant t_0 à la position \vec{r}_0 , et reçue à l'instant t —soit après une durée de trajet de $t - t_0$ — à une position \vec{r} est identique à la réponse obtenue à la position \vec{r}_0 à l'instant $-t_0$, émise à l'instant $-t$ de la position \vec{r} , soit toujours après une durée de trajet de $t - t_0$. C'est une application directe du principe de retour inverse des ondes, qu'elles soient lumineuses ou acoustiques.

Nous voulons maintenant démontrer, de manière analogue au paragraphe précédent, que la fonction de Green permet d'écrire de manière générale la solution de l'équation des ondes en tenant automatiquement compte des termes sources, des conditions initiales et des conditions aux limites. De la même manière que dans le paragraphe précédent, nous inversons le rôle des points sources et des points d'observations. En nous aidant de la relation de réciprocity, nous obtenons le système d'équations suivantes :

$$\begin{cases} \nabla_0^2 p(\vec{r}_0, t_0) - \frac{1}{c^2} \frac{\partial^2 p(\vec{r}_0, t_0)}{\partial t_0^2} = -4\pi s(\vec{r}_0, t_0) \\ \nabla_0^2 G(\vec{r}, t | \vec{r}_0, t_0) - \frac{1}{c^2} \frac{\partial^2 G(\vec{r}, t | \vec{r}_0, t_0)}{\partial t_0^2} = -4\pi \delta(\vec{r} - \vec{r}_0) \delta(t - t_0) \\ t \text{ et } \vec{r} \text{ fixés} \\ \forall t_0, \forall \vec{r}_0 \in \mathcal{V} \end{cases} \quad (1.29)$$

Multiplions la première équation par $G(\vec{r}, t | \vec{r}_0, t_0)$ et la deuxième équation par $p(\vec{r}_0, t_0)$ et soustrayons-les, nous obtenons le résultat suivant :

$$\begin{aligned} & G(\vec{r}, t | \vec{r}_0, t_0) \nabla_0^2 p(\vec{r}_0, t_0) - p(\vec{r}_0, t_0) \nabla_0^2 G(\vec{r}, t | \vec{r}_0, t_0) \\ & + \frac{1}{c^2} \left[p(\vec{r}_0, t_0) \frac{\partial^2 G(\vec{r}, t | \vec{r}_0, t_0)}{\partial t_0^2} - G(\vec{r}, t | \vec{r}_0, t_0) \frac{\partial^2 p(\vec{r}_0, t_0)}{\partial t_0^2} \right] \\ & = -4\pi [s(\vec{r}_0, t_0) G(\vec{r}, t | \vec{r}_0, t_0) - \delta(\vec{r} - \vec{r}_0) \delta(t - t_0) p(\vec{r}_0, t_0)] \end{aligned} \quad (1.30)$$

Intégrons maintenant cette équation selon la variable spatiale \vec{r}_0 sur tout \mathcal{V} , et selon la variable temporelle t_0 de 0 jusque t^+ (afin de prendre en compte tout le pic du à la présence d'impulsion de Dirac) :

$$\begin{aligned} & \int_0^{t^+} dt_0 \iiint_{\mathcal{V}} dv_0 \left\{ G(\vec{r}, t | \vec{r}_0, t_0) \nabla_0^2 p(\vec{r}_0, t_0) - p(\vec{r}_0, t_0) \nabla_0^2 G(\vec{r}, t | \vec{r}_0, t_0) \right\} \\ & + \int_0^{t^+} dt_0 \iiint_{\mathcal{V}} dv_0 \left\{ \frac{1}{c^2} \left(p(\vec{r}_0, t_0) \frac{\partial^2 G(\vec{r}, t | \vec{r}_0, t_0)}{\partial t_0^2} - G(\vec{r}, t | \vec{r}_0, t_0) \frac{\partial^2 p(\vec{r}_0, t_0)}{\partial t_0^2} \right) \right\} \\ & = 4\pi \left\{ p(\vec{r}, t) - \int_0^{t^+} dt_0 \iiint_{\mathcal{V}} dv_0 s(\vec{r}_0, t_0) G(\vec{r}, t | \vec{r}_0, t_0) \right\} \end{aligned} \quad (1.31)$$

De manière analogue au paragraphe précédent, appliquons maintenant le deuxième théorème de Green, et simplifions l'intégrale du deuxième terme du membre de gauche :

$$\begin{aligned}
& 4\pi \int_0^{t^+} dt_0 \iiint_{\mathcal{V}} dv_0 s(\vec{r}_0, t_0) G(\vec{r}, t | \vec{r}_0, t_0) \\
& \int_0^{t^+} dt_0 \iint_{\mathcal{S}} d\vec{S}_0 \cdot \{G(\vec{r}, t | \vec{r}_0, t_0) \text{grad}_0 p(\vec{r}_0, t_0) - p(\vec{r}_0, t_0) \text{grad}_0 G(\vec{r}, t | \vec{r}_0, t_0)\} \\
& + \frac{1}{c^2} \iiint_{\mathcal{V}} dv_0 \left[p(\vec{r}_0, t_0) \frac{\partial G(\vec{r}, t | \vec{r}_0, t_0)}{\partial t_0} - G(\vec{r}, t | \vec{r}_0, t_0) \frac{\partial p(\vec{r}_0, t_0)}{\partial t_0} \right]_{t_0=0}^{t_0=t^+} \\
& = \begin{cases} 4\pi p(\vec{r}, t) & , \text{ si } \vec{r} \text{ à l'intérieur de } \mathcal{V} \\ 0 & , \text{ si } \vec{r} \text{ à l'extérieur de } \mathcal{V} \end{cases}
\end{aligned} \tag{1.32}$$

L'intégrande du troisième terme du membre de gauche de l'équation précédente s'annule pour $t_0 = t^+$ car $\left[\frac{\partial G(\vec{r}, t | \vec{r}_0, t_0)}{\partial t_0} \right]_{t_0=t^+} = 0$ et $G(\vec{r}, t | \vec{r}_0, t^+) = 0$, ce sont les conditions initiales sur les fonctions de Green qui nous le disent —le système est causal. Ainsi, l'équation précédente se simplifie pour obtenir l'équation intégrale suivante :

$$\begin{aligned}
& 4\pi \int_0^{t^+} dt_0 \iiint_{\mathcal{V}} dv_0 s(\vec{r}_0, t_0) G(\vec{r}, t | \vec{r}_0, t_0) \\
& + \int_0^{t^+} dt_0 \iint_{\mathcal{S}} d\vec{S}_0 \cdot \{G(\vec{r}, t | \vec{r}_0, t_0) \text{grad}_0 p(\vec{r}_0, t_0) - p(\vec{r}_0, t_0) \text{grad}_0 G(\vec{r}, t | \vec{r}_0, t_0)\} \\
& - \frac{1}{c^2} \iiint_{\mathcal{V}} dv_0 \left[p(\vec{r}_0, 0) \left(\frac{\partial G(\vec{r}, t | \vec{r}_0, t_0)}{\partial t_0} \right)_{t_0=0} - (G(\vec{r}, t | \vec{r}_0, t_0))_{t_0=0} \left(\frac{\partial p(\vec{r}_0, t_0)}{\partial t_0} \right)_{t_0=0} \right] \\
& = \begin{cases} 4\pi p(\vec{r}, t) & , \text{ si } \vec{r} \text{ à l'intérieur de } \mathcal{V} \\ 0 & , \text{ si } \vec{r} \text{ à l'extérieur de } \mathcal{V} \end{cases}
\end{aligned} \tag{1.33}$$

Dans cette équation intégrale, toutes les données sont connues, excepté $\text{grad}_0 p(\vec{r}_0, t_0)$ sur la surface \mathcal{S} , car nous disposons uniquement de conditions de Dirichlet —et réciproquement pour les conditions de Neumann, on ne connaît pas $p(\vec{r}_0, t_0)$. Or, nous avons dit à la section précédente A.2 que l'équation des ondes admettait une solution unique, si les conditions initiales étaient de Cauchy, et si on disposait de conditions aux frontières de type Dirichlet ou Neumann.

Comme la fonction de Green satisfait le même type de conditions aux limites que celles satisfaites par la pression acoustique [20], nous pouvons en déduire que $G(\vec{r}, t | \vec{r}_0, t_0) = 0$ pour tout $\vec{r} \in \mathcal{S}$ s'il s'agit de l'équation initiale 1.26, ou pour tout $\vec{r}_0 \in \mathcal{S}$ si nous avons échangé le rôle des points d'observation et source, ce qui est le cas dans l'équation précédente, si nous nous rappelons les hypothèses introduites par l'équation 1.29. Ainsi, l'équation précédente se simplifie de la manière suivante finalement :

$$\begin{aligned}
& 4\pi \int_0^{t^+} dt_0 \iiint_{\mathcal{V}} dv_0 s(\vec{r}_0, t_0) G(\vec{r}, t | \vec{r}_0, t_0) \\
& - \int_0^{t^+} dt_0 \iint_{\mathcal{S}} d\vec{S}_0 \cdot (p(\vec{r}_0, t_0) \text{grad}_0 G(\vec{r}, t | \vec{r}_0, t_0)) \\
& - \frac{1}{c^2} \iiint_{\mathcal{V}} dv_0 \left[p(\vec{r}_0, 0) \left(\frac{\partial G(\vec{r}, t | \vec{r}_0, t_0)}{\partial t_0} \right)_{t_0=0} - (G(\vec{r}, t | \vec{r}_0, t_0))_{t_0=0} \left(\frac{\partial p(\vec{r}_0, t_0)}{\partial t_0} \right)_{t_0=0} \right] \\
& = \begin{cases} 4\pi p(\vec{r}, t) & , \text{ si } \vec{r} \text{ à l'intérieur de } \mathcal{V} \\ 0 & , \text{ si } \vec{r} \text{ à l'extérieur de } \mathcal{V} \end{cases}
\end{aligned} \tag{1.34}$$

L'équation intégrale démontrée précédemment permet donc de résoudre aussi bien l'équation des ondes homogène ou inhomogène, en spécifiant des conditions aux limites homogènes ou inhomogènes.

A.3.3 Conclusion sur les fonctions de Green

Les fonctions de Green permettent de donner des expressions générales de solutions d'équations aux dérivées partielles intégrant automatiquement la présence de sources et de conditions aux limites, sous forme d'équations intégrales. Elles constituent en cela un outil théorique très puissant afin de voir si les travaux expérimentaux réalisés s'appuient sur de solides bases théoriques. En revanche, elles sont très peu commodes à utiliser dans des problèmes pratiques, car il existe peu de formes analytiques à notre disposition de la fonction de Green, excepté pour des contours simples. C'est ce que nous allons essayer de développer dans la prochaine partie sur les équations intégrales. Toutefois, la perspective du recours des fonctions de Green dans la pratique n'est pas totalement réduite à néant, car celles-ci peuvent être calculées de manière approchée à l'aide des fonctions propres, qui peuvent être évaluées de manière numérique.

A.4 Equations intégrales

Dans la partie précédente, nous avons introduit les fonctions de Green, et nous avons donné quelques formulations théoriques d'équations intégrales permettant de calculer le champ de pression acoustique en tout point M intérieur au volume de restitution \mathcal{V} . Dans cette partie, nous allons essayer d'explicitier un peu plus ces équations intégrales, en explicitant les fonctions de Green pour certains contours particuliers. Comme précédemment, nous ferons ceci pour l'équation de Helmholtz, puis pour l'équation des ondes.

A.4.1 Equation intégrale de Kirchoff-Helmholtz

Dans toute cette section, nous supposons l'absence de sources dans le volume de restitution, afin d'alléger l'écriture des différentes expressions. La prise en compte des sources ne pose pas de problèmes en soi. Il s'agit juste d'un gain en lisibilité. Nous avons vu à la section précédente A.3 que la solution $P(\vec{r})$ de l'équation de Helmholtz pouvait être exprimée sous la forme de l'équation intégrale suivante :

$$\begin{aligned} & \frac{1}{4\pi} \iint_{\mathcal{S}} [G_k(\vec{r} | \vec{r}_0^s) \text{grad}_0 P(\vec{r}_0^s) - P(\vec{r}_0^s) \text{grad}_0 G_k(\vec{r} | \vec{r}_0^s)] \cdot d\vec{S}_0 \\ & = \begin{cases} P(\vec{r}) & \text{si } \vec{r} \text{ est à l'intérieur de } \mathcal{V} \\ 0 & \text{si } \vec{r} \text{ est à l'extérieur de } \mathcal{V} \end{cases} \end{aligned} \quad (1.35)$$

Dans le cas idéal, nous connaissons une fonction de Green $G_k(\vec{r} | \vec{r}_0)$, qui satisfait soit des conditions de Dirichlet — $G_k(\vec{r} | \vec{r}_0) = 0$ si $\vec{r}_0 \in \mathcal{S}$ — soit des conditions de Neumann — $\frac{\partial G_k(\vec{r} | \vec{r}_0)}{\partial n} = 0$ si $\vec{r}_0 \in \mathcal{S}$ — ce qui permet de simplifier l'équation intégrale précédente :

$$\begin{aligned} -\frac{1}{4\pi} \iint_{\mathcal{S}} P(\vec{r}_0^s) \text{grad}_0 G_k(\vec{r} | \vec{r}_0^s) \cdot d\vec{S}_0 & = \begin{cases} P(\vec{r}) & \vec{r} \text{ intérieur à } \mathcal{V} \\ 0 & \vec{r} \text{ extérieur à } \mathcal{V} \end{cases} & \text{Dirichlet} \\ \frac{1}{4\pi} \iint_{\mathcal{S}} G_k(\vec{r} | \vec{r}_0^s) \text{grad}_0 P(\vec{r}_0^s) \cdot d\vec{S}_0 & = \begin{cases} P(\vec{r}) & \vec{r} \text{ intérieur à } \mathcal{V} \\ 0 & \vec{r} \text{ extérieur à } \mathcal{V} \end{cases} & \text{Neumann} \end{aligned} \quad (1.36)$$

Malheureusement, il est très difficile dans la pratique d'aboutir à la forme analytique de fonctions de Green satisfaisant des conditions aux limites de type Dirichlet ou Neumann pour

des surfaces \mathcal{S} délimitant des volumes de restitution quelconques. En pratique, nous utilisons toujours des surfaces simples —espace libre, sphère, cylindre infini— qui permettent d’aboutir à des expressions analytiques de la fonction de Green. Malheureusement, elles ne vérifient plus les conditions aux limites pour un contour quelconque, d’où l’utilité d’avoir gardé une expression générale pour l’équation intégrale de Kirchoff-Helmholtz 1.35.

A.4.1.1 Fonction de Green en espace libre

La fonction de Green en espace libre satisfait la condition d’annulation à l’infini. Elle est de la forme suivante :

$$G_k(\vec{r} | \vec{r}_0) = \frac{e^{jk\|\vec{r}-\vec{r}_0\|}}{\|\vec{r}-\vec{r}_0\|} \quad (1.37)$$

Si nous injectons cette expression de la fonction de Green dans l’équation 1.35, nous obtenons le résultat suivant :

$$\frac{1}{4\pi} \iint_{\mathcal{S}} \left[\frac{e^{jk\|\vec{r}-\vec{r}_0\|}}{\|\vec{r}-\vec{r}_0\|} \text{grad}_0 P(\vec{r}_0) - P(\vec{r}_0) \text{grad}_0 \frac{e^{jk\|\vec{r}-\vec{r}_0\|}}{\|\vec{r}-\vec{r}_0\|} \right] \cdot d\vec{S}_0 = \begin{cases} P(\vec{r}) & \vec{r} \text{ intérieur à } \mathcal{V} \\ 0 & \vec{r} \text{ extérieur à } \mathcal{V} \end{cases}$$

Si nous utilisons la fonction de Green en espace libre pour expliciter l’intégrale de Kirchoff-Helmholtz, il est nécessaire de connaître à la fois la pression, mais aussi sa dérivée par rapport à la normale sur la surface \mathcal{S} pour pouvoir extrapoler le champ de pression acoustique en tout point du volume de restitution \mathcal{V} . Nous n’aurions uniquement besoin de la pression comme condition aux limites si le volume de restitution \mathcal{V} était identique au volume dans lequel la fonction de Green était calculée.

A.4.1.2 Intégrales de Rayleigh [1]

Il doit être possible, sous certaines conditions, de s’affranchir de l’une de ces deux conditions aux limites afin d’obtenir des expressions qui ne font intervenir que la pression $P(\vec{r}_0)$ ou son gradient dans le processus d’extrapolation. C’est le cas si nous considérons une surface composée d’une demi-sphère de rayon infini Γ , et d’un plan infini Π , tel qu’en témoigne la figure suivante. Dans ce cas il est possible de trouver une fonction de Green satisfaisant soit des conditions de Dirichlet, soit des conditions de Neumann sur la surface \mathcal{S} .

Nous cherchons une fonction de Green qui satisfait à la fois l’équation de Helmholtz, mais aussi des conditions de Dirichlet homogènes sur la surface $\mathcal{S} = \Pi \cup \Gamma$. Si nous considérons la fonction suivante :

$$G_k(\vec{r} | \vec{r}_0) = \frac{e^{jk\|\vec{R}\|}}{\|\vec{R}\|} - \frac{e^{jk\|\vec{R}'\|}}{\|\vec{R}'\|} \quad (1.38)$$

où $\vec{R} = \vec{r} - \vec{r}_0$ et où \vec{R}' désigne le symétrique de \vec{R} par rapport au plan Π , nous pouvons vérifier qu’elle est bien solution de l’équation de Helmholtz, et qu’elle s’annule lorsque $\vec{r} \in \Pi, \forall \vec{r}_0 \in \mathcal{S}$, ce qui est le cas représenté sur la figure —en effet, dans ce cas nous avons $\vec{R} = \vec{R}'$.

Nous pouvons donc utiliser la première formule de l’équation 1.36 concernant les conditions aux limites de Dirichlet, en inversant son signe car le vecteur $d\vec{S}_0 = ds_0 \vec{n}$ est orienté vers l’intérieur cette fois-ci :

$$\frac{1}{4\pi} \iint_{\mathcal{S}=\Pi \cup \Gamma} P(\vec{r}_0^*) \text{grad}_0 G_k(\vec{r} | \vec{r}_0^*) \cdot d\vec{S}_0 = P(\vec{r}) \quad (1.39)$$

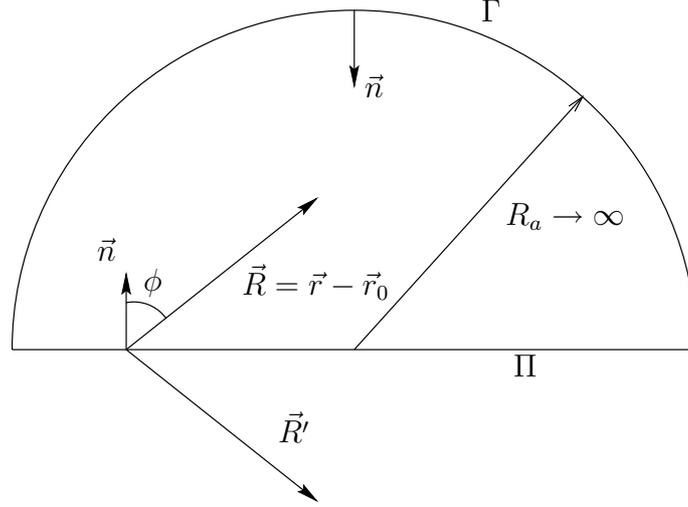


FIG. 1.4 – Intégrales de Rayleigh

Cette intégrale est nulle sur Γ où l'on a $\text{grad}_0 G_k(\vec{r} | \vec{r}_0^s) = \vec{0}$. Il ne reste donc plus que l'intégrale sur le plan Π . Pour l'expliciter, calculons $\text{grad}_0 G_k(\vec{r} | \vec{r}_0^s) \cdot d\vec{S}_0$, pour tout \vec{r}_0^s de Π :

$$\begin{aligned}
 \text{grad}_0 G_k(\vec{r} | \vec{r}_0^s) \cdot d\vec{S}_0 &= \frac{\partial G_k(\vec{r} | \vec{r}_0^s)}{\partial n} dS_0 \\
 &= \frac{\partial}{\partial n} \left(\frac{e^{jk\|\vec{R}\|}}{\|\vec{R}\|} - \frac{e^{jk\|\vec{R}'\|}}{\|\vec{R}'\|} \right) dS_0 \\
 &= \frac{1+jkR}{R^2} \left(\frac{\partial R}{\partial n} - \frac{\partial R'}{\partial n} \right) e^{-jkR} dS_0 \\
 &= 2 \frac{1+jkR}{R^2} \cos \phi e^{-jkR} dS_0
 \end{aligned} \tag{1.40}$$

Finalement, en réintroduisant cette expression, dans l'équation intégrale, nous obtenons la *deuxième intégrale de Rayleigh* :

$$\frac{1}{2\pi} \iint_{\Pi} P(\vec{r}_0^s) \frac{1+jk\|\vec{r}-\vec{r}_0^s\|}{\|\vec{r}-\vec{r}_0^s\|^2} \cos \phi e^{-jk\|\vec{r}-\vec{r}_0^s\|} dS_0 = P(\vec{r}) \tag{1.41}$$

Pour obtenir la première intégrale de Rayleigh, il suffit de chercher la fonction de Green qui satisfait des conditions De Neumann homogènes sur la surface $\mathcal{S} = \Pi \cup \Gamma$. La fonction de Green correspondante est :

$$G_k(\vec{r} | \vec{r}_0) = \frac{e^{jk\|\vec{R}\|}}{\|\vec{R}\|} + \frac{e^{jk\|\vec{R}'\|}}{\|\vec{R}'\|} \tag{1.42}$$

avec les mêmes notations utilisées auparavant. Tout calcul effectué, en considérant toujours que la normale pointe vers l'intérieur, l'expression de la *première intégrale de Rayleigh* est :

$$\frac{1}{2\pi} \iint_{\Pi} \frac{e^{jk\|\vec{r}-\vec{r}_0^s\|}}{\|\vec{r}-\vec{r}_0^s\|} \text{grad}_0 P(\vec{r}_0^s) \cdot d\vec{S}_0 = P(\vec{r}) \tag{1.43}$$

remarque : Nous avons dit à la section A.2 qu'il existait des problèmes d'unicité de la solution à l'équation de Helmholtz selon les conditions aux limites utilisées. En effet, nous avons dit que des conditions de Dirichlet ou de Neumann étaient suffisantes pour garantir

l'unicité en dehors des valeurs propres de k correspondantes au domaine considéré. Aux modes propres, il était nécessaire d'avoir des conditions de Cauchy pour garantir l'unicité de la solution. Dans les expressions des deux différentes intégrales de Rayleigh, elles ne font intervenir que la connaissance de la pression, ou celle de son gradient. Cela n'est pas en contradiction avec ce qui a été dit précédemment, car le volume de restitution considéré est infini, et ne possède donc pas de modes de résonance pour la pression, ce qui explique que des conditions de Dirichlet ou de Neumann sont suffisantes pour extrapoler le champ dans toute la zone.

A.4.1.3 Equations intégrales pour un cylindre infini, une sphère, et contours arbitraires

Il serait idéal d'avoir une formulation simple et analytique de la fonction de Green satisfaisant à la fois l'équation de Helmholtz, mais aussi des conditions aux limites de Dirichlet, ou de Neumann, sur des contours particuliers, tels qu'un cylindre ou bien une sphère, ou des contours quelconques. Malheureusement, cela n'est pas possible. Toutefois, le constat n'est pas si pessimiste qu'il en a l'air, et des solutions existent pour extrapoler la pression acoustique à l'intérieur d'un cylindre, d'une sphère, ou de tout contour quelconque, à partir de la connaissance de la pression acoustique sur la surface : il s'agit de toutes les techniques de l'holographie acoustique. Pour des contours cylindriques, il est possible, à partir de la connaissance de la pression sur la surface, de reconstruire le champ à l'intérieur de celui-ci en décomposant cet hologramme sur la base des fonctions propres du domaine, c'est-à-dire, soit les harmoniques cylindriques, soit les harmoniques sphériques. Pour ces deux contours particuliers, il existe des expressions analytiques sous forme de séries du champ reconstitué à l'intérieur. Pour des contours quelconques, des fonctions propres peuvent être évaluées de manière numérique et la même méthodologie peut être employée pour reconstruire le champ à l'intérieur du volume de restitution. En effet, pour des contours quelconques, la fonction de Green ne peut pas être explicitée directement, mais elle peut s'exprimer sous forme d'une série dans laquelle interviennent les fonctions propres du domaine. Tout lecteur intéressé par le sujet pourra se reporter aux publications [20][21][22][17][18][16] pour obtenir plus de détails.

A.4.2 Equation intégrale de Kirchoff

Dans toute cette section, nous supposons aussi l'absence de sources sonores dans le volume de restitution, afin d'alléger un peu l'écriture des équations. Nous avons vu à la section précédente A.3 que la solution $p(\vec{r}, t)$ de l'équation des ondes pouvait s'exprimer de manière générale par l'équation intégrale suivante :

$$\begin{aligned}
& + \int_0^{t^+} dt_0 \iint_S d\vec{S}_0 \cdot \{ G(\vec{r}, t | \vec{r}_0, t_0) \text{grad}_0 p(\vec{r}_0, t_0) - p(\vec{r}_0, t_0) \text{grad}_0 G(\vec{r}, t | \vec{r}_0, t_0) \} \\
& - \frac{1}{c^2} \iiint_{\mathcal{V}} dv_0 \left[p(\vec{r}_0, 0) \left(\frac{\partial G(\vec{r}, t | \vec{r}_0, t_0)}{\partial t_0} \right)_{t_0=0} - (G(\vec{r}, t | \vec{r}_0, t_0))_{t_0=0} \left(\frac{\partial p(\vec{r}_0, t_0)}{\partial t_0} \right)_{t_0=0} \right] \quad (1.44) \\
& = \begin{cases} 4\pi p(\vec{r}, t) & , \text{ si } \vec{r} \text{ à l'intérieur de } \mathcal{V} \\ 0 & , \text{ si } \vec{r} \text{ à l'extérieur de } \mathcal{V} \end{cases}
\end{aligned}$$

Dans le cas idéal, la fonction de Green $G(\vec{r}, t | \vec{r}_0, t_0)$ satisfait des conditions initiales de Cauchy, et des conditions aux limites de Dirichlet sur les frontières du domaine. L'équation précédente devient alors :

$$\begin{aligned}
& - \int_0^{t^+} dt_0 \iint_{\mathcal{S}} d\vec{S}_0 \cdot (p(\vec{r}_0, t_0) \text{grad}_0 G(\vec{r}, t | \vec{r}_0, t_0)) \\
& - \frac{1}{c^2} \iiint_{\mathcal{V}} dv_0 \left[p(\vec{r}_0, 0) \left(\frac{\partial G(\vec{r}, t | \vec{r}_0, t_0)}{\partial t_0} \right)_{t_0=0} - (G(\vec{r}, t | \vec{r}_0, t_0))_{t_0=0} \left(\frac{\partial p(\vec{r}_0, t_0)}{\partial t_0} \right)_{t_0=0} \right] \\
& = \begin{cases} 4\pi p(\vec{r}, t) & , \text{ si } \vec{r} \text{ à l'intérieur de } \mathcal{V} \\ 0 & , \text{ si } \vec{r} \text{ à l'extérieur de } \mathcal{V} \end{cases} \quad (1.45)
\end{aligned}$$

Malheureusement, il est très difficile d'aboutir à une forme analytique simple de ces fonctions de Green pour des contours quelconques. Ce cas est encore plus compliqué que le cas précédent. On ne détaillera que la fonction de Green en espace libre.

La fonction de Green en espace libre est solution de l'équation des ondes 1.26 et satisfait des conditions initiales de Cauchy homogènes $-G(\vec{r}, t | \vec{r}_0, t_0) = 0$ et $\frac{\partial G(\vec{r}, t | \vec{r}_0, t_0)}{\partial t} = 0$ pour tout $t < t_0$ — ainsi que des conditions homogènes sur les frontières du domaine $-G(\vec{r}, t | \vec{r}_0, t_0) = 0$ pour tout $\vec{r} \in \mathcal{S}$ — situées à l'infini. La fonction remplissant ces conditions est :

$$G(\vec{r}, t | \vec{r}_0, t_0) = \frac{\delta\left((t - t_0) - \frac{\|\vec{r} - \vec{r}_0\|}{c}\right)}{\|\vec{r} - \vec{r}_0\|} \quad (1.46)$$

Nous ne détaillerons pas la forme de l'équation intégrale de Kirchoff prise lorsque nous choisissons cette fonction de Green pour tout contour quelconque \mathcal{S} comme nous l'avons fait dans le cas précédent pour l'équation de Helmholtz. Cette équation représente à ce jour plus un intérêt théorique que pratique.

A.5 Conclusion

Beaucoup de choses ont été dites dans cette partie théorique, et il peut être nécessaire d'en résumer les points importants, qui serviront pour la suite des événements. Nous avons vu que la pression acoustique $p(\vec{r}, t)$ est solution de l'équation des ondes, accompagnée des conditions aux limites et initiales suivantes :

$$\left\{ \begin{array}{l} \nabla^2 p(\vec{r}, t) - \frac{1}{c^2} \frac{\partial^2 p(\vec{r}, t)}{\partial t^2} = -4\pi s(\vec{r}, t) \\ p(\vec{r}, t)|_{t=0} = f(\vec{r}) \\ \left(\frac{\partial p(\vec{r}, t)}{\partial t} \right)_{t=0} = g(\vec{r}) \\ p(\vec{r}, t) \text{ connue } \forall \vec{r} \in \mathcal{S} \end{array} \right. \quad (1.47)$$

Le problème ainsi posé possède une solution unique. Cela a deux conséquences très importantes :

- Si nous parvenons à imposer les mêmes conditions aux limites qu'une source sonore initiale par un dispositif auxiliaire, nous sommes assurés de l'identité des deux champs reproduits, à l'intérieur du volume de restitution.
- Seule la connaissance de la pression sur la surface est nécessaire pour enregistrer toute l'information spatio-temporelle du champ sonore.

En revanche, nous avons vu que l'équation des ondes était rarement utilisée dans le domaine de l'holographie acoustique, qui consiste à extrapoler le champ acoustique dans une zone tridimensionnelle à partir de sa connaissance sur une surface uniquement, et que nous préférons lui substituer l'équation de Helmholtz à la place, qui se prête mieux aux traitements numériques. Malheureusement, nous avons vu que la résolution de l'équation de Helmholtz était délicate, à cause du problème des conditions aux limites à imposer sur le contour délimitant la zone de restitution. Cette discipline est toujours en cours de développement actuellement, afin d'améliorer ses performances.

Ainsi, plusieurs stratégies sont possibles pour aborder la thématique de la restitution du champ sonore sur une zone étendue de l'espace :

- Nous pouvons tout d'abord vouloir avancer sur deux plans de cette thématique, celui de la reconstruction réelle du champ sonore par un réseau de transducteurs, et celui de la simulation numérique de champ sonore. Le recours à la simulation peut être fort utile, aussi bien pour la mise en place d'algorithmes permettant d'achever notre but, que pour comparer les différences entre le champ initial de la source et celui reconstruit par le réseau de transducteurs. Si nous choisissons cette problématique, il sera alors nécessaire de connaître aussi bien la pression acoustique, que la dérivée de la pression acoustique par rapport à la normale à la surface délimitant le volume de restitution. L'autre alternative est d'essayer de développer des outils directement sur l'équation des ondes, et non sur l'équation de Helmholtz.
- Nous pouvons tout aussi bien vouloir travailler en aveugle. En effet, nous savons que la connaissance de la pression acoustique sur la frontière du domaine est suffisante pour caractériser l'identité d'un champ sonore, il suffit d'essayer de reconstruire au mieux ce champ, afin qu'il soit le plus proche possible du champ initial. Les outils à notre service pour y parvenir sont alors la théorie des systèmes, la théorie des problèmes inverses, la théorie du contrôle des systèmes, etc. Cette étude se base alors sur le fait que le problème auquel nous sommes confrontés est bien-posé, c'est-à-dire que le champ reconstruit sera d'autant plus proche du champ initial que la pression contrôlée par le réseau de transducteurs sera proche de la consigne.

Nous allons maintenant nous intéresser plus en détail aux problématiques particulières liées à l'enregistrement du champ sonore sur une zone étendue, et la restitution du champ sonore sur une zone étendue.

B Enregistrement du champ sonore

Nous ne pouvons pas affirmer que les techniques de prise de son actuelles sont des méthodes d'enregistrement du champ sonore. En effet, les seules techniques utilisées en prise de son sont des techniques ayant recours à un seul microphone pour enregistrer une source donnée, s'il s'agit d'une prise de son monophonique, ou deux microphones tout au plus s'il s'agit d'une prise de son stéréophonique. Certes, le but de ce document n'est pas de remettre en cause ce qui a mis du temps à s'établir, et les méthodes de prise de son actuelles n'ont plus à faire leur preuve. Nous avons su substituer un côté artistique à la prise de son, là où la physique est restée en échec jusqu'à présent. Le but de ce document n'est pas non plus de faire un cours sur la prise de son, c'est pourquoi nous n'irons pas plus loin dans cette voie. Mais, pour en revenir à ce qui était annoncé un peu plus tôt, nous ne pouvons nier le fait que les enregistrements actuels n'enregistrent pas la scène sonore dans sa totalité : ils enregistrent seulement un point de vue, celui de l'ingénieur du son, qui a disposé son matériel de sorte à obtenir la 'meilleure séquence sonore' que pouvait lui offrir la scène. L'enregistrement d'un champ sonore contient plus que cela : il doit aussi bien enregistrer le déroulement temporel de la scène sonore, mais il doit aussi

être capable d’enregistrer son organisation spatiale, car l’organisation spatiale du champ sonore reflète les caractéristiques des lieux d’écoutes dans lesquels les sources acoustiques évoluent. Actuellement, nous ne faisons que privilégier le point de vue d’un seul observateur de la scène, en plaçant un microphone ou un couple stéréophonique. Tout au plus sommes-nous capables de mélanger les points de vue de plusieurs auditeurs en faisant une balance adéquate entre un couple stéréophonique principal et des microphones d’appoint.

Le domaine de la restitution du champ sonore est inévitablement couplé à celui de l’enregistrement du champ sonore, c’est pourquoi si nous décidons de faire évoluer les techniques de restitution du champ sonore afin qu’elles soient plus fidèles au champ initial, il est nécessaire de développer les outils correspondants dans le domaine de l’enregistrement du champ sonore, sinon les performances d’un tel système seraient très vite limitées : imaginez, ce serait aussi blasphématoire que de placer des enregistrements monophoniques dans un univers stéréophonique —certes, c’est comme ça que sont produits la majorité des disques des médias les plus importants— si nous voulons faire une comparaison avec les moyens actuels. Dans cette partie nous essayons d’envisager des nouvelles méthodes pour l’enregistrement du champ sonore, qui seront associées avec notre futur dispositif de restitution de champ sonore sur une zone étendue de l’espace. Ce dispositif d’enregistrement fera appel principalement à la technique des réseaux de microphones. Dans une première partie, nous étudions en quoi le recours aux réseaux de microphones permet d’avoir une meilleure compréhension de l’organisation spatiale du champ sonore. Puis, nous essaierons de faire le lien avec la partie théorique précédente, afin de voir quelles sont les différentes possibilités concernant le développement des réseaux de microphones. Dans cette même partie, nous mettrons aussi en évidence les défauts d’un tel système par rapport au système idéal, et notamment le défaut principal de ces réseaux, qui est de discrétiser l’enregistrement du champ sonore, et nous verrons quelles sont les conditions nécessaires afin d’enregistrer toute l’information contenue dans le champ sonore initial. Dans une dernière partie, nous évoquerons les différentes techniques de prise de son associées avec ces dispositifs de mesure, et nous verrons qu’ils n’enlèveront pas forcément la dimension artistique de la discipline, bien au contraire.

B.1 Préliminaires

Dans cette section, nous démontrons que le recours à un réseau de microphones lors de l’enregistrement de la scène sonore nous apporte une meilleure image de l’organisation spatiale du champ sonore que celle fournie en ayant recours à des méthodes de prise de sons classiques [11]. En effet, lors d’une séance d’enregistrement, les professionnels s’accordent à dire que le choix du lieu de l’enregistrement est déterminant pour la qualité finale de l’enregistrement. Ceci est dû au fait que l’acoustique des différents lieux d’écoute joue un rôle perceptif très important sur la qualité de l’enregistrement. L’outil physique le plus largement utilisé dans le domaine de l’acoustique des salles est la réponse impulsionnelle. A partir de cette réponse impulsionnelle sont dérivés des critères scientifiques de plus haut niveau (tels que le temps de réverbération, le taux de premières réflexions, etc.), censés correspondre plus ou moins à des critères perceptifs de la qualité acoustique du lieu. D’un point de vue perceptif toujours, l’opinion générale est que la structure précise de ces réponses impulsionnelles est fortement redondante lorsque nous nous déplaçons dans la salle [11]. Pourtant, il existe des différences significatives sur l’allure temporelle et fréquentielle des signaux recueillis pour différentes positions. C’est pourquoi les méthodes d’analyse employées pour traiter ces réponses sont d’ordre statistiques. Nous allons voir que le recours à un réseau de microphones pour l’étude de l’acoustique d’une salle permet de caractériser l’organisation spatiale des événements sonores, ce que nous ne parvenons pas à faire en plaçant les microphones de manière arbitraire et subjective [11].

Dans toute la suite du paragraphe, les figures sont tirées de [11]. La figure 1.5 montre des réponses impulsionnelles relevées en plusieurs positions différentes, situées sur un même axe, et espacées de 50 cm chacune. Si nous examinons seulement la première et la dernière réponse

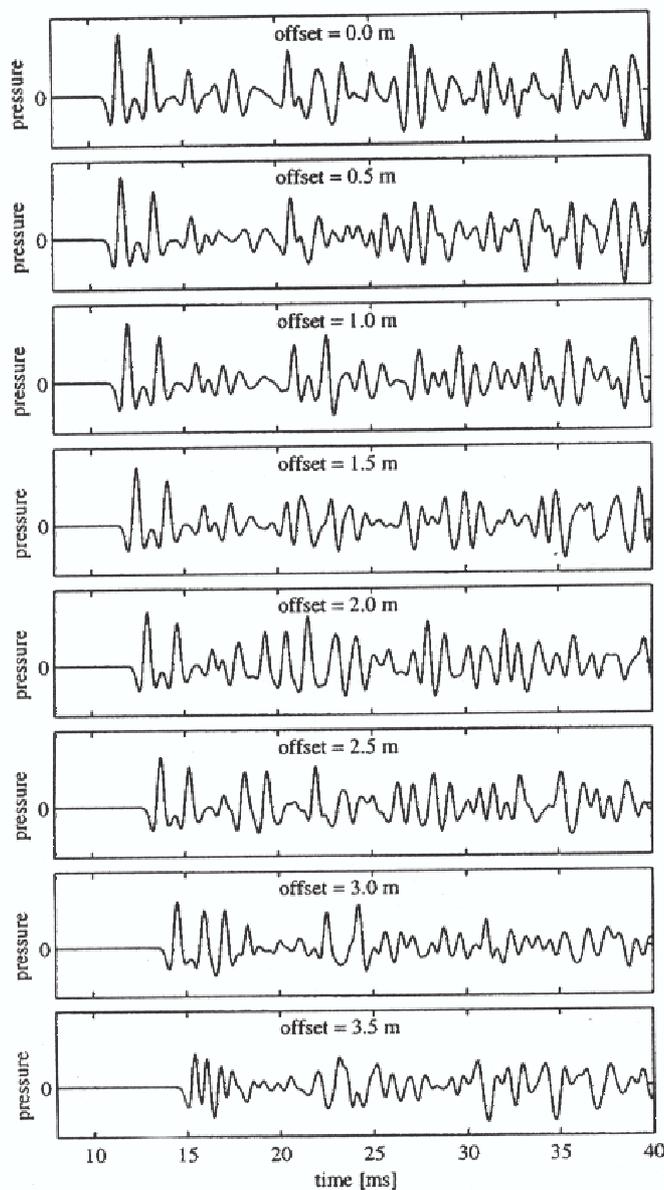


FIG. 1.5 – Séries de réponses impulsionnelles mesurée par huit microphones distant de 50 cm, pour une position donnée de la source

impulsionnelle sur cette figure, nous avons du mal à croire que la même source est à l'origine du phénomène. Toutefois, nous voyons qu'il existe quelques ressemblances entre ces différents tracés, mais que des différences notables apparaissent lorsque les microphones sont de plus en plus éloignés. Ces représentations temporelles successives rendent difficilement compte de l'organisation spatiale du champ. Une représentation mieux adaptée, consiste à tracer ceci en trois dimensions, la première étant la dimension spatiale, la deuxième étant le temps, et la troisième étant l'amplitude de l'onde relevée par le microphone (voir figure 1.6). Nous observons tout de suite qu'il existe une forte corrélation entre les différentes réponses relevées.

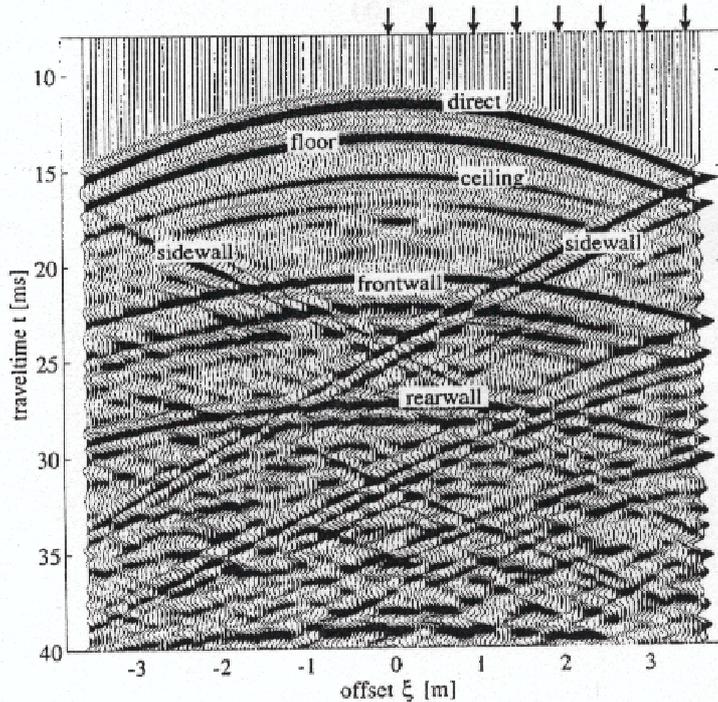


FIG. 1.6 – Représentation tridimensionnelle du champ sonore dans un espace clos enregistrée par un réseau de microphones rapprochés. Les réponses impulsionnelles indiquées par des flèches sont celles affichées sur la figure 1.5

L'analyse de cette réponse nous permet d'identifier les principales formes géométriques présentes sur cette figure. La première courbe représente le front d'onde dû au son direct. Le deuxième front d'onde qui parvient juste après est dû à la réflexion sur le sol. Le suivant est dû à la réflexion sur le plafond de la salle. De plus, nous parvenons clairement à identifier les réflexions sur les parois latérales. En effet, tout se passe comme si le front d'onde se repliait à l'endroit où se situent les parois. En conclusion, nous voyons bien que le recours à un réseau de microphones lors de l'enregistrement nous permet d'avoir une meilleure vision de l'espace sonore. Cette représentation a donné naissance à une nouvelle technique utilisée pour l'étude de l'acoustique des lieux d'écoute, baptisée Wave Field Analysis (WFA) par son créateur Mr Berkhout, officiant à l'université des technologies de Delft [13]. Pour les simulations de champ sonore, il peut être utile de développer des logiciels d'acoustique prédictive, ce qui serait très utile pour l'étude de la restitution de champ sonore en lieu fermé. Plusieurs approches ont déjà été proposées dans [23], [24] et [25]. Néanmoins il ne s'agissait ici que d'une présentation des réseaux de microphones. Maintenant, il nous faut décrire les contraintes à satisfaire permettant d'aboutir à des systèmes d'enregistrement utilisant des réseaux de microphones qui soient capables d'enregistrer toute l'information spatio-temporelle d'une scène sonore pour pouvoir ensuite la restituer sur un système de reproduction adapté.

B.2 Des différentes méthodologies pour l'enregistrement du champ sonore

Nous avons vu que le recours à des réseaux de microphones permet d'avoir une meilleure vision de l'organisation spatiale du champ sonore. Nous allons maintenant essayer de décrire les caractéristiques d'un dispositif permettant d'enregistrer le champ sonore dans sa globalité, et nous ferons largement appel aux résultats théoriques développées dans la partie A. Une méthode

naïve pour enregistrer toute l'information spatio-temporelle du champ sonore serait de disposer notre réseau de capteurs de manière homogène dans tout le volume où l'on souhaite connaître le champ. Cette méthode n'est pas envisageable sur un plan pratique, car elle mettrait en jeu un nombre trop important de capteurs. La solution pour réduire le nombre de capteurs à utiliser est donnée par la partie théorique, dont le résultat majeur est que le comportement du champ à l'intérieur d'un volume peut être déduit de son comportement sur la surface délimitant ce volume, donc nous n'aurons qu'à placer le réseau de capteurs de pression uniquement sur une surface.

B.2.1 Les conditions idéales

B.2.1.1 Méthodologie dérivée de l'équation intégrale de Kirchoff-Helmholtz

A titre de rappel, nous rappelons l'expression de l'intégrale de Kirchoff-Helmholtz, démontrée dans le cas où la fonction de Green utilisée est celle en espace libre :

$$\frac{1}{4\pi} \iint_{\mathcal{S}} \left[\frac{e^{jk\|\vec{r}-\vec{r}_0\|}}{\|\vec{r}-\vec{r}_0\|} \text{grad}_0 P(\vec{r}_0) - P(\vec{r}_0) \text{grad}_0 \frac{e^{jk\|\vec{r}-\vec{r}_0\|}}{\|\vec{r}-\vec{r}_0\|} \right] \cdot d\vec{S}_0 = \begin{cases} P(\vec{r}) & \vec{r} \text{ intérieur à } \mathcal{V} \\ 0 & \vec{r} \text{ extérieur à } \mathcal{V} \end{cases}$$

Cette équation nous informe que la mesure de la pression, et du gradient de pression, en tout point de la surface \mathcal{S} délimitant le volume \mathcal{V} est suffisante pour connaître toute l'organisation spatiale du champ sonore. Nous avons déjà beaucoup discuté des conditions aux limites à utiliser pour garantir l'existence et l'unicité de la solution de l'équation de Helmholtz. Nous avons notamment dit que des conditions de Cauchy imposaient trop de contraintes, et que l'existence de la solution n'était plus forcément assurée. En effet, il existe naturellement une relation entre la pression et son gradient en dehors des modes de résonance, alors que les deux grandeurs évoluent de manière indépendante pour un mode de résonance —en pratique $P(\vec{r}) = 0$ et $\text{grad} P(\vec{r}) \cdot d\vec{S} \neq 0$. Nous pouvons donc en déduire que, si notre système de mesure est parfait, c'est-à-dire, entre autres, qu'il n'est pas sujet au bruit, alors le recours à des conditions de Cauchy est envisageable. En pratique, tout système de mesure est entaché de bruit, il faut donc étudier les conséquences de ce défaut sur notre système d'extrapolation. En théorie, pour les modes propres du domaine, la pression s'annule en tout point de la surface \mathcal{S} , et c'est le gradient qui impose l'unicité de la solution pour ces valeurs particulières de k . Notre système de mesure nous donnera forcément une pression différente de 0 pour ces valeurs de k , mais cela n'empêche pas l'évaluation de l'intégrale ci-dessus. Le deuxième terme correspondra alors à un terme d'erreur dû à la présence de bruit dans le système de mesure.

L'avantage de ce dispositif mesurant à la fois la pression et le gradient de pression est qu'il est possible *a posteriori* de changer le volume \mathcal{V} où nous souhaitons connaître le champ. L'intérêt est double :

- Les transducteurs qui serviront au moment de la restitution du champ ne se situent pas forcément sur la même surface \mathcal{S} que les capteurs qui ont servi lors de l'enregistrement. Une étape d'extrapolation sera nécessaire au préalable pour connaître le champ de pression acoustique au niveau des transducteurs, notamment si les techniques utilisées lors de la restitution sont basées sur la Wave Field Synthesis, que nous décrirons plus en détail dans la partie C sur la restitution du champ sonore sur une zone étendue de l'espace.
- Il peut être utile de vouloir réduire la connaissance de l'organisation spatiale du champ à un volume inférieur par rapport au volume d'enregistrement. Ce sera utile notamment lorsque les volumes d'enregistrement seront grands, alors que les volumes de restitution seront petits, ce qui sera le cas chez l'utilisateur lorsqu'il utilisera le dispositif de restitution à la maison. Dans ce cas, un processus d'extrapolation sera nécessaire pour connaître le

champ sur la surface \mathcal{S}' délimitant le volume de restitution \mathcal{V}' , à partir de la connaissance du champ sur la surface \mathcal{S} utilisée au moment de l'enregistrement.

L'inconvénient de ce dispositif est qu'il faille utiliser un système de mesure capable de mesurer simultanément et la pression, et la composante normale à la surface du gradient de pression.

B.2.1.2 Méthodologie dérivée de l'équation intégrale des ondes

Nous avons vu que le problème suivant :

$$\left\{ \begin{array}{l} \nabla^2 p(\vec{r}, t) - \frac{1}{c^2} \frac{\partial^2 p(\vec{r}, t)}{\partial t^2} = -4\pi s(\vec{r}, t) \\ p(\vec{r}, t)|_{t=0} = f(\vec{r}) \\ \left(\frac{\partial p(\vec{r}, t)}{\partial t} \right) |_{t=0} = g(\vec{r}) \\ p(\vec{r}, t) \text{ connue } \forall \vec{r} \in \mathcal{S} \end{array} \right.$$

admettait une solution unique pour la pression acoustique. Il suffit donc d'enregistrer la pression acoustique sur une surface fermée \mathcal{S} pour pouvoir caractériser de manière unique le champ sonore. L'avantage d'une telle méthode de mesure est qu'elle ne nécessite pas le recours à la mesure du gradient de pression. L'inconvénient majeur est que nous n'aurons pas accès à tous les outils de l'holographie acoustique permettant d'extrapoler le champ à l'intérieur, et cela inhibera les manipulations que nous aurions pu faire afin d'obtenir un système flexible au niveau de l'utilisation. La principale conséquence de cette rigidité est que le volume d'enregistrement sera nécessairement le volume de restitution.

Nous avons démontré à la section A.4 que la pression acoustique pouvait être calculée uniquement par la connaissance des conditions initiales, des conditions aux frontières et de la fonction de Green. Nous pouvons donc penser que tout espoir n'est pas perdu, et qu'il est aussi possible d'extrapoler le champ par ce dispositif de mesure. Néanmoins cela nécessite la connaissance des conditions initiales, qui requerraient d'enregistrer la pression sur tout le volume, ce qui n'est pas possible. Nous pouvons toujours faire l'hypothèse que ces conditions initiales sont nulles, au bruit de fond près, mais il faudrait tout de même connaître la fonction de Green qui satisfait des conditions initiales homogènes, et des conditions de Dirichlet homogènes pour la surface \mathcal{S} considérée, ce qui posera les mêmes difficultés que pour les fonctions de Green pour l'équation de Helmholtz pour des contours quelconques. Dans ce cas, l'extrapolation du champ de pression à l'intérieur du volume ne ferait plus appel qu'à la connaissance de la pression acoustique en tout point de la surface \mathcal{S} . Cette piste de recherche n'a pas encore été explorée... Elle méritera peut-être qu'on lui accorde quelque considération.

B.2.2 Les conditions réelles et les limites associées

Nous venons d'envisager les deux dispositifs possibles pour enregistrer parfaitement l'organisation spatiale et temporelle du champ sonore, et nous en avons décrit les limites théoriques. Mais d'autres causes viennent limiter les performances de notre système dans la pratique, la principale étant qu'il est impossible de mesurer de manière continue la pression et/ou le gradient de pression sur une surface fermée. Une étape nécessaire du processus sera donc de discrétiser le problème. De plus, les capteurs utilisés possèdent des limitations qui font que le système de mesure discrétisé ne sera pas idéal non plus. La dernière limitation vient de la présence éventuelle de sources sonores parasites dans le volume d'enregistrement du champ. Nous allons revenir successivement sur chacune de ses limitations.

B.2.2.1 Limites dues à la discrétisation

Les seuls capteurs disponibles pour mesurer la pression ou le gradient de pression sont les microphones, et ils ne sont capables d'effectuer cette mesure qu'en un point de l'espace. Une étape déterminante de l'enregistrement d'une scène sonore sera la discrétisation des différentes équations intégrales mentionnées jusqu'ici, de telle sorte qu'elles n'altèrent pas les performances du système de manière trop importante. Comme tout système échantillonné, des limites vont apparaître. La première est la résolution de notre système d'enregistrement sonore vis-à-vis des différentes longueurs d'onde que nous pouvons détecter. Cette première limite affecte tous les réseaux de microphones. Il existe une deuxième limitation, mais qui ne concerne que le système d'enregistrement utilisant des réseaux linéaires ou plans (basés sur les intégrales de Rayleigh), car en pratique le réseau n'est jamais infini, il va donc y avoir de la diffraction sur les bords, ayant des effets analogues au phénomène de Gibbs, qui survient dès lors que l'horizon d'observation est temporellement, ou ici spatialement limité.

Pour mettre en évidence le type de problèmes auxquels nous allons être confrontés lors du processus de discrétisation, nous allons étudier un cas simple, celui du réseau de microphones linéaire, c'est-à-dire dont les microphones sont régulièrement espacés. Il correspond à un échantillonnage régulier et les mêmes limitations que pour les systèmes temporellement échantillonnés vont apparaître, mais dans le domaine spatial cette fois-ci. Les sources étudiées seront des ondes planes de longueur d'onde λ . Ainsi, dans le domaine spatial, supposons que nous ayons mis en place un réseau de microphones linéairement espacés d'une quantité Δx constante. La périodicité spatiale de l'onde est la longueur d'onde λ . Pour une onde arrivant avec un angle d'incidence rasant par rapport au réseau de microphones, les conditions de Shannon énoncent que le système est correctement échantillonné si nous disposons au minimum de deux points pour chaque longueur d'onde. Ainsi, si les microphones sont espacés de Δx , la longueur d'onde minimale détectée est $\lambda_{min} = 2\Delta x$, ce qui correspond à une fréquence $f_{max} = \frac{c}{2\Delta x}$, ou à un nombre d'onde $k_{max} = \frac{\pi}{\Delta x}$. Ainsi, pour un angle d'incidence rasant, le système est correctement échantillonné tant que le nombre d'onde d'onde respecte la condition suivante :

$$-k_{max} \leq k \leq k_{max} \quad (1.48)$$

Lorsque l'angle d'incidence devient quelconque, seule la composante du vecteur d'onde colinéaire à la direction est à prendre en compte. Ainsi, si nous appelons θ l'angle d'incidence de l'onde (voir figure 1.7) par rapport à la normale au réseau de microphone ($-\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}$), la condition sur le nombre d'onde devient :

$$|k_y| \leq \frac{\pi}{\Delta x |\sin \theta|} = \frac{2\pi f_{max}}{c |\sin \theta|} \quad (1.49)$$

Mise en évidence du recouvrement spatial : la manière la plus simple pour le mettre en évidence est de considérer une onde plane de fréquence f et ayant un angle d'incidence θ . Pour une onde plane de fréquence donnée, la composante du vecteur d'onde échantillonnée $k_y = k \sin \theta = \frac{2\pi f}{c} \sin \theta$ est uniquement fonction de l'angle d'incidence de l'onde, nous pouvons donc directement associer k_y à la direction de l'onde incidente. Tant que $f \leq f_{max}$, il ne peut y avoir de recouvrement spectral spatial. En revanche, dès que $f > f_{max}$, il y a recouvrement spectral dans le domaine spatial, ce qui implique que le signal reçu sur le réseau de microphones peut avoir été enregistré pour des ondes planes de même fréquence, mais de directions différentes. En effet, l'échantillonnage spatial a pour effet de rendre périodique le spectre spatial de période $2k_{max}$. Donc, si nous procédons de manière analogue au domaine temporel, où le domaine de fréquence était restreint à $[-\frac{F_e}{2}, \frac{F_e}{2}]$, cela a pour conséquences de limiter les angles d'incidences en fonction de la fréquence de l'onde incidente pour se prémunir d'un éventuel repliement spectral.

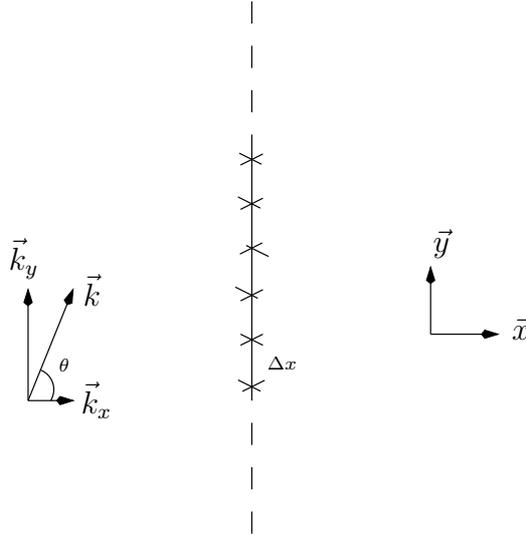


FIG. 1.7 – illustration des notations pour le repliement spectral

Prenons un exemple. Soit $\Delta x = 31.4 \text{ cm}$, le nombre d'onde maximal vaut $k_{max} = 10$, ce qui correspond à une fréquence maximale en incidence rasante $f_{max} = \frac{c}{2\Delta x} = 541 \text{ Hz}$. Toutes les composantes fréquentielles d'ondes incidentes sont correctement enregistrées si leurs fréquences sont inférieures à cette limite. En revanche, si $f > f_{max}$, il faut restreindre le domaine d'incidence pour que $k_y = \frac{2\pi f}{c} \sin \theta \leq k_{max} = 10$. Voici un tableau présentant ces limites pour différentes fréquences :

Fréquence (en Hz)	541	1000	3000	5000	10000
$ \theta_{lim} $ (en degrés)	90	33	10	6.2	3.1

Si nous ne prenons pas nos précautions, nous ne saurons pas avec exactitude de quelle direction provient l'onde plane incidente. Cela nécessite un filtrage spatial. Nous essaierons de quantifier les effets du recouvrement spectral spatial lors de la restitution de la scène sonore dans le cas où ces précautions n'ont pas été prises.

Cette démonstration a été effectuée dans le cas simple d'un réseau linéaire et régulièrement espacé, et dans le cas où la source était une onde plane. Maintenant, il faudrait essayer de mettre en évidence les mêmes limitations qui devraient survenir pour une source quelconque (monopole et dipôle, ce serait déjà bien en pratique), pour des réseaux qui ne sont plus forcément linéaires (pour des réseaux circulaires par exemple), ou bien espacés de manière non constante. Un début d'étude est indiqué pour les réseaux linéaires espacés de manière logarithmique dans [10]. Normalement, tout champ peut être décomposé comme une superposition d'ondes planes, donc le paragraphe précédent est valide pour les réseaux de microphones réguliers. De manière générale, pour un réseau quelconque, il faut s'assurer que nous sommes capables de déterminer la localisation de la source de manière unique uniquement par les données enregistrées par le réseau.

B.2.2.2 Limitations dues aux capteurs utilisés

Revenons sur les limitations intrinsèques aux capteurs employés lors des mesures : les microphones. Ces capteurs sont supposés avoir une grande plage de linéarité, et nous supposons donc qu'ils sont parfaitement linéaires pour notre étude. Il en existe deux types, les microphones sensibles à la pression, et les microphones sensibles au gradient de pression. Les microphones de la première catégorie sont omnidirectionnels tandis que les ceux de la deuxième sont directifs. Si nous souhaitons mesurer la pression, nous aurons donc logiquement recours à des microphones

de pression, supposés omnidirectionnels sur toute leur bande passante. Mais cette hypothèse est de moins en moins vérifiée aux hautes fréquences, à cause de phénomènes d'interférence qui se produisent sur la capsule : le comportement de la capsule pour une onde incidente normale ne sera plus identique à celui correspondant à une onde en incidence rasante, car les dimensions de la capsule ne sont alors plus négligeables devant la longueur d'onde de l'onde. C'est une limitation de ces microphones. Pour mesurer le gradient de pression, nous aurons recours à des microphones dipôlaire, spécialement conçus pour mesurer la composante du gradient de pression selon l'axe du micro. Il suffit alors de placer le microphone perpendiculairement à la surface fermée au point considéré.

B.2.2.3 Influence des sources parasites

Un autre problème qui surviendra inévitablement dans la pratique et qui n'a pas encore été mentionné est la présence inévitable de sources parasites lors de la prise de sons. Par source parasite, nous n'entendons pas ici des sources de bruit extérieures qui viendraient dégrader le signal enregistré, mais des sources qui seraient présentes dans le volume de restitution par exemple, et dont nous ne souhaiterions pas la présence. En toute théorie, si nous utilisons l'intégrale de Kirchoff-Helmholtz, nous savons que le résultat est nul dans le volume délimité par la surface fermée et ne contenant pas les sources. Ainsi, si des sources se situent de part et d'autre de la surface fermée délimitée, nous reconstruisons le champ sonore des sources externes dans le volume intérieur, et nous reconstruisons le champ sonore des sources internes dans le volume extérieur. Si le dispositif de restitution est placé en chambre anéchoïque, la présence de ces sources parasites dans le volume de restitution n'est donc pas gênante, si le reste du système est parfait. Par contre, dans une salle réelle, les réflexions dues aux différentes parois dévoileront de manière indirecte la présence de ces sources parasites dans le volume de restitution. Il faudra donc essayer de quantifier cet effet.

B.3 Prospectives d'applications à la prise de son

Les équations intégrales de Kirchoff et de Kirchoff-Helmholtz affirment que nous sommes capables d'extrapoler le champ de pression acoustique dans un volume délimité par une surface fermée \mathcal{S} , fini ou non, mais ne contenant pas de sources. Ainsi, nous pouvons en déduire deux perspectives d'application des réseaux de microphones pour la prise de son, qui vont être détaillées tour à tour.

B.3.1 Prise de son pour les scènes sonores réelles

La première possibilité consiste à placer toutes les sources acoustiques à l'extérieur de la surface fermée \mathcal{S} , et de disposer tous les capteurs autour de cette surface. Donnons un exemple d'application de cette technique. Nous souhaitons enregistrer un orchestre symphonique, et nous jugeons, arbitrairement peut-être, que le chef d'orchestre occupe la meilleure place pour écouter le concert. Nous considérons donc que le volume à proximité du chef d'orchestre, et ne contenant pas de musiciens — nous considérons que le chef d'orchestre n'est pas une source acoustique — est un volume privilégié pour jouir de l'œuvre exécutée et de profiter de l'acoustique de la salle d'enregistrement. Nous plaçons donc en conséquence tout un réseau de capteurs idéaux autour du chef d'orchestre et enregistrons la scène sonore. Dans ces conditions, si nous plaçons le réseau de transducteurs sur la même surface que celle utilisée pour le réseau de capteurs, il nous est possible de restituer parfaitement l'œuvre initiale, et cette technique porte le nom d'holophonie. En toute théorie, en dehors du volume de restitution, le champ sonore serait nul, ce qui limiterait les problèmes de voisinage. En pratique, il faudra certainement aussi s'affranchir de l'effet de la salle hôte du système de restitution, qui génère des réflexions venant interférer avec le champ

initialement synthétisé, afin de parvenir au meilleur résultat possible. De plus, il nous est possible de restituer le champ de pression dans tout volume contenu dans le volume initial délimité par la surface \mathcal{S} , en extrapolant le champ sonore sur une surface \mathcal{S}' , contenue dans le volume de restitution initial, à partir de celui mesuré sur la surface \mathcal{S} . Ainsi, l'objectif d'une telle prise de son serait de se faire dans un volume le plus grand possible, mais ne contenant pas de sources, et l'auditeur final aurait possibilité de le réduire à souhait pour sa propre écoute personnelle dans son salon, qui ne possède pas les mêmes dimensions, il suffirait juste d'indiquer la position des transducteurs à un ordinateur, qui calculerait les nouveaux signaux de commande à envoyer aux transducteurs pour obtenir le champ initial mais dans un volume plus restreint.

B.3.2 Prise de son pour les scènes sonores virtuelles

La deuxième possibilité est de placer la source acoustique que l'on souhaite restituer à l'intérieur de la surface fermée \mathcal{S} et de placer les microphones tout autour de celle-ci. Les microphones devront servir de capteurs de pression, mais on doit aussi mesurer la dérivée de cette pression par rapport à la normale \vec{n} à la surface \mathcal{S} , il sera donc nécessaire d'utiliser les deux types de microphones différents. L'équation de Kirchoff ne permet pas de calculer le champ de pression en un point du volume intérieur délimité par la surface fermée, mais elle permet en revanche de le calculer en tout point extérieur à \mathcal{S} , ce qui permet d'accéder à la structure du champ de pression en champ lointain. La seule hypothèse faite pour que ce calcul soit réalisable est qu'il n'y ait pas de sources acoustiques situées à l'extérieur de \mathcal{S} , ce qui implique qu'il n'y ait pas non plus de sources secondaires dues à des réflexions sur des parois, et qui se situeraient obligatoirement à l'extérieur de \mathcal{S} : un tel enregistrement est uniquement possible en salle anéchoïque. Dans le cas idéal d'une salle parfaitement anéchoïque, et de capteurs nous permettant de mesurer à la fois la pression, et la dérivée de la pression par rapport à la normale, avec un nombre suffisant de capteurs pour pouvoir nous prémunir d'un éventuel recouvrement spatial, nous sommes alors capables d'extrapoler la valeur du champ de pression de la source en tout point de l'espace extérieur, ce qui est un avantage extraordinaire ! En effet, considérons un volume \mathcal{V}' n'ayant aucune partie commune avec le volume \mathcal{V} , et dans lequel nous souhaitons restituer le champ de la source. Pour parvenir à cet objectif, il suffit de placer un réseau de transducteurs tout autour de la surface \mathcal{S}' délimitant le volume \mathcal{V}' , que nous alimentons par des signaux de commande fournis par l'extrapolation du champ de pression sur la surface \mathcal{S}' calculé à partir du champ mesuré sur la surface \mathcal{S} . Cette technique fonctionne tant que l'on souhaite restituer la source sonore en champ libre. Pour la transposer en espace clos, si nous supposons que les réflexions sur les parois sont de nature spéculaire, il est possible de calculer le champ de pression sur la surface \mathcal{S}' dû à la présence de la source, mais aussi dû à la présence des sources secondaires images de la source principale par rapport aux différentes parois de la salle. Le calcul devient de plus en plus long et fastidieux, car le nombre de sources secondaires à prendre en compte peut vite devenir très important si nous voulons obtenir une bonne restitution (à titre de rappel, nous nous souvenons que la densité de réflexions dans une salle augmente proportionnellement à t^2). Si nous ne faisons plus l'hypothèse que les réflexions sont de nature spéculaire, mais diffuse, le problème devient très complexe, et mérite plus amples réflexions. . .

B.3.3 Conclusion

Ces deux possibilités d'enregistrement de la scène sonore ont des finalités diamétralement opposées. La deuxième se veut particulièrement favorable aux développements concernant les réalités virtuelles, tandis que la première se veut beaucoup plus proche de la réalité et permet d'enregistrer fidèlement la signature acoustique du lieu d'écoute dans lequel se situe la scène sonore. Un des avantages est qu'on a accès avec le même matériel de production et de restitu-

tion à ces deux aspects ! La deuxième technique possède beaucoup plus de possibilités que la première, mais le traitement à effectuer est beaucoup plus fastidieux. En toute théorie, il serait possible de restituer un enregistrement effectué par la première méthode en utilisant la deuxième méthode : il suffirait d’enregistrer chacun des instruments séparément en chambre anéchoïque, de les disposer virtuellement à la même place qu’ils auraient occupé s’ils avaient joué dans la salle de concert, et de construire la salle tout autour, par des traitements informatiques. On conçoit combien un tel traitement est beaucoup plus complexe que la première possibilité. En revanche, la première possibilité permettrait de supplanter rapidement les systèmes classiques actuels, car elle est analogue aux enregistrements stéréophoniques, mais le résultat serait beaucoup plus fidèle par rapport à la scène initiale. Ce n’est pas pour autant qu’il faille stopper le développement de la deuxième possibilité.

C Reconstruction du champ sonore

Dans cette partie, nous supposons que nous disposons des informations suffisantes pour pouvoir contrôler le champ sonore dans une certaine zone de l’espace. En particulier, nous nous placerons dans le cas idéal où nous connaissons la pression et la dérivée de la pression par rapport à la normale en tout point de la surface fermée délimitant la zone de restitution, c’est-à-dire que nous supposons le problème de l’enregistrement spatio-temporel d’une scène sonore résolu. Dans une première partie, nous ferons le bilan des différentes techniques de spatialisation existantes à ce jour. Puis nous développerons les deux méthodes permettant d’obtenir un contrôle du champ sonore sur une zone étendue de l’espace. Finalement, nous essaierons de décrire les différents domaines d’application possibles dans lequel un tel système de synthèse de champ sonore pourrait être utilisé.

C.1 Panorama des différentes méthodes de spatialisation des sons

En matière de spatialisation des sons, plusieurs stratégies peuvent être employées [19] :

- La reproduction du champ physique, c’est-à-dire la reproduction du champ produit par la source dans tous ses détails.
- La deuxième stratégie consiste à approcher ce résultat en donnant une certaine image de la scène sonore à l’auditeur, qui essaie de tromper notre perception auditive de telle sorte que l’illusion produite soit la plus réaliste possible. Cette technique est basée sur les spécificités de la perception auditive, et donc sur des critères psychoacoustiques. Elle ne retient que les éléments déterminants dans le processus de localisation des sources ; la spatialisation ainsi construite est artificielle.

Des exemples de systèmes appartenant à la deuxième catégorie sont la stéréophonie, les techniques binaurales et transaurales, et les techniques développées pour le cinéma, telles que le Dolby Stereo, le Dolby Surround, etc. Les techniques stéréophoniques reposent fortement sur l’effet psychoacoustique créé par les sources “fantômes”, situées entre les deux haut-parleurs, et correctement localisées en fonction des différences interaurales de temps et d’intensité entre les signaux venant de la première et de la deuxième enceinte. Par son principe même, cette reproduction du champ sonore ne se fait pas dans une zone étendue de l’espace. Le plus souvent même, il n’existe qu’un seul point de bonne audition dans une pièce pour un système stéréophonique. Les techniques de Dolby Surround et assimilées sont basées sur le même principe, sauf que cinq enceintes sont utilisées au lieu de deux, et que deux dimensions peuvent être spatialisées au lieu d’une seule pour la stéréophonie. Pour l’industrie cinématographique, ce n’est pas gênant car le support visuel est important, il faut que nous ayons l’impression que le son vienne majoritairement de l’image, les autres enceintes ne servant qu’à recréer des bruits d’ambiance.

Les techniques binaurales et transaurales sont basées sur la mesure des HRTF (Head Related Transfer Function), qui traduisent les différences de comportement de l'oreille selon que le son vienne d'une direction ou d'une autre. Ces HRTF sont souvent très différentes d'un individu à un autre, et il est difficile de standardiser cette méthode pour un groupe général d'individus. Des recherches sont en cours pour améliorer cette technique capable de recréer les trois dimensions de l'espace sonore, avec des restrictions toutefois, la majeure étant que l'auditeur déplace la scène sonore de manière simultanée avec ces mouvements de tête. Une première approche pour reconstruire le champ physique a été donnée par la technique Ambisonic, dont le principe est de décomposer le champ en ondes planes (harmoniques cylindriques ou sphériques). Elle donne des bons résultats lorsque nous sommes en hypothèse de champ lointain ($kr \gg 1$ pour toutes les longueurs d'ondes présentes dans la source sonore). De plus, elle facilite tout le système d'encodage et de décodage de l'organisation spatiale du champ sonore, ce qui réduit la complexité du traitement à effectuer. Nous pouvons démontrer que la technique Ambisonic, et la technique holophonique (basées sur le concept de Wave Field Synthesis décrit dans la prochaine partie de ce document) sont semblables [18], et un système de reproduction de champ sonore basée sur une méthode hybride entre ces deux techniques a déjà été développé [18], [17]. La principale restriction de la technique Ambisonic est que la zone d'écoute correctement restituée diminue en fonction de la fréquence, à tel point qu'en pratique, la zone d'écoute correctement restituée est uniquement centrée sur un seul auditeur, tellement la zone restituée fidèlement diminue lorsque la fréquence augmente. Les deux approches qui seront développées dans la prochaine partie pour la reconstruction physique du champ sonore dans une zone étendue de l'espace sont la synthèse de fronts d'onde (ou Wave Field Synthesis, WFS) ou bien le principe du contrôle des conditions aux limites (ou Boundary Surface Control principe, BSC). Le tableau ci-dessous résume les principales caractéristiques de chaque système de spatialisation sonore :

Méthode	Stéréophonie	Surround	Techniques binaurales et transaurales	Ambisonic	WFS et BSC
Principe	Différences interaurales de temps et d'intensité	Différences interaurales de temps et d'intensité	H.R.T.F (Head Related Transfer Functions)	Décomposition en modes propres (harmoniques sphériques ou cylindriques)	Intégrale de Kirchoff-Helmholtz
Dimensions spatialisées	1	2	3	2/3	3
Nombre de haut-parleurs	2	5+1	2	$N \geq 3$	Liée à la taille de la zone d'écoute et aux performances souhaitées
Zone d'écoute	1 auditeur	1 auditeur	1 auditeur	1 auditeur	Étendue.

C.2 Des différentes méthodologies pour la synthèse d'un champ sonore

Comme nous l'avons mentionné dans la partie précédente, il existe deux différentes possibilités pour contrôler le champ sonore sur une zone étendue de l'espace. Toutes deux sont basées sur l'équation intégrale de Kirchoff-Helmholtz démontrée dans la partie I.B de ce document. Elles ne sont finalement que deux visions différentes du même phénomène, et c'est ce que nous allons mettre en évidence. Chacune de ces deux méthodes possède ses avantages et ses inconvénients, et toutes deux souffrent de limitations communes dues à la discrétisation de l'équation intégrale de Kirchoff-Helmholtz. Une comparaison de ces deux méthodes sera effectuée une fois que nous les aurons décrites.

C.2.1 Le principe de la Wave Field Synthesis

Le principe de la Wave Field Synthesis, dénommé ainsi par ses concepteurs Berkhout [6] et al., travaillant à l'université des technologies de Delft, est basé sur l'équation intégrale de Kirchoff-Helmholtz :

$$\frac{1}{4\pi} \iint_{\mathcal{S}} \left[\frac{e^{jk\|\vec{r}-\vec{r}_0\|}}{\|\vec{r}-\vec{r}_0\|} \text{grad}_0 P(\vec{r}_0) - P(\vec{r}_0) \text{grad}_0 \frac{e^{jk\|\vec{r}-\vec{r}_0\|}}{\|\vec{r}-\vec{r}_0\|} \right] \cdot d\vec{S}_0 = \begin{cases} P(\vec{r}) & \vec{r} \text{ intérieur à } \mathcal{V} \\ 0 & \vec{r} \text{ extérieur à } \mathcal{V} \end{cases}$$

Il spécifie que nous sommes capables de restituer le champ sonore d'une source située à l'extérieur du volume de restitution par la connaissance (par la mesure ou l'extrapolation des mesures, en temps réel ou en temps différé) de la pression et de la composante de la vitesse particulière normale à la surface fermée délimitant le volume de restitution, en tout point de cette surface fermée. En effet, le premier terme de cette intégrale marque la contribution de monopoles acoustiques élémentaires placés en chaque point de la surface fermée, et alimentés par les signaux $\frac{\text{grad}_0 P(\vec{r}_0)}{4\pi} \cdot d\vec{S}_0$, qui sont les conditions aux limites imposées par la source initiale lors de l'enregistrement sur la surface \mathcal{S} , tandis que le deuxième terme marque la contribution de dipôles acoustiques placés en chaque point de la surface fermée, et alimentés par les signaux $\frac{P(\vec{r}_0)}{4\pi} d\mathcal{S}_0$ [8].

Plusieurs réalisations ont été faites à l'aide de ce dispositif, qui sont décrites dans [7], [8], et [9].

C.2.2 Le Boundary Surface Control Principle (BSC)

Le principe de contrôle du champ sonore basé sur le contrôle des conditions aux limites de l'équation intégrale de Kirchoff-Helmholtz a été proposé par S. Ise dans [4]. L'équation intégrale de Kirchoff-Helmholtz peut être vue comme une équation centrale pour les programmes permettant la résolution numérique de l'équation de propagation des ondes, qui fournissent une solution lorsque les conditions aux limites sur les frontières du domaine sont spécifiées. Elle est basée sur une autre interprétation de l'équation de Helmholtz :

$$\frac{1}{4\pi} \iint_{\mathcal{S}} [G_k(\vec{r} | \vec{r}_0^s) \text{grad}_0 P(\vec{r}_0^s) - P(\vec{r}_0^s) \text{grad}_0 G_k(\vec{r} | \vec{r}_0^s)] \cdot d\vec{S}_0 = \begin{cases} P(\vec{r}) & \text{si } \vec{r} \text{ intérieur à } \mathcal{V} \\ 0 & \text{si } \vec{r} \text{ extérieur à } \mathcal{V} \end{cases}$$

Maintenant, supposons que nous divisions la surface \mathcal{S} en N surfaces élémentaires \mathcal{S}_i , et faisons l'hypothèse supplémentaire que la pression et la dérivée de la pression par rapport à la normale soient constantes sur chaque surface \mathcal{S}_i . Dans ce cas, l'équation précédente peut être approchée par la somme suivante :

$$\sum_{i=1}^N \left[g_i \frac{\partial P(\vec{r}_i)}{\partial n} - g'_i P(\vec{r}_i) \right] = \begin{cases} P(\vec{r}) & \text{si } \vec{r} \text{ int\'erieur \`a } \mathcal{V} \\ 0 & \text{si } \vec{r} \text{ ext\'erieur \`a } \mathcal{V} \end{cases} \quad (1.50)$$

o\`u

$$\begin{cases} g_i = \iint_{\mathcal{S}_i} G_k(\vec{r} | \vec{r}_0^s) dS_0 \\ g'_i = \iint_{\mathcal{S}_i} \vec{r} \vec{a} d_0 G_k(\vec{r} | \vec{r}_0^s) \cdot d\vec{S}_0 \end{cases}$$

Dans l'\'equation pr\'ecedente g_i et g'_i sont des constantes d\'etermin\'ees par l'int\'egrale de la fonction de Green et de la d\'eriv\'ee de la fonction de Green par rapport \`a la normale sur chaque surface \'el\'ementaire \mathcal{S}_i . Les pressions $P(\vec{r}_i)$ et les gradients de pression $\frac{\partial P(\vec{r}_i)}{\partial n}$ sont les entr\'ees du probl\'eme. L'\'equation pr\'ecedente implique donc que le champ sonore \`a l'int\'erieur du volume \mathcal{V} est enti\'erement contr\^ol\'e si nous parvenons \`a contr\^oler la pression et le gradient de pression sur les N surfaces \'el\'ementaires \mathcal{S}_i . Dans le cas de l'int\'egrale de Kirchoff-Helmholtz, une m\'ethode a \'et\'e d\'evelopp\'ee pour contr\^oler la pression et l'intensit\'e sonore en quelques points du volume de restitution, cette m\'ethode visant \`a \'economiser le nombre de contraintes \`a remplir pour garantir l'unicit\'e du gradient de pression dans le probl\'eme de Dirichlet propre \`a la surface \mathcal{S} . Cette m\'ethode est d\'evelopp\'ee dans [5] pour les lecteurs int\'eress\'es par le sujet. Le principe du contr\^ole des conditions aux limites sur une surface est donc plut\^ot bas\'e sur les pr\'eceptes du contr\^ole actif et de la th\'eorie des probl\'emes inverses : si nous parvenons \`a imposer les conditions aux limites sur les fronti\'eres du domaine, alors nous disposons de la connaissance du champ en tout point int\'erieur \`a ce domaine.

C.2.3 Comparaison des deux m\'ethodes

Le tableau suivant r\'ecapitule les deux points de vue diff\'erents adopt\'es par les deux m\'ethodes pr\'ecedentes :

	Principe de la WFS	Principe de la BSC
$\begin{cases} G_k(\vec{r} \vec{r}_0^s) \\ \frac{\partial G_k(\vec{r} \vec{r}_0^s)}{\partial n} \end{cases}$	Vue comme des sources secondaires	Constantes d\'etermin\'ees par la forme des surfaces \'el\'ementaires
$\begin{cases} P(\vec{r}_0^s) \\ \frac{\partial P(\vec{r}_0^s)}{\partial n} \end{cases}$	Signaux de commande des sources secondaires	Contraintes \`a imposer sur les surfaces \'el\'ementaires pour contr\^oler le champ sonore

Nous allons maintenant examiner les cons\'equences de l'adoption d'un point de vue ou de l'autre pour le d\'eveloppement d'un syst\'eme de restitution, et voir quels sont les avantages et les inconv\'enients de chacun d'eux, ainsi que leurs limites communes.

C.2.3.1 Limites dues \`a la discr\'etisation

Dans les deux cas, il y aura des limites propres au dispositif dues \`a la discr\'etisation du probl\'eme. Nous avons d\'ej\`a mentionn\'e les limites concernant les syst\'emes \'echantillonn\'es dans le domaine spatial \`a la partie B.2.2.1, les syst\'emes ayant recours \`a la WFS souffriront des m\^emes limites que les r\'eseaux de microphones utilis\'es pour l'enregistrement de la sc\'ene sonore. En ce qui concerne les syst\'emes ayant recours au principe de la BSC, l'origine des probl\'emes provient de la taille des surfaces \'el\'ementaires \mathcal{S}_i . En effet, pour que l'expression d\'emontr\'ee au paragraphe pr\'ecedent soit valide, nous avons suppos\'e que la pression et le gradient de pression \'etaient uniformes sur chacune des surfaces \'el\'ementaires. Cette hypoth\`ese est de moins en moins v\'erifi\'ee lorsque la taille de ces surfaces \'el\'ementaires augmente, c'est-\`a-dire lorsqu'on essaie de diminuer le nombre de transducteurs pour contr\^oler le champ sonore.

C.2.3.2 Contraintes sur les transducteurs

Dans le cas de la Wave Field Synthesis, nous sommes obligés d'avoir recours à des transducteurs ayant une directivité bien particulière, monopolaire ou dipolaire, ou les deux, tout dépend de la géométrie du réseau de restitution, ce qui est assez contraignant dans la pratique car les enceintes ont souvent une directivité monopolaire à basses fréquences, mais ce n'est plus le cas en hautes fréquences, et ce n'est pas pour autant que leur directivité soit dipolaire aux hautes fréquences. Il est toutefois possible d'adapter l'opérateur de synthèse de front d'ondes aux caractéristiques des enceintes, on se reportera à [9] pour plus de détails. Pour le principe du contrôle des conditions aux limites, aucune spécificité n'est demandée particulièrement aux transducteurs, ce qui est un avantage de cette méthode.

C.2.3.3 Contraintes sur le placement des transducteurs

Dans le cas de la WFS, les transducteurs sont placés sur la surface fermée délimitant le volume de la restitution, ce qui peut parfois être très gênant dans la pratique, si le système utilisé obstrue la vision des spectateurs lors d'un spectacle, ou si le réseau de transducteurs doit avoir une géométrie figée pour un utilisateur individuel. Toutefois une extrapolation est toujours possible sur un volume de restitution plus petit que le volume qui a servi pour l'enregistrement. C'est un inconvénient de cette méthode. Dans le cas de la BSC, les transducteurs sont placés à l'extérieur du volume de restitution, et tout le problème réside dans les consignes que l'on doit calculer (problèmes inverses) pour satisfaire les conditions aux limites. En pratique, une fois le système développé, il suffirait d'une séance d'étalonnage, utilisant à la fois des transducteurs et des capteurs, les capteurs étant disposés sur la surface délimitant le volume de restitution, aux points où nous souhaitons contrôler la pression et le gradient de pression, si nous voulons utiliser ce dispositif en temps différé, puis nous n'avons qu'à ranger les microphones pour la suite des opérations. Il faut laisser les capteurs en position si nous voulons utiliser ce dispositif en temps réel (cela ressemble alors très fortement à du contrôle actif). C'est donc un avantage de cette méthode.

C.2.3.4 Comportement vis-à-vis des sources parasites - influence de l'effet de salle

Dans le cas de la WFS, si le système se comporte idéalement au regard de l'intégrale de Kirchoff, alors le champ restitué est correct dans le volume de restitution, et nul en dehors du volume de restitution. Donc nous n'avons pas à priori à nous soucier de l'effet de salle du lieu d'écoute hôte du système de restitution. Malheureusement, il est fort peu probable que le système soit idéal à cause de la discrétisation, si bien que le champ à l'extérieur du volume de restitution ne sera plus nul. Il faudra donc que nous le prenions en compte au moment de la synthèse du champ sonore, et la stratégie adoptée rejoindra celle de la méthode de la BSC, à savoir recourir à la théorie des problèmes inverses pour corriger les signaux de commande envoyés à notre distribution de monopoles et de dipôles et obtenir le bon champ dans le volume de restitution. Il en est de même si nous considérons la présence de sources parasites au moment de l'enregistrement de la scène sonore dans le volume de restitution. Dans ce cas le champ des sources sonores externes est correctement restitué dans le volume de restitution, et nul à l'extérieur, et réciproquement, le champ des sources sonores internes, parasites, est correctement restitué à l'extérieur du volume de restitution et nul dans le volume de restitution. Donc nous rencontrons des problèmes car le champ émis à l'extérieur du volume de restitution subit des réflexions sur les parois de la salle, à moins que la salle ne soit parfaitement absorbante, et vient perturber le champ initialement émis. C'est un inconvénient, car cette fois-ci, la différence entre le champ escompté et le champ produit est à la fois due à l'effet de salle, mais aussi à la signature des sources sonores parasites. Il se peut que nous soyons capables de corriger l'effet de salle, mais il est peu probable que nous puissions diminuer l'influence des sources parasites sans connaître

leur signature. Dans le cas de la BSC, l'effet de salle est naturellement pris en compte lors du calcul des signaux de commande servant à alimenter les transducteurs, ce qui peut être à la fois un avantage et un inconvénient. En effet, supposons que le dispositif utilisé soit commandé en temps réel, donc d'une manière analogue à celle du contrôle actif. Dans ce cas, tout message qui ne ferait pas partie de la consigne serait purement et simplement supprimé par le dispositif, étant donné que les conditions aux limites sur la surface fermée imposent le champ sonore, visant à annuler tout ce qui ne fait pas partie de la consigne. Pour la communication entre des artistes et leur public, il n'est pas possible d'utiliser un tel dispositif. Par contre, si le dispositif est utilisé en temps différé, ou plutôt devrait-on dire de manière passive, après une séance d'étalonnage, le problème précédent est réglé. Il faut alors supposer que le comportement de la salle ne change pas trop entre deux étalonnages successifs du dispositif (hypothèse de stationnarité).

C.2.4 Conclusion

Pour conclure sur ce passage, nous pouvons dire qu'il semble plus prometteur de construire un système de restitution de champ sonore sur une zone étendue de l'espace à partir du principe de contrôle des conditions aux limites sur une surface fermée, car le système résultant posséderait beaucoup plus d'avantages qu'un système ayant recours à la technique de la Wave Field Synthesis. En effet, les systèmes basés sur la WFS sont très proches du modèle physique mais sont compliqués à mettre en œuvre, et ils savent difficilement s'adapter aux contraintes qui se présentent naturellement en pratique. Quant aux systèmes basés sur le principe de la BSC, ils adoptent plus une approche de la théorie des systèmes et du contrôle, sans pour autant limiter la modélisation physique du phénomène, facilitant ainsi la résolution du problème car nous disposons de tous les outils concernant la théorie des systèmes et des problèmes inverses, et s'adaptant mieux à des contraintes pratiques, telles que la disposition des transducteurs dans la salle, etc. Le seul inconvénient de cette méthode est qu'il faut disposer aussi d'un réseau de capteurs en plus du réseau de transducteurs, mais ce réseau peut être limité à un capteur dans le cas du développement d'un système en temps différé, car si nous supposons l'effet de salle comme stationnaire, les réponses impulsionnelles mesurées peuvent être effectuées successivement plutôt que simultanément.

C.3 Perspectives d'applications à la restitution du son

Une grande importance a toujours été attachée à l'organisation temporelle du son, et tous les systèmes de restitution sonore actuels le prouvent bien. L'organisation spatiale de la scène sonore n'est pas absente, mais elle est restée et reste toujours en arrière-plan pour le moment, les systèmes de restitution préférant abuser nos capacités auditives par des artifices acoustiques plutôt que de rendre réellement compte de l'organisation spatiale du champ sonore initial. Il est temps de voir quelles peuvent être les perspectives de systèmes de restitution sonore spatio-temporels, car ces systèmes marquent une réelle avancée dans le domaine. En effet, nous ne pouvons nous empêcher de constater les progrès accomplis pour améliorer l'organisation temporelle des sons entre le premier phonographe et les systèmes existant aujourd'hui. Les systèmes de restitution de champ sonore sur une zone étendue de l'espace constateront un progrès analogue en ce qui concerne l'organisation spatiale du champ sonore, et ce dans des domaines multiples que nous allons maintenant détailler.

C.3.1 Restitution du champ sonore

Le domaine de la reproduction à l'identique d'un champ sonore donné sera certainement celui qui se développera le plus rapidement, car nous avons déjà mentionné les différentes techniques de l'enregistrement de la scène sonore qui seront possibles, à la section B.3. Celle qui nécessitera le moins de développements technologiques sera typiquement celle où on placera le réseau de microphones autour d'une zone ne contenant pas de sources acoustiques, ce qui permettra de reproduire le champ à l'identique dans un volume contenu dans le volume ayant servi pour l'enregistrement. Cela présente déjà néanmoins une réelle avancée dans le domaine de la restitution sonore car les informations relatives au lieu d'écoute, et notamment l'effet de salle, sont correctement enregistrées et restituées par un tel système. Il ne s'agit pas ici de remettre en question le caractère artistique d'une prise de son traditionnelle, qui consiste à placer les microphones à l'endroit qui correspondait le mieux à la sensibilité du preneur de son, et à appliquer des traitements postérieurs pour aboutir au résultat attendu par le directeur artistique. Le système présenté sera capable de reproduire à l'identique le champ initial, mais rien n'empêchera les producteurs, compositeurs et preneurs de son de le modifier postérieurement selon leur volonté afin que le résultat final colle à leurs exigences. En revanche, il sera toujours possible d'apprécier chez soi les qualités acoustiques des salles les plus renommées, et cela ne nécessite aucune modification si nous considérons le point de vue des mélomanes en quête du son le plus pur possible.

C.3.2 Correction/Renforcement/Enrichissement du champ sonore

Le deuxième type d'applications possibles sera certainement celui de la correction acoustique des salles par des dispositifs électroacoustiques. Une salle ne doit pas avoir le même comportement selon qu'elle accueille une conférence de presse, un salon d'exposition, un opéra ou bien un concert de musique symphonique, pourtant on utilise souvent une même salle pour des usages aussi multiples que ceux mentionnés précédemment. Les solutions actuelles sont loin d'être idéales, en témoigne l'acoustique des gares par exemple. Les dispositifs de contrôle du champ sonore peuvent donc apporter une nouvelle donne pour tous ces problèmes, et selon les circonstances, corriger, renforcer, ou enrichir le champ sonore initial. Dans le cas d'une pièce de théâtre par exemple, on pourrait amplifier la voix des acteurs sans pour autant avoir l'impression que le son ne provienne des enceintes. Pour la musique symphonique, on pourrait rajouter des réflexions précoces, ce qui est souvent mentionné comme critère de qualité de la part des musiciens, car ils ont l'impression de mieux entendre leurs partenaires. De plus, il serait possible de renforcer ou d'inhiber la composante réverbérée du champ sonore, et ainsi mieux satisfaire les exigences de la salle selon l'événement qu'elle accueille. Les solutions mécaniques amovibles retenues jusqu'à maintenant ne sont pas pratiques à manipuler. Les nouvelles possibilités offertes par ces dispositifs électroacoustiques pourraient les supplanter à l'avenir. Néanmoins, les traitements électroniques à mettre en œuvre seront déjà de nature plus complexe et ce domaine se développera moins vite que celui de la reproduction à l'identique du champ sonore. Néanmoins, il ouvre les portes du suivant.

C.3.3 Synthèse de champ sonore

S'il est possible d'apporter des modifications acoustiques au comportement acoustique des lieux existants, il est aussi tout à fait possible d'en construire des nouveaux, et les dispositifs

de restitution de champ sonore sur une zone étendue ouvre les voies de l'acoustique virtuelle. Tout d'abord, ils pourraient déjà simuler le comportement de salles et apporter un nouvel outil très performant pour la discipline de l'acoustique prédictive. Ils pourraient de plus servir de simulateurs d'environnement acoustique, ce qui est souvent nécessaire pour la réalisation d'expériences dans le domaine de la psychoacoustique. Un exemple de ce genre d'expérience psychoacoustique a été l'étude du confort acoustique d'une rame de TGV. Les tests se sont déroulés en chambre anéchoïque par la diffusion d'enregistrements réalisés dans une rame de TGV circulant sur les voies. Certains auditeurs ont pu confondre le bruit entendu à celui émis par un aspirateur, donc cela peut venir de deux causes : soit le support visuel est important, soit les signaux diffusés ne rendent pas compte de l'organisation spatiale du champ sonore. Mais pourquoi se contenter de simuler des choses existantes ou tester les acoustiques de futurs lieux d'écoute devant être construits dans l'avenir ? L'information sur l'organisation spatiale du champ sonore n'a plus besoin de support physique et réel pour être véhiculée, étant donné que nous serons capables de simuler le comportement d'une salle particulière ailleurs que dans le lieu original. L'organisation spatiale des sons n'est donc plus limitée que par l'imagination de l'utilisateur, et devient un outil de création artistique. Les compositeurs pourront simuler l'acoustique de lieux virtuels qui ne rempliront pas forcément les contraintes de lieux réels ou pouvant exister, car il n'est pas nécessaire de les bâtir. Ces lieux pourraient ne pas pouvoir être construits dans le monde réel à la limite, ce qui nous importe est leur acoustique. De plus, à partir de la deuxième technique de prise de son mentionnée à la partie II.B.4, nous serons capable de faire évoluer des sources réelles dans cet espace virtuel, en plaçant le réseau de microphones tout autour de la source acoustique, le champ sonore pouvant alors être extrapolé dans le reste de l'espace par les données enregistrées par ce réseau. Mais, de même, nous ne sommes pas limités à faire évoluer des sources réelles dans une acoustique virtuelle, et nous pouvons faire la synthèse complète du champ sonore d'une source virtuelle placée dans une acoustique virtuelle. Ainsi, la discipline de la synthèse sonore qui se cantonnait à retransmettre uniquement l'organisation temporelle des sons et non l'organisation spatiale des sons pourra généraliser ces modèles afin qu'ils prennent en compte les caractéristiques spatio-temporelles de la source d'une part, et celle du lieu d'écoute d'autre part, ces deux parties étant dissociées. Ainsi, l'artiste disposera de tous les outils nécessaires à la création artistique, et l'univers des possibilités s'offrant à lui n'aura plus de limites. . .

Chapitre 2

Conception d'un nouveau système de restitution sonore

Le premier chapitre de ce rapport a présenté les rappels théoriques nécessaires pour pouvoir appréhender le domaine de la restitution sonore sur une zone étendue, ainsi que l'état de l'art sur les différentes techniques d'enregistrement et de reproduction du champ sonore utilisées actuellement. Dans le présent chapitre, nous allons mettre à contribution tous les résultats du chapitre précédent afin d'établir les bases d'un nouveau système de restitution sonore, qui, je l'espère, apportera un point de vue nouveau dans la discipline. Le but n'est pas ici de critiquer les systèmes actuellement en cours de développement dans d'autres laboratoires ou organismes, ce qui serait fortement présomptueux de notre part, étant donné que la solution retenue n'est pas idéale. En revanche, la technique qui est présentée dans ce document s'appuie sur des hypothèses qui sont rarement faites dans ce domaine, et nous avons bon espoir que cela puisse aboutir à des résultats bénéfiques, qui pourront être mis à profit par la suite par d'autres techniques.

L'étude réalisée sur la spatialisation sonore relatée dans ce chapitre s'incorpore à un projet plus général, le projet Demix-Remix, interne du groupe des écoles des télécommunications (GET), auquel appartient l'ENST. Dans un premier temps, nous présenterons les objectifs de ce projet. Puis nous nous focaliserons sur le développement de notre nouveau système de restitution sonore. Pour cela, nous présenterons tout d'abord quelques-uns des paramètres de 'haut niveau' caractérisant tout système de restitution sonore. Puis nous détaillerons quelle est la position retenue pour notre système de restitution vis-à-vis de chacun de ces paramètres, ce qui a pour objectif de répondre à la question : Quels sont les objectifs du dispositif, et, indirectement, quelles sont ses limitations ? Ces objectifs établis, nous ferons un inventaire des techniques et traitements susceptibles de permettre la résolution partielle ou totale de notre problème, que nous n'avons malheureusement pas eu le temps de mettre en place mais qui seront testées lors de la poursuite de ce travail en thèse, qui débutera à la prochaine rentrée universitaire.

A Présentation du projet Demix-Remix

Le projet Demix-Remix a pour but de faire se rejoindre deux axes de recherche dans le domaine de l'audio :

- La spatialisation des sons qui donne à l'auditeur la sensation du relief sonore par le placement de sources sonores dans un espace tridimensionnel, en utilisant plusieurs enceintes pour la restitution et en filtrant les signaux par des filtres adaptés.
- La séparation des sources qui part d'un signal dans lequel plusieurs sources sont mélangées, et extrait les différentes sources.

En jargon d'ingénieur du son, "remix" est le raccourci pour signifier une opération de post-production qui vise à reconstituer un mixage différent de celui précédemment réalisé à partir des bandes originales.

L'objectif du projet serait de proposer à l'auditeur d'effectuer lui-même cette opération de mixage en lui laissant la possibilité de placer les sons à sa convenance. L'étape préparatoire d'un tel processus consiste à isoler les sources acoustiques présentes sur un enregistrement original. Pour y parvenir, on s'appuie sur des techniques de séparation de sources dans le domaine fréquentiel en mono-capteur, ou bien des méthodes spatiales ou spatio-temporelles, si les signaux disponibles sont multivoies.

"Demix" et "Remix" peuvent se faire conjointement, au niveau de l'équipement de l'auditeur. On peut aussi envisager qu'ils se fassent aux deux bouts d'un canal véhiculant des signaux MPEG-4, car ce codage permet de coder un signal en le décomposant en plusieurs composantes, chacune pouvant d'ailleurs être codée avec une technique différente des autres composantes.

Nous allons détailler succinctement les grandes idées concernant la séparation de sources et l'indexation multimédia, et nous omettrons le domaine de la spatialisation, étant donné qu'il sera développé de manière beaucoup plus approfondie dans le reste du chapitre.

A.1 Séparation de sources

Le de-mixage audio s'appuie sur la séparation des sources. L'objectif est de reconnaître dans le signal plusieurs composantes indépendantes les unes des autres et de les extraire. Ce problème peut se formuler de deux manières différentes suivant que le signal à traiter est monophonique, ou qu'il a été acquis avec plusieurs capteurs. Le vocabulaire des chercheurs distingue ainsi la séparation de sources, qui est une technique multivoies, et l'analyse de scènes audio, reposant sur des traitements monophoniques. Dans un format monophonique, on ne peut distinguer les instruments différents que par les caractéristiques du son de ces instruments. Dans une représentation multivoies, chaque instrument obtient en plus une signature spatiale unique. Si on imagine un enregistrement stéréo effectué par deux microphones placés de manière judicieuse, alors chaque instrument, de part sa place dans l'orchestre, a une fonction de transfert unique vers chacun des microphones. Cette fonction de transfert représente la propagation acoustique entre l'instrument et chacun des microphones. La propagation comprend le chemin direct et une multitude de chemins réfléchis.

A.1.1 Séparation multivoies

Des techniques de séparation performantes ont été mises au point ces dernières années, surtout dans le contexte des communications numériques. Mais l'adaptation de ces techniques aux signaux audio n'est pas immédiate, à cause de la complexité des fonctions de transfert mises en jeu. Les méthodes de séparation qui fonctionnent bien en communications exploitent certaines hypothèses sur la nature des sources à séparer, notamment que de telles sources appartiennent toujours à un alphabet fini. Un signal audio, en revanche, ne correspond pas à un signal à alphabet fini, et les méthodes classiques de séparation ont tendance à restituer un signal filtré, voire partialement blanchi. La qualité de la séparation est très décevante à l'écoute, nécessitant en pratique un post-filtrage pour restaurer le spectre d'origine. La réussite de cette technique de post-filtrage suppose des connaissances des spectres à restaurer, connaissance difficilement conciliable avec les caractéristiques transitoires des signaux audio. Pour ces raisons, la séparation des sources appliquée à l'audio n'a pas rencontré les mêmes gains que la séparation appliquée aux communications numériques.

A.1.2 Séparation monophonique

Les traitements monophoniques s'appuient le plus souvent sur le caractère harmonique des signaux musicaux. La note fondamentale (de fréquence ou pitch F) est accompagnée de la série de ses harmoniques ($2F$, $3F$, ...). Quand la fréquence a été estimée, l'estimation des amplitudes ne soulève pas de gros problèmes. Elle se fera par des estimateurs classiques, tels que les moindres carrés ou le maximum de vraisemblance. La difficulté provient en général de l'estimation des pitchs multiples ainsi que de l'inharmonicité (légers écarts des fréquences de partiels, par rapport à ce que seraient ces fréquences dans la série exactement harmonique). A cette difficulté s'ajoute la nécessité de modéliser la partie des signaux qui n'est pas purement harmonique, par exemple les attaques des sons. Chaque composante s'obtient à la fin du traitement par une synthèse à partir de ses caractéristiques.

A.2 Indexation multimédia

Une séparation de sources efficace permettrait également une identification plus robuste des signaux en présence et représente ainsi l'une des directions possibles pour l'indexation automatique de signaux audio (reconnaissance des instruments de musique, extraction de fréquences fondamentales multiples, ...). Chaque composante extraite peut se caractériser indépendamment des autres. On extrairait la partition musicale à partir de la succession des notes fondamentales détectées. L'identification de la signature spectrale ouvrirait la voie à la reconnaissance de l'instrument joué. La détection des rythmes serait également facilitée par la séparation en composantes élémentaires : il sera plus facile sur chaque composante de détecter les débuts et fins de notes (ou d'objets sonores, par exemple de percussions) ; la détection des synchronismes entre composantes rendra plus aisée l'extraction des rythmes.

B Paramètres synthétiques des dispositifs de restitution sonore

Avant de développer dans les détails la solution que nous avons retenue pour notre nouveau système de restitution sonore, il semblerait intéressant d'extraire quelques-uns des paramètres de 'haut niveau' de tout système de restitution sonore, visant à décrire leurs performances en termes de qualité de la scène sonore restituée. Après quelques réflexions, nous en avons retenu quatre, et il se pourrait qu'il y en ait d'autres :

- La taille de la zone d'écoute.
- La modélisation des sources sonores.
- L'égalisation du système de reproduction.
- L'implantation d'un effet de salle artificiel.

Dans toute la suite de cette partie, nous revenons successivement sur chacun des points mentionnés ci-dessus, et évoquons les principales difficultés inhérentes à chacune de ces catégories. Une fois ces paramètres explicités, cela nous permettra de mieux comprendre la démarche qui nous a conduit à adopter la solution décrite à la section C, qui sera un compromis entre les performances désirées et les difficultés sous-jacentes à un tel traitement, afin d'obtenir une meilleure restitution du champ sonore que les dispositifs actuels, ce qui demandera une confirmation *a posteriori* par des expériences perceptives.

B.1 Taille de la zone d'écoute

Un des premiers paramètres permettant de caractériser les performances et la qualité d'un procédé de restitution sonore est la taille de la zone d'écoute dans laquelle le champ sonore est restitué de manière *optimale*. Pour ce critère, il existe deux grandes familles de systèmes de

restitution de champ sonore : ceux dont la taille de la zone d'écoute est limitée, généralement, à un seul auditeur, et ceux dont la taille de la zone d'écoute est étendue, et qui permettent ainsi de véhiculer une bonne information concernant le relief sonore de la scène initiale dans un certain volume donné de l'espace, que l'on souhaitera le plus grand possible. Ainsi, même si la restitution stéréophonique par deux haut-parleurs d'une scène sonore est effectuée sur une zone étendue de l'espace, appelée le lieu d'écoute, ce procédé ne permet une restitution 'fidèle' de la scène initiale qu'en un unique point, communément appelé '*sweet spot*' dans la littérature, et la restitution stéréophonique sera donc classée dans la catégorie des techniques dont la taille de la zone de restitution est réduite. Dans cette catégorie figurent aussi tous les systèmes de restitution basés sur les HRTF (Head Related Transfer Functions), dont le but est de filtrer de manière appropriée un signal source initial en fonction de la direction de laquelle il provient, et les systèmes transoraux, qui visent à reproduire de manière identique chacun des deux signaux d'un enregistrement stéréophonique sur chacune des deux oreilles de l'auditeur, ce qui nécessite d'annuler les trajets croisés par un filtrage adapté. En revanche, des procédés de restitution sonore tels que les méthodes ambisoniques, qui cherchent à enregistrer et restituer chacune des composantes du champ sonore décomposé selon une certaine base (base cylindrique ou sphérique), ou bien tels que la Wave Field Synthesis et les méthodes dérivées de cette technique, qui cherchent à restituer le front d'onde d'une source initiale par un réseau de transducteurs, appartiennent à la catégorie des techniques dont la taille de la zone de restitution est étendue, car il s'agit des fondements mêmes de ces méthodes. Toutes ces méthodes de restitution sur une zone étendue, reposent, de manière directe ou indirecte, sur l'équation intégrale de Kirchhoff, qui généralise le principe de Huygens énoncé dans le domaine de l'optique, dans le cas où les sources sonores secondaires servant à reconstruire le champ initial ne sont plus situées sur un front d'onde de la source primaire. Les traitements sous-jacents à de telles techniques sont forcément plus complexes que pour les dispositifs à zone d'écoute limitée, mais le jeu en vaut peut-être la chandelle d'un point de vue perceptif.

B.2 Modélisation des sources acoustiques

Ce point est rarement mentionné lorsqu'il s'agit de qualifier les performances d'un système de restitution sonore, et pourtant, il gagnerait à l'être. En effet, si nous nous remémorons les premiers âges de la prise de son, un seul microphone était utilisé au cours de l'enregistrement, ce qui n'était pas grave, étant donné qu'un seul transducteur était utilisé lors de la restitution. Avec l'arrivée des techniques stéréophoniques se sont offertes deux possibilités :

- Perpétuer l'enregistrement monophonique. Dans beaucoup d'oeuvres musicales, notamment la majorité des produits de l'industrie musicale, les enregistrements sont souvent effectués avec un seul microphone lors de la prise de son. C'est un traitement au moment de la séance de mixage qui permet de spatialiser la source, en affectant des poids ou des retards différents aux deux signaux de sortie. Une telle source est nommée *source fantôme*.
- Mettre au point une nouvelle méthode d'enregistrement, capable de tirer bénéfice du nouveau système de restitution. C'est ainsi que sont apparues les techniques de prise de son stéréophoniques, qui ont permis de rendre une 'meilleure image' perceptive du relief sonore de la scène initiale.

Cette réflexion peut être mûrie dans un contexte plus général, qui est celui de la modélisation des sources, et que nous allons développer maintenant.

B.2.1 sources ponctuelles

Si nous considérons les sources physiques idéales, le monopole et le dipôle, la théorie nous enseigne qu'un seul signal est nécessaire pour les caractériser : il s'agit de la signature temporelle

de la source sonore. Ainsi, si nous disposons de suffisamment de caractéristiques permettant de rendre compte du comportement acoustique du lieu d'enregistrement, nous sommes capables d'extraire cette signature temporelle à partir d'un enregistrement monophonique réalisé dans ce lieu d'écoute.

Nous pouvons raffiner la représentation des sources ponctuelles, en essayant d'intégrer des renseignements concernant leur directivité, ce qui est fait dans le cas du dipôle, et ce qui peut être généralisé à des diagrammes de directivité quelconques. Dans tous les cas, la connaissance de la directivité de la source, et de sa signature temporelle suffit à caractériser la source ponctuelle.

Le modèle de la source acoustique ponctuelle, aussi complexe soit-il (dans le cas où son diagramme de directivité est complexe), est utopique pour rendre compte du comportement de sources sonores réelles, qui appartiennent quasiment toutes à la catégorie des sources étendues. Le recours à une modélisation ponctuelle des sources sonores réelles aura donc certaines limites quant à la qualité de l'image sonore restituée : pour s'en convaincre facilement, il suffit de comparer deux enregistrements stéréophoniques d'une même oeuvre, l'un effectué à partir d'une prise de son monophonique et spatialisé en jouant sur le panoramique au moment du mixage, et l'autre effectué par une prise de son stéréophonique. Le relief de la scène sonore est beaucoup mieux rendu, d'un point de vue perceptif, dans le cas d'une prise de son stéréophonique. Il semble nécessaire de s'intéresser à la caractérisation de sources sonores étendues.

B.2.2 sources étendues

Contrairement aux sources ponctuelles, la signature temporelle d'une source étendue ne peut pas se réduire à un unique signal. Il n'y a qu'à considérer le cas élémentaire d'une source étendue constituée de deux monopoles de signatures temporelles indépendantes. Dans ce dernier cas, deux signaux sont nécessaires pour caractériser de manière unique la signature temporelle de la source. Ainsi, si nous connaissons avec perfection le comportement du lieu d'enregistrement, ce qui est un autre problème, et ce n'est pas l'objectif ici de le résoudre, nous serions capables, par un traitement approprié, d'extraire les deux signatures temporelles de chacun des monopoles constituant la source étendue. Mais qu'en serait-il pour une source sonore réelle, tel qu'un violon, ou un piano ? Il s'agit d'un cas beaucoup plus complexe d'étude, et ce n'est pas non plus notre prétention de le résoudre ici, alors que des chercheurs tentent de modéliser les instruments musicaux depuis la nuit des temps, en obtenant seulement une connaissance partielle de leur fonctionnement. Le cas des sources étendues peut toujours être considéré comme un cas limite d'une superposition d'une multitude de sources ponctuelles, et il faudrait donc théoriquement la même multitude d'enregistrements pour extraire la signature de la source sonore. On voit que cette méthode n'est pas envisageable. Une autre procédure de caractérisation des sources sonores a été envisagée à la section B.3 de l'étude bibliographique. Nous retiendrons que la caractérisation des sources sonores est de plus en plus fidèle à la source initiale lorsque le nombre de microphones utilisés lors de la prise de son est important. Ce problème reste toutefois un problème ouvert, et n'a pas encore reçu de solution définitive.

B.3 Implantation d'un effet de salle artificiel

La question de la modélisation des sources sonores était un paramètre important sur lequel il était nécessaire de s'attarder lors du développement d'un système de restitution sonore, car il s'agit de la composante active responsable de l'évolution temporelle du champ. La deuxième composante responsable de l'évolution du champ, qualifiée de passive, est la connaissance du comportement acoustique du lieu à restituer. Il est donc nécessaire de coder le comportement d'une salle, d'une manière laissée aux soins du lecteur, qui ne sera pas développée ici, car le sujet n'a pas encore reçu de réponse définitive, et constitue un sujet à part entière. Cela concerne

à la fois les disciplines de l'acoustique prédictive et le développement de programmes simulant l'acoustique de réalités virtuelles. En théorie, le comportement de la salle est indépendant de celui des sources sonores. En pratique, il est très difficile de dissocier l'un de l'autre dans les enregistrements, car ces deux problèmes sont fortement couplés.

B.4 Égalisation du système de reproduction

L'égalisation du système de restitution est une caractéristique importante des dispositifs de restitution sonore, qui est malheureusement rarement effectuée, ce qui est compréhensible en regard des traitements à effectuer pour y parvenir. Elle possède deux objectifs :

- S'affranchir dans une certaine mesure du matériel utilisé au moment de la restitution — Amplificateur de puissance, mais surtout enceintes.
- S'affranchir de l'effet de salle, dans laquelle le système de restitution a été placé. En effet, dans le cas d'enregistrements stéréophoniques par exemple, les deux signaux enregistrés viennent alimenter directement les deux transducteurs utilisés au moment de la restitution. Ils évoluent par la suite dans un canal acoustique qui vient les déformer par des réflexions successives, de la diffusion, et de la diffraction, caractéristiques de la salle hôte, avant d'atteindre les oreilles de l'auditeur. Le processus d'égalisation essaie de compenser cet effet de salle.

De manière générale, le processus d'égalisation a pour objectif de s'affranchir du matériel ainsi que des caractéristiques du lieu d'écoute, dans lequel le système de restitution a été placé. Il permet donc théoriquement d'écouter le même enregistrement sur des matériels différents, en des lieux différents. Le processus d'égalisation est plus ou moins complexe, selon que l'on se place dans un contexte monovoie, ce qui est le cas si nous utilisons l'hypothèse de ponctualité pour la modélisation des sources sonores réelles, ou dans un contexte multivoies, si nous essayons de modéliser des sources sonores étendues.

B.5 Conclusion

Maintenant que nous avons détaillé les paramètres importants permettant de juger de la qualité des systèmes de restitution sonore, examinons un exemple de dispositif de restitution sonore en détails : étudions le cas de la restitution stéréophonique. Pour ce qui concerne la taille de la zone d'écoute, nous avons vu qu'elle était limitée en pratique à un unique auditeur, placé au niveau du *sweet spot*. En ce qui concerne le problème de la modélisation des sources sonores, et celui du contrôle de l'effet de salle, ils sont tous les deux couplés et contenus dans les deux signaux enregistrés au moment de la prise de son, ce qui empêche de faire des traitements performants par rapport au cas où les deux paramètres seraient découplés. La meilleure restitution que nous pouvons obtenir au moyen d'un tel procédé serait obtenue à la fin du processus d'égalisation, s'il était effectué : on recevrait alors le premier signal enregistré sur l'oreille gauche, et le deuxième signal sur l'oreille droite. Pour conclure, le système stéréophonique ne possède pas une zone d'écoute étendue, ne permet pas de manipulations raffinées sur le contrôle de l'effet de salle, et est fortement tributaire de la salle de restitution. Son seul avantage est d'utiliser deux microphones lors de l'enregistrement, et cela n'est malheureusement pas toujours le cas. Ce constat nourrit notre espoir de développer un dispositif de restitution sonore fournissant de meilleurs résultats.

Examinons maintenant les dispositifs basés sur les HRTF. Il s'agit d'un dispositif à zone d'écoute limitée, étant donné qu'il est réalisé au casque. Les sources sonores sont modélisées de manière ponctuelle, et la technique de spatialisation a pour but de mettre en place un filtrage adapté de la signature temporelle en fonction de la direction d'incidence de la source. Le contrôle de l'effet de salle peut être mis en place, par l'ajout de sources secondaires par exemple, rendant compte de réflexions multiples. Il est réalisé de manière artificielle, ce qui est un avantage : la

modélisation des sources et le contrôle de l'effet de salle sont découplés. Le dernier avantage est qu'il nécessite un léger processus d'égalisation des écouteurs uniquement, et qu'aucune inversion de canal acoustique n'est à faire. Cette méthode de spatialisation est en cours de généralisation pour pouvoir être appliquée sur deux enceintes au lieu de l'être au casque, comme en témoigne un autre axe de recherche de l'ENST, toujours dans le cadre du projet Demix-Remix. Cela nécessite un processus d'égalisation plus complexe, car il doit aussi inverser le canal acoustique, en éliminant les trajets croisés de l'enceinte gauche à l'oreille droite, et de l'enceinte droite à l'oreille gauche. C'est donc une généralisation des systèmes transoraux, qui possédera toujours l'inconvénient de n'être valable que dans une zone limitée de l'espace. Le dernier dispositif en cours de développement, à l'IRCAM, est la généralisation du dispositif précédent afin qu'il soit réalisé sur une zone étendue. La technique de spatialisation utilisée est celle de la Wave Field Synthesis, les sources sont modélisées de manière ponctuelle (à directivité variable), et une batterie de filtres est utilisée pour modéliser différentes directions d'incidence de la source primaire, en s'affranchissant de l'effet de salle de la salle hôte.

Toutes les méthodes précédentes permettront d'obtenir une excellente localisation des sources dans un environnement tridimensionnel, sur une zone d'écoute limitée ou étendue, et permettant de plus un certain contrôle de l'effet de salle. Leur seul inconvénient réside dans le fait que les sources sonores restituées sont ponctuelles, ce qui dégradera la signature temporelle de la source initiale.

Le dispositif idéal permettrait de garder les avantages du système basé sur les techniques de WFS, zone d'écoute étendue, modélisation des sources découplée de la modélisation de l'effet de salle, et il ne ferait plus non plus l'hypothèse de ponctualité des sources sonores restituées. Malheureusement, les traitements à effectuer lorsque cette hypothèse n'est plus vérifiée sont beaucoup plus ardues, car ils s'inscrivent dans un cadre multivoies, mais nous aurons le temps d'y revenir dans la suite de ce document. Nous pouvons synthétiser toutes les données permettant de juger des performances de chacun des systèmes de restitution sonore sous la forme d'un tableau :

Procédé de restitution	Taille de la zone d'écoute optimale	Modélisation des sources	Contrôle de l'effet de salle	Egalisation
Stéréophonie	limitée	'étendue' (2 voies)	Non	Non
Transaural simple	limitée	'étendue' (2 voies)	Non	Oui
HRTF (binaural ou transaural)	limitée	ponctuelle (1 voie)	Oui	Oui
Ambisonique / WFS	étendue	ponctuelle (1 voie)	Oui	Oui

C Présentation de la solution retenue

C.1 Description générale

L'objectif du dispositif qui sera développé à l'avenir à l'ENST est d'être capable de restituer des champs sonores sur une zone étendue de l'espace, grâce à un réseau de transducteurs. Les sources sonores seront modélisées par des signaux multivoies, afin de rendre compte, dans une certaine mesure, de leur caractère étendu. Le contrôle de l'effet de salle ne sera pas implanté directement au commencement de ce projet — l'effet de salle du lieu d'enregistrement sera quand même pris en compte implicitement, mais cela sera expliqué ci-après — mais il pourra faire l'objet d'études ultérieures si les premiers résultats obtenus sont concluants. De plus, une égalisation sera effectuée afin d'essayer de s'affranchir du matériel et de compenser l'effet de salle du lieu où est disposé le système de restitution. Pour tenter de comparer ce système à un dispositif

déjà existant, il s'agirait d'une amélioration de la restitution stéréophonique, dans le sens où la plupart des inconvénients seraient ôtés —zone limitée, absence d'égalisation— en conservant l'avantage de pouvoir modéliser les sources sonores par des signaux multivoies. Malheureusement, ce système n'est pas le système idéal décrit à la fin de la section précédente, et il existe bien un défaut qui vient limiter ses performances : la modélisation des sources sonores, et celle du lieu d'enregistrement sont couplées, et il est impossible pour le moment de les manipuler de manière indépendante. Ainsi, le système, tel qu'il est décrit, sera au mieux capable de restituer un champ initial enregistré sur plusieurs voies dans un autre lieu : nous conserverons normalement la signature temporelle des sources et la signature du lieu d'écoute, ce qui serait déjà un énorme progrès. En revanche, des modifications seront nécessaires si on veut adapter ce système pour qu'il puisse faire de la synthèse de champ sonore, ce qui fait partie du domaine des réalités virtuelles, car il est nécessaire de contrôler de manière indépendante les sources et l'effet de salle que l'on souhaite restituer pour y parvenir.

C.2 Présentation du problème direct

Dans cette partie, nous supposons que le champ sonore initial a été enregistré en I points, et nous connaissons de plus la géométrie du réseau d'enregistrement. Nous souhaitons contrôler le champ sonore restitué par un réseau de I transducteurs en K points. Le système présenté est illustré par un schéma sur la figure 2.1.

Nous pouvons relier les K signaux représentant la valeur de la pression acoustique à contrôler à J signaux d'entrée du réseau de transducteurs par l'intermédiaire de la relation suivante :

$$\begin{pmatrix} s_1(t) \\ \vdots \\ s_K(t) \end{pmatrix} = \begin{bmatrix} h_{11}(t) & \dots & h_{1J}(t) \\ \vdots & \ddots & \vdots \\ h_{K1}(t) & \dots & h_{KJ}(t) \end{bmatrix} \star \begin{pmatrix} y_1(t) \\ \vdots \\ y_J(t) \end{pmatrix} \quad (2.1)$$

où les signaux $s_{[1\dots K]}$ désignent l'évolution temporelle de la pression acoustique aux K points de contrôle, où les signaux $y_{[1\dots J]}$ désignent les signaux d'entrée des J transducteurs, et où les $h_{kj}(t)$ désignent les réponses impulsionnelles du canal acoustique entre le transducteur j et le point de contrôle k . Pour obtenir cette formule où \star désigne le produit de convolution, nous avons supposé que le système était linéaire, ce qui est une approximation, car la chaîne de restitution comporte des non-linéarités, notamment au niveau des transducteurs, mais nous les négligeons dans le reste du traitement.

De plus, tous les traitements effectués sur ces signaux seront faits de manière numérique, nous travaillerons donc plutôt sur des données discrètes, issues d'un processus d'échantillonnage des différentes grandeurs de l'équation précédente, qui devient :

$$\begin{pmatrix} s_1(n) \\ \vdots \\ s_K(n) \end{pmatrix} = \begin{bmatrix} h_{11}(n) & \dots & h_{1J}(n) \\ \vdots & \ddots & \vdots \\ h_{K1}(n) & \dots & h_{KJ}(n) \end{bmatrix} \star \begin{pmatrix} y_1(n) \\ \vdots \\ y_J(n) \end{pmatrix} \quad (2.2)$$

où nous avons remplacé la variable temporelle t par l'indice du signal n , qui désigne la valeur du signal pour $t = nT_e$ où T_e désigne la période d'échantillonnage de notre système. Les signaux échantillonnés ont été préalablement filtrés par un filtre anti-repliement de sorte qu'ils soient à bande limitée, et que la condition de Shannon sur l'échantillonnage soit respectée. Il faut bien faire attention que cette écriture ne peut pas être implantée directement sur l'ordinateur car le produit de convolution sur des signaux discrets fait intervenir aussi bien les échantillons antérieurs que postérieurs à n . Il s'agit d'une écriture analytique.

Nous pouvons aussi relier les signaux reçus aux K points de contrôle du champ sonore aux I signaux enregistrés, par la relation :

$$\begin{pmatrix} s_1(n) \\ \vdots \\ s_K(n) \end{pmatrix} = \mathbf{A} \star \begin{pmatrix} x_1(n) \\ \vdots \\ x_I(n) \end{pmatrix}, \text{ où } \mathbf{A} = \begin{bmatrix} a_{11}(n) & \dots & a_{1I}(n) \\ \vdots & \ddots & \vdots \\ a_{K1}(n) & \dots & a_{KI}(n) \end{bmatrix} \quad (2.3)$$

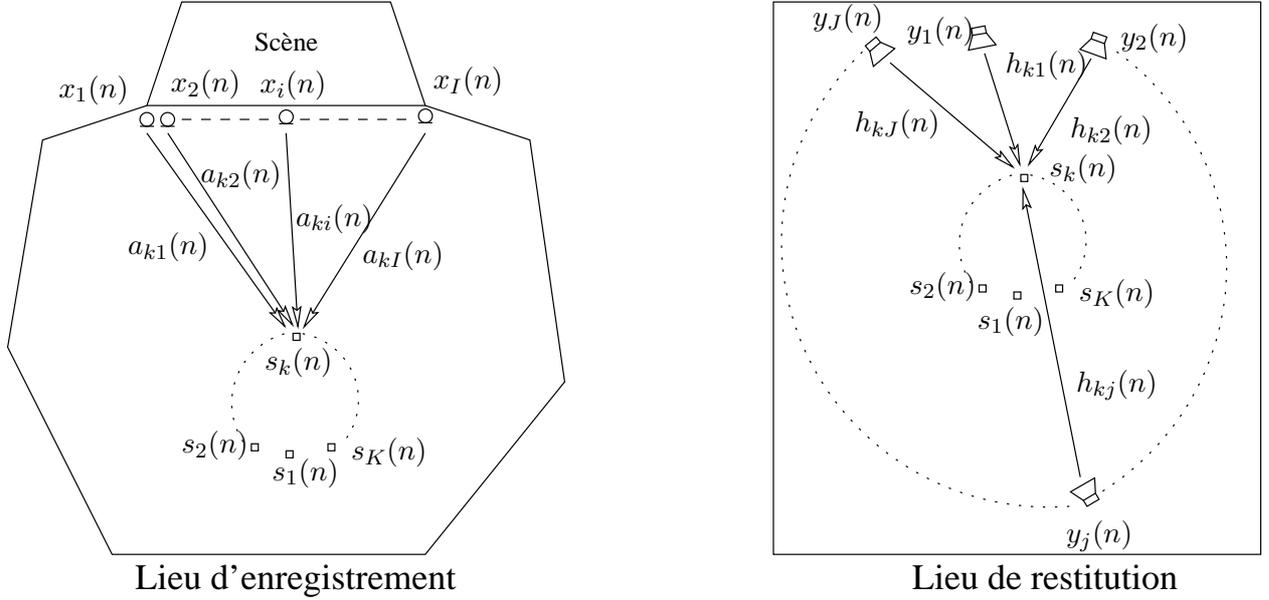


FIG. 2.1 – Représentation schématique du système de restitution

où \mathbf{A} désigne une matrice $K \times I$, dont l'effet est d'extrapoler le champ sonore en K points par la connaissance de celui-ci en I points. On nommera pour cela \mathbf{A} matrice d'extrapolation, afin de conserver un formalisme général, mais, dans la suite de ce document, nous nous contenterons d'avoir $I = K$, c'est-à-dire le même nombre de points de contrôle que le nombre de signaux enregistrés, et nous ferons de plus l'hypothèse supplémentaire que la géométrie du réseau de microphones qui a servi lors de l'enregistrement est la même que la géométrie du réseau de points où nous souhaitons contrôler le champ. Ce n'est pas le cas représenté sur la figure 2.1. Ce serait vrai si nous avions placé directement les microphones dans le lieu d'enregistrement au niveau des K points de contrôle, ce qui supprimerait l'étape d'extrapolation. La matrice \mathbf{A} contient alors typiquement des séries de Dirac sur la diagonale et est nulle ailleurs, de sorte que les signaux de consigne soient identiques aux signaux enregistrés (en pratique, à un délai prêt, mais ce sera expliqué ultérieurement).

C.3 Formulation du problème inverse

La formulation du problème direct nous permet de calculer les signaux représentant la pression acoustique aux K points de contrôle en fonction des J signaux d'entrée des transducteurs, mais ce n'est pas le problème qui nous intéresse. Le problème qui nous préoccupe est de calculer les signaux d'entrée des transducteurs à partir de la connaissance des K signaux de contrôle, que cette connaissance soit le fruit de mesures directes —dans le cas où $K = I$ et où les configurations géométriques des réseaux d'enregistrement et de restitution sont identiques— élaborée

par l'extrapolation de ces mesures en d'autres points de l'espace, ou bien créées artificiellement. Ainsi, nous retiendrons que le formalisme de la méthode est général : il pourra aussi bien s'appliquer à notre étude, qui est un cas particulier, que pour le cas du système de restitution idéal mentionné à la partie B.5. En effet, les K signaux de contrôle pourraient modéliser le caractère étendu de la source, et la matrice d'extrapolation \mathbf{A} contrôlerait l'effet de salle à restituer. La seule nouveauté est de modéliser les sources sonores initiales par plusieurs signaux, ce qui va nous entraîner dans un contexte multivoies. Pour visualiser les changements que cela apporte, nous traiterons les deux cas : le cas où la signature temporelle de la source acoustique est représentée par un signal unique et le cas où la signature temporelle de la source est représentée par un ensemble de signaux.

C.3.1 Signature temporelle unidimensionnelle

Pour ce cas, nous supposons que nous disposons d'un unique signal pour représenter la signature temporelle de la source, ce qui permet uniquement de simuler le comportement de sources ponctuelles comme nous l'avons mentionné à la partie B.2.1. Ainsi, l'équation 2.3 se résume à :

$$\begin{pmatrix} s_1(n) \\ \vdots \\ s_K(n) \end{pmatrix} = \begin{pmatrix} a_1(n) \\ \vdots \\ a_K(n) \end{pmatrix} \star x(n) \quad (2.4)$$

Le vecteur A caractérise la propagation virtuelle entre la position de la source que nous souhaitons implanter et chacun des K points de contrôle du champ. Il reflète l'effet de salle que nous souhaitons implanter. Pour obtenir le champ escompté à partir de cet enregistrement, nous le filtrons par J filtres différents dont les sorties alimentent chacune un transducteur. Toutes ces manipulations sont représentées sous forme d'un schéma bloc sur la figure 2.2.

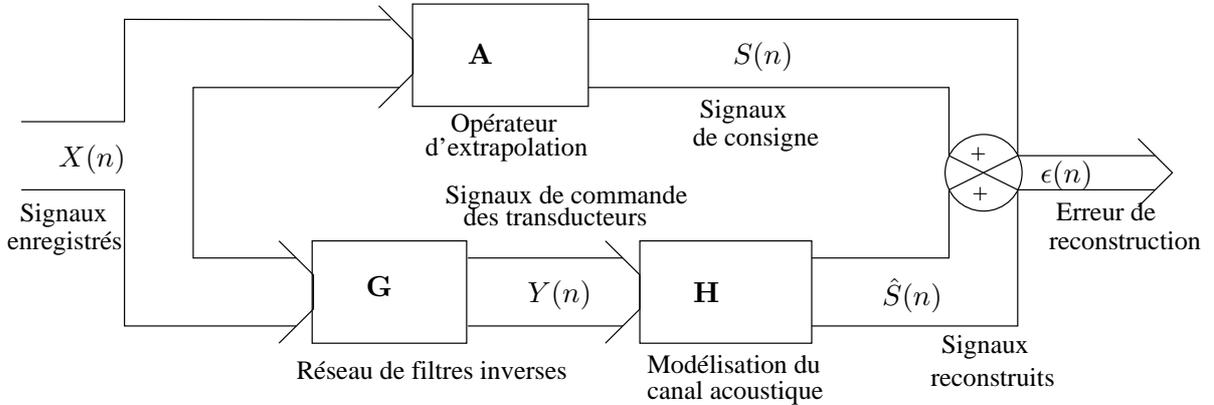


FIG. 2.2 – Représentation du système sous forme de schéma bloc

$$\begin{pmatrix} y_1(n) \\ \vdots \\ y_J(n) \end{pmatrix} = \begin{pmatrix} g_1(n) \\ \vdots \\ g_J(n) \end{pmatrix} \star x(n) \quad (2.5)$$

Nous voulons que le champ reconstruit par le réseau de transducteurs soit identique au champ

de consigne, il faut donc que :

$$\begin{pmatrix} \hat{s}_1(n) \\ \vdots \\ \hat{s}_K(n) \end{pmatrix} = \mathbf{H} \star \begin{pmatrix} g_1(n) \\ \vdots \\ g_J(n) \end{pmatrix} \star x(n) = \begin{pmatrix} a_1(n) \\ \vdots \\ a_K(n) \end{pmatrix} \star x(n) = \begin{pmatrix} s_1(n) \\ \vdots \\ s_K(n) \end{pmatrix} \quad (2.6)$$

Nous sommes donc capables de nous affranchir de l'entrée $x(n)$, et nous pouvons écrire de manière symbolique :

$$\begin{pmatrix} g_1(n) \\ \vdots \\ g_J(n) \end{pmatrix} = \mathbf{H}^{-1} \star \begin{pmatrix} a_1(n) \\ \vdots \\ a_K(n) \end{pmatrix} \quad (2.7)$$

où \mathbf{H}^{-1} désigne symboliquement l'opération d'inversion du canal acoustique entre chaque transducteur j et chaque position k . Pour réaliser cette inversion de canal acoustique, il est d'abord nécessaire dans un premier temps de le caractériser. Une série de mesures a été mise en place afin de calculer chacune des réponses impulsionnelles dans un cas expérimental, et c'est l'objet de l'annexe A de développer quelles sont les méthodes disponibles pour extraire les réponses impulsionnelles d'un système dans le domaine de l'acoustique.

Pour implanter ce résultat en pratique, même si les détails de cette opération seront décrits dans la prochaine partie, nous pouvons d'ores et déjà annoncer qu'il existe deux stratégies possibles :

- Inverser directement la matrice \mathbf{H} , si elle est carrée ($J = K$) ou calculer sa pseudo-inverse. Cette méthode peut demander un temps de calcul assez conséquent, dépendant de la taille de la matrice \mathbf{H} , mais une fois cette opération faite, il ne reste plus qu'une simple convolution à calculer. En pratique, la taille des réponses impulsionnelles est trop importante pour s'orienter vers cette solution, et on utilise alors la deuxième.
- Calculer par des méthodes adaptatives les coefficients des filtres inverses, ce qui demande un temps de calcul moindre que la méthode précédente. En revanche, les coefficients calculés pour le filtre inverse prennent en compte automatiquement les $(a_1(n) \dots a_K(n))^T$, et il est nécessaire de recalculer ces filtres pour chaque position de la source que l'on souhaite restituer. On constitue alors une banque de filtres inverses, dépendant de la position de la source à simuler, ce qui est analogue au principe de restitution basé sur les HRTF.

C.3.2 Signature temporelle multidimensionnelle

Nous considérons maintenant que nous disposons de I enregistrements pour caractériser les sources sonores. Nous élaborons les K signaux de contrôle à partir de ces I enregistrements à partir de l'équation 2.3. De même, pour obtenir au mieux le champ escompté en ces K points de contrôle, nous mettons en place un réseau de filtres en aval des signaux enregistrés, toujours conformément au schéma bloc de la figure 2.2, donnant en sortie les signaux de consigne devant alimenter les transducteurs. Cela est résumé dans l'équation suivante :

$$\begin{pmatrix} y_1(n) \\ \vdots \\ y_J(n) \end{pmatrix} = \begin{bmatrix} g_{11}(n) & \dots & g_{1I}(n) \\ \vdots & \ddots & \vdots \\ g_{J1}(n) & \dots & g_{JI}(n) \end{bmatrix} \star \begin{pmatrix} x_1(n) \\ \vdots \\ x_I(n) \end{pmatrix} \quad (2.8)$$

Nous souhaitons que le champ reconstitué par le réseau de transducteurs soit égal au champ de consignes, ce qui se traduit par l'équation :

$$\hat{S}(n) = \mathbf{H} \star \mathbf{G} \star X(n) = \mathbf{A} \star X(n) = S(n) \quad (2.9)$$

Dans l'équation précédente, nous pouvons simplifier par $X(n)$, et s'affranchir ainsi du signal d'entrée. De manière symbolique, nous obtenons notre réseau de filtre inverse toujours en inversant le canal acoustique \mathbf{H} :

$$\mathbf{G} = \mathbf{H}^{-1} \star \mathbf{A} \quad (2.10)$$

C'est l'objectif de la prochaine section de donner les différentes méthodes possibles pour calculer ce réseau de filtres inverses.

remarque : Il peut être utile de plonger le cas monovoie dans un contexte multivoies. En effet, supposons que les K signaux extrapolés à partir de $x(n)$ soient indépendants et constituent les $K = I$ signaux de référence d'un enregistrement multivoies. Dans ce cas, nous pouvons faire l'hypothèse que la matrice d'extrapolation \mathbf{A} est égale à l'identité. La méthode permet d'inverser le canal acoustique et de calculer un réseau de filtres inverses, et la valeur ajoutée du procédé réside dans le fait que ce réseau de filtres est indépendant de la position de la source que l'on souhaite restituer, contrairement aux conclusions faites à la partie précédente. Ainsi, il n'est plus nécessaire de constituer une banque de filtres contenant un banc de filtres pour chaque position à restituer. Ce procédé devient avantageux à partir du moment où l'on veut simuler plus de $K = I$ positions différentes, car cela nécessiterait le calcul d'un nombre plus important de filtres que si l'on plonge le problème monovoie dans un contexte multivoies. En revanche, il existe quand même un revers à la médaille, c'est qu'il est nécessaire de filtrer $J \times K$ signaux au lieu de J pour tout enregistrement. Il reste à déterminer quelle serait la meilleure stratégie, mais pour cela, il est d'abord nécessaire de donner une solution à ce problème de filtrage de signaux multidimensionnels.

D Méthodes de calcul des filtres inverses

Dans la section précédente, nous avons présenté, de manière globale et sans rentrer dans les détails, les différentes méthodes permettant d'obtenir le champ escompté en nos K points de contrôle à partir de la connaissance de I signaux de référence, et nous sommes aboutis à la conclusion qu'il était nécessaire, dans tous les cas, d'inverser le canal acoustique du lieu d'écoute dans lequel se situe le système de restitution. C'est l'objectif de cette partie de voir quelles méthodes sont à notre disposition pour pouvoir atteindre ce but. Pour structurer la présentation, nous présenterons d'abord sur un cas simple (mono-entrée et mono-sortie) en quoi l'inversion d'un canal acoustique, qui est une opération de déconvolution, n'est pas chose aisée. Ensuite, nous présenterons les premières méthodes qui viennent à l'esprit lorsque l'on souhaite inverser le canal acoustique, et nous verrons pourquoi elles ne peuvent pas s'appliquer en pratique. Finalement, nous mentionnerons la solution fournie à ce problème par des méthodes adaptatives, qui permettent de calculer la solution de manière itérative, sans avoir à inverser directement la matrice.

D.1 Préliminaires : Complexité de l'inversion d'un canal acoustique

L'objectif de cette partie est de montrer, sur un exemple élémentaire, que le problème de l'inversion d'un canal acoustique n'est pas simple. Nous allons considérer le cas où nous souhaitons contrôler le champ acoustique en un point à l'aide d'un seul transducteur. Pour y parvenir, nous supposerons connue la réponse impulsionnelle entre le point d'émission et le point de réception. Cette réponse impulsionnelle tient compte des caractéristiques du système de restitution, à savoir l'enceinte et l'amplificateur, de l'effet de salle, mais aussi du capteur utilisé au moment de l'identification. Nous ne mesurons donc pas la réponse du canal acoustique à proprement parler, car la mesure est fortement tributaire du matériel utilisé lors de l'acquisition. Il n'est pas gênant

que les caractéristiques du système de transduction soient prises en compte au moment de la mesure, car cela fait partie du processus d'égalisation expliqué à la section B.4. En revanche, il serait souhaitable que la mesure effectuée soit indépendante du capteur utilisé, il existe pour cela deux solutions :

- Utiliser pour toutes les mesures des microphones ayant les mêmes caractéristiques, ce qui est envisageable pour des travaux de recherche, mais pas pour un développement à grande échelle de ce procédé.
- Faire une étape de calibration au préalable des microphones utilisés lors de la séance de mesure en vue de s'affranchir de leurs caractéristiques. C'est cette méthode qu'il faudra utiliser si le procédé doit se développer.

A la fin du processus d'identification, explicité dans l'annexe A, nous obtenons une réponse impulsionnelle d'une longueur de N échantillons. La modélisation la plus naturelle qui vient à l'esprit est d'utiliser un filtre à réponse impulsionnelle finie (RIF), et c'est celle qui est retenue dans tous les articles cités en bibliographie. Nous utiliserons aussi bien la représentation temporelle de cette réponse impulsionnelle $h(n)$, que sa transformée en z :

$$H(z) = \sum_{i=1}^N h(n)z^{-n} \quad (2.11)$$

La solution qui vient immédiatement à l'esprit pour réaliser l'inversion de cette réponse impulsionnelle est d'utiliser un filtre dont la transformée en z est la suivante :

$$G(z) = \frac{1}{\sum_{i=1}^N h(n)z^{-n}} \quad (2.12)$$

Il s'agit d'un filtre à réponse impulsionnelle infinie. Les cours de traitement de signal numérique nous enseignent que le filtre représenté par $G(z)$ possède un développement causal stable si et seulement si tous ses pôles se situent rigoureusement à l'intérieur du cercle trigonométrique. Ce n'est souvent pas le cas pour les canaux acoustiques qui possèdent généralement la propriété d'être à non minimum de phase, c'est-à-dire que la réponse impulsionnelle finie modélisant la réponse impulsionnelle du canal possède des zéros à l'extérieur du cercle trigonométrique [29], [27], [33]. Les zéros de la réponse impulsionnelle $h(n)$ sont transformés en pôles pour le filtre inverse, qui ne possède alors pas de développement causal stable. Peut-on faire quelque chose pour pallier ce problème? La réponse est oui, et elle est explicitée en détails dans [29]. Il faut dans un premier temps décomposer la réponse impulsionnelle en deux composantes, la composante à minimum de phase, et le reste, appelé composante à maximum de phase :

$$\begin{aligned} G(z) &= \prod_{i=1}^k (1 - a_i z^{-1}) \prod_{i=k+1}^{N-1} (1 - a_i z^{-1}) \quad , \text{ où } \begin{cases} |a_i| < 1, & i = 1, 2, \dots, k \\ |a_i| > 1, & i = k + 1, k + 2, \dots, N - 1 \end{cases} \\ &= G_{min}(z) G_{max}(z) \end{aligned} \quad (2.13)$$

A la partie $G_{min}(z)$ correspond un filtre causal stable, tandis qu'à la partie $G_{max}(z)$ correspond un filtre anticausal stable. Le filtre inverse n'est donc pas causal, mais il existe. Il est explicité par la relation suivante :

$$g(n) = g_{min}(n) \star g_{max}(n) \quad (2.14)$$

En pratique, il est possible de manipuler un tel filtre en tronquant sa réponse impulsionnelle à un certain ordre, on obtient alors un filtre à réponse impulsionnelle finie, non causal. On le rend

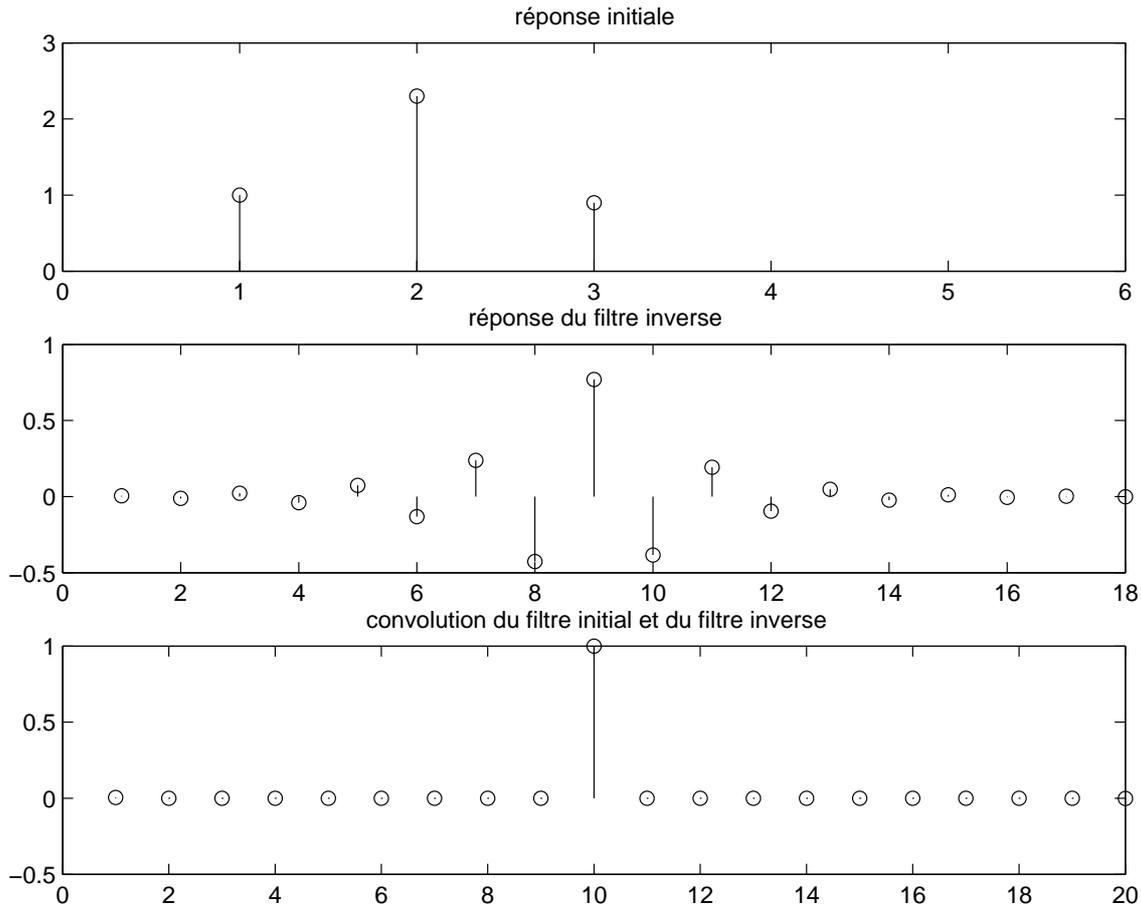


FIG. 2.3 – Inversion d’une réponse impulsionnelle mixte

causal en décalant les échantillons de sa réponse impulsionnelle dans le temps, ce qui correspond finalement à un retard de groupe constant. L’impulsion résultante de la convolution du filtre direct et du filtre inverse sera simplement décalée d’un retard de L échantillons par rapport à l’origine, où L désigne la longueur du filtre anticausal tronqué. Appliquons ce traitement dans un cas simple où la réponse impulsionnelle possède deux zéros : $z_1 = -0.5$ et $z_2 = -1.8$. Elle est donc mixte, car elle possède à la fois une partie à minimum de phase et une partie à maximum de phase. Nous avons tronqué la réponse du filtre acausal à 9 coefficients, celle du filtre causal à 10 coefficients. Le filtre inverse est la convolution des deux précédents, il possède 18 coefficients. Le résultat de la convolution entre la réponse impulsionnelle initiale et le filtre inverse est montré sur la figure 2.3. On observe que l’impulsion est bien restituée, à un délai prêt, qui dépend de la longueur de la troncature de la partie anticausale du filtre inverse, soit ici 9 échantillons, et en négligeant les effets de bord —la valeur des autres échantillons n’est pas rigoureusement nulle.

Cet exemple nous enseigne quelque chose qui sera général par la suite : nous obtiendrons de meilleurs résultats dans le calcul des filtres inverses si nous choisissons de restituer au final une impulsion décalée, ce qui permet de prendre en compte une certaine partie de la fraction anticausale du filtre inverse réel, sans décalage, et ce traitement sera toujours fait par la suite. De plus on s’aperçoit que cette méthode fonctionne bien, alors pourquoi ne pas la généraliser ? Il est impossible de la généraliser pour la bonne et simple raison qu’il est impossible d’évaluer avec suffisamment de précision tous les zéros d’un RIF réel lorsque l’ordre devient élevé, ce qui fait diverger la méthode décrite ici. C’est pour cela qu’on a considéré un cas simple où il n’y a que

deux zéros. Typiquement, une réponse impulsionnelle d'une durée de 1 seconde échantillonnée à la fréquence 44.1 kHz contiendrait 44100 échantillons. La méthode n'est plus fiable pour des RIF de longueur de l'ordre de quelques dizaines d'échantillons, donc ce n'est pas envisageable. Cette section servait juste à présenter le problème.

D.2 Méthode directe d'inversion du canal acoustique

Si nous considérons toujours un cas mono-entrée et mono-sortie, l'opération d'inversion se résume à une opération de déconvolution. Nous avons la relation initiale $s(n) = h \star y(n)$ et le signal de commande est calculé à partir de l'enregistrement initial par $y(n) = g \star x(n)$. Le signal désiré $s(n)$ est identique au signal $x(n)$ à un délai prêt, nous cherchons donc à résoudre l'équation suivante :

$$x(n - \Delta) = h \star g \star x(n) \Rightarrow \delta(n - \Delta) = h \star g(n) \quad (2.15)$$

Soit N la longueur de la réponse impulsionnelle h , et L le nombre de coefficients souhaités pour le filtre inverse. Nous pouvons poser matriciellement le problème précédent :

$$N + L - 1 \left\{ \left(\begin{array}{c} \Delta - 1 \left\{ \begin{array}{c} 0 \\ \vdots \\ 0 \end{array} \right. \\ 1 \\ 0 \\ \vdots \\ 0 \end{array} \right) \right\} = \begin{bmatrix} h(0) & 0 & \dots & 0 \\ h(1) & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \vdots & & \ddots & h(0) \\ \vdots & & & h(1) \\ h(N-1) & & & \vdots \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & h(N-1) \end{bmatrix} \begin{pmatrix} g(0) \\ \vdots \\ g(L-1) \end{pmatrix} \quad (2.16)$$

Il s'agit d'un système de $N + L - 1$ équation à L inconnues. Il n'existe donc pas de solution exacte à ce problème. On utilise une méthode de résolution approchée, par moindres carrés, afin d'obtenir la solution de ce problème qui minimise l'erreur quadratique :

$$G = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \delta(n - \Delta) \quad (2.17)$$

Ce calcul nécessite l'inversion de la matrice $\mathbf{H}^T \mathbf{H}$, qui est de dimension $L \times L$, ce qui demande $L^3/2$ opérations si la matrice est de structure quelconque. Ici, la matrice à inverser est de structure toeplitz, et il existe des algorithmes d'inversion rapide qui requièrent L^2 opérations seulement. Si nous calculons l'erreur correspondante à la solution optimale G , nous obtenons :

$$\epsilon = (\mathbf{I} - \mathbf{H}\mathbf{G}) \delta(n - \Delta) = (\mathbf{I} - \mathbf{H}(\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T) \delta(n - \Delta) \quad (2.18)$$

Ainsi, nous pouvons déterminer le délai Δ optimal qui minimise l'énergie de l'erreur. Pour cela, il suffit de calculer $\epsilon^T \epsilon$, et on vérifie aisément que l'erreur minimale est obtenue pour $\Delta = k$ où k désigne l'indice du terme minimal situé sur la diagonale de $(\mathbf{I} - \mathbf{H}\mathbf{G})$. La question qui reste en suspens est la détermination de l'ordre L du filtre inverse à utiliser. Le plus simple est encore une fois d'utiliser un critère énergétique : si nous supposons que le terme d'erreur est perçu comme un bruit supplémentaire, on pourrait prendre un rapport signal/bruit d'au moins 20 dB par exemple, et souhaiter ainsi que l'énergie de l'erreur soit inférieure à 0.01. Il faudra ensuite tester perceptivement la validité d'un tel critère.

Un tel traitement a été effectué pour une réponse impulsionnelle mesurée lors de la campagne de mesures. La longueur initiale est de 256 échantillons, ce qui correspond à 5 ms pour une fréquence d'échantillonnage de 44.1 kHz. C'est certes ridicule de faire cette manipulation, car la réponse impulsionnelle a été très fortement tronquée, mais ça allège le temps de calcul, pour l'inversion. Et puis la réponse impulsionnelle a été effectuée en chambre sourde, donc elle est normalement très courte.

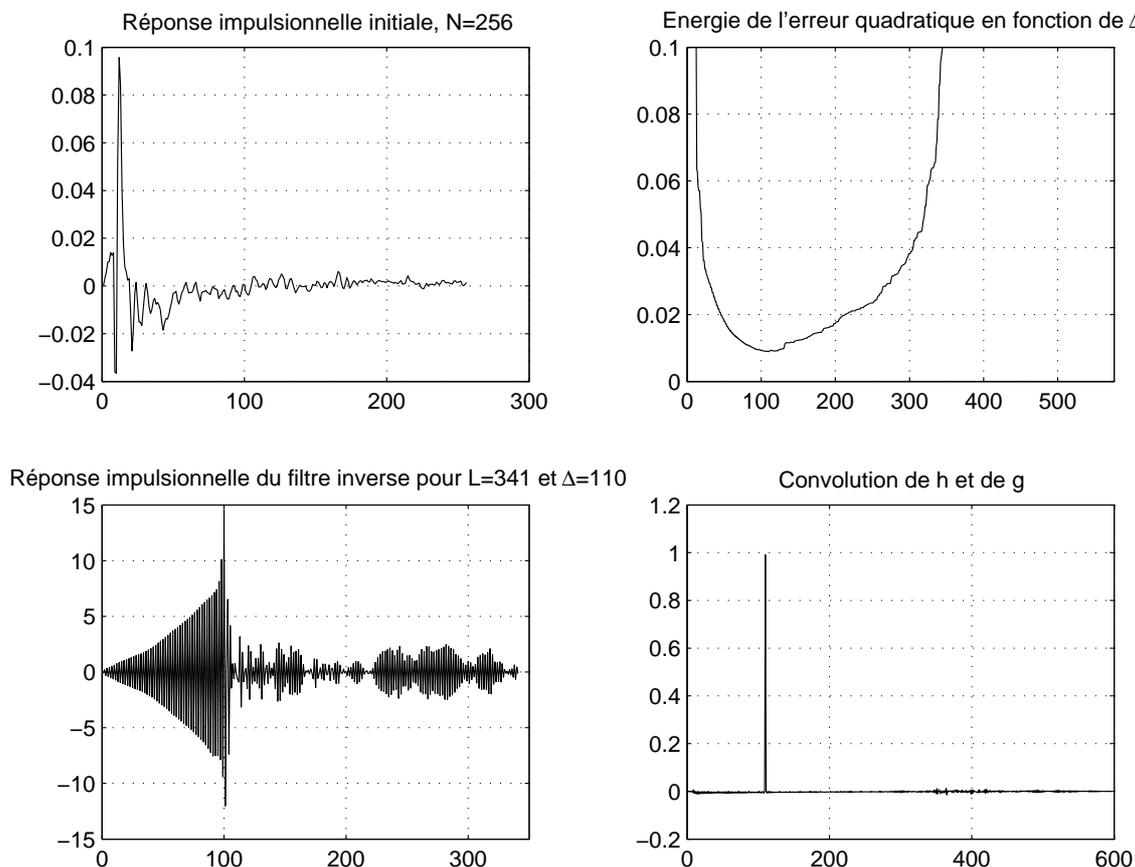


FIG. 2.4 – Inversion d'une réponse impulsionnelle par la méthode des moindres carrés

On a représenté sur la figure 2.4 la réponse impulsionnelle utilisée pour le traitement. Juste à côté est représenté l'énergie du signal d'erreur en fonction du délai Δ que l'on souhaite imposer au système. On observe que celle-ci est minimale pour $\Delta = 110$, et qu'elle est inférieure à 0.01, on a choisi l'ordre L en conséquence pour remplir ce critère. Le filtre correspondant est tracé dans le troisième cadre, et la convolution entre la réponse impulsionnelle initiale et le filtre inverse est tracée dans le dernier cadre. Pour mieux visualiser les performances de cette égalisation, on a tracé le module de la transformée de Fourier de l'impulsion restituée sur la figure 2.5. On observe sur cette figure que les performances de l'égalisation se situent dans la bande $\pm 1dB$ à partir du moment où la fréquence se situe entre 100 Hz, et 10 kHz. Pour mieux visualiser les défauts du système d'égalisation, nous avons aussi tracé sur la figure 2.6 le module de la TF de l'erreur, et le retard de groupe correspondant. Cette figure permet de voir le niveau de l'erreur en dB par rapport à celui du signal initial. On observe que celui-ci est en dessous de 20 dB pour des fréquences supérieures à 300 Hz, et ce, sur une large bande de fréquences. Le retard de groupe permet de quantifier le temps de retard de chaque composante fréquentielle de l'erreur par rapport au signal initial.

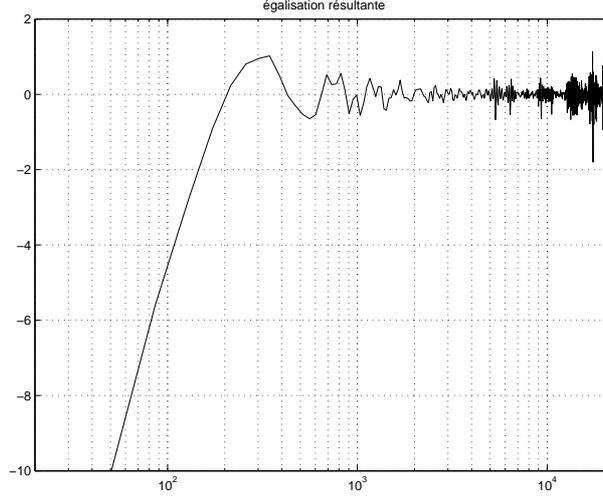


FIG. 2.5 – Module de la Transformée (en dB) de Fourier de l’impulsion restituée en fonction de la fréquence (en Hz)

Maintenant, essayons de prospecter les performances de cette méthode pour des systèmes plus ‘réels’. On peut dans une première approximation considérer que le temps de réverbération moyen d’une salle est de une seconde, ce qui représente 44100 échantillons. Il faut donc créer une matrice \mathbf{H} de dimension $44100 + L - 1 \times L$, ce qui demande des capacités de stockage assez conséquentes déjà et un temps de création non négligeable. Ce qui fait que cette méthode n’est pas envisageable pour des réponses impulsionnelles longues. Néanmoins, une idée pour s’en sortir serait de placer un banc de filtres en entrée du dispositif qui viendrait étudier le signal par bandes de fréquences successives, ce qui autoriserait alors de ralentir la cadence d’échantillonnage dans chaque sous-bande du signal [34]. Il ne resterait plus qu’à appliquer l’algorithme dans chaque sous-bande, ce qui accélérerait le calcul d’un ordre de grandeur normalement, mais cela reste à vérifier.

Maintenant, établissons une version multidimensionnelle de cet algorithme. Il suffit juste d’écrire l’équation 2.15 sous forme matricielle dans un contexte multidimensionnel :

$$\begin{bmatrix} \delta(n - \Delta) & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \delta(n - \Delta) \end{bmatrix} = \begin{bmatrix} \mathbf{H}_{11} & \dots & \mathbf{H}_{1J} \\ \vdots & \ddots & \vdots \\ \mathbf{H}_{K1} & \dots & \mathbf{H}_{KJ} \end{bmatrix} \begin{bmatrix} G_{11} & \dots & G_{1K} \\ \vdots & \ddots & \vdots \\ G_{J1} & \dots & G_{JK} \end{bmatrix}$$

où

$$\delta(n - \Delta) = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{H}_{kj} = \begin{bmatrix} h_{kj}(0) & 0 & \dots & 0 \\ h_{kj}(1) & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \vdots & & \ddots & h_{kj}(0) \\ \vdots & & & h_{kj}(1) \\ h_{kj}(N-1) & & & \vdots \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & h_{kj}(N-1) \end{bmatrix}, \quad G_{kj} = \begin{pmatrix} g_{jk}(0) \\ \vdots \\ g_{jk}(L-1) \end{pmatrix} \quad (2.19)$$

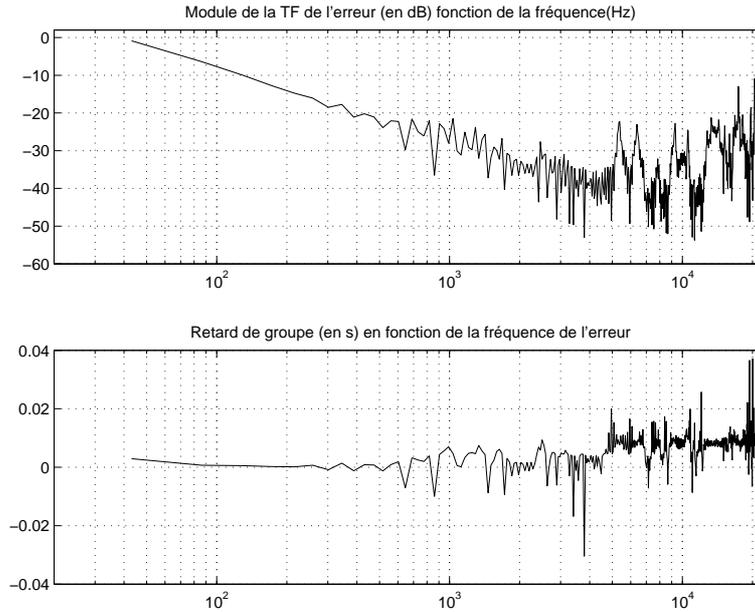


FIG. 2.6 – Caractérisation de l'erreur du système d'égalisation dans le domaine fréquentiel : module de la TF de l'erreur (en dB) en fonction de la fréquence (en Hz)

On peut résumer la dernière équation de manière synthétique :

$$\delta(\mathbf{n} - \mathbf{\Delta}) = \mathbf{H}\mathbf{G} \quad (2.20)$$

où $\delta(\mathbf{n} - \mathbf{\Delta})$ désigne une matrice $K \cdot (N + L - 1) \times K$, \mathbf{H} désigne une matrice $K \cdot (N + L - 1) \times J \cdot L$ et où \mathbf{G} désigne une matrice $J \cdot L \times K$. La solution est toujours donnée par la formule 2.17, mais la matrice \mathbf{H} n'est malheureusement pas de structure Toeplitz, ni Toeplitz par bloc, ce qui empêche d'utiliser les algorithmes rapides de résolution, type Levinson-Whittle-Robinson. Sous certaines hypothèses, il est possible de calculer l'inverse exacte de la matrice \mathbf{H} , ce qui est possible si le nombre de points de contrôles K est strictement inférieur au nombre de transducteurs J , cette méthode porte alors le nom de *multiple input/output inverse theorem* (MINT), d'après ses créateurs, Miyoshi et Kaneda [26],[27]. Pour calculer une inverse directe, il suffit que la matrice \mathbf{H} soit bien conditionnée d'une part, et carrée d'autre part, ce qui est le cas si la relation suivante est vérifiée : $K(N + L - 1) = JL$. En pratique, on utilise un point de contrôle de moins que de transducteurs, $K = J - 1$, et donc, pour une réponse impulsionnelle de longueur N , il faut considérer des filtres inverses dont la longueur est $L = (J - 1)(N - 1)$, [30]. Autant dire qu'il s'agit d'une formulation théorique du problème, mais qu'il est impossible d'implanter directement cette solution dans la pratique. En effet, nous disposons de $J = 24$ transducteurs, si nous voulons imposer $K = 23$ contraintes pour le champ sonore, avec $N = 44100$ échantillons par réponse impulsionnelle, il nous faut alors plus d'un million de coefficients par filtre inverse, qui nécessitent l'inversion d'une matrice 24342648×24342648 , ce qui est utopique à réaliser en pratique.

Pour la beauté du geste, appliquons la méthode MINT, qui permet d'accomplir de manière exacte l'inversion du canal acoustique, qu'il soit à minimum de phase ou non, dans le cas de systèmes transauraux : nous voulons dans ce cas reproduire à l'identique 2 signaux de consignes à partir de 3 transducteurs. Nous pouvons voir sur la figure 2.7 les réponses impulsionnelles du canal acoustique entre les trois transducteurs et les deux capteurs. Elles sont au nombre de 6, et possèdent 201 échantillons chacune. Nous pouvons voir que les temps d'arrivée sur chaque microphone ne sont pas identiques, en témoignent les différents temps d'arrivée sur chaque

graphe du pic principal. Elles ne sont pas de même niveau non plus, étant donné que les points ne sont pas équidistants de chaque transducteur.

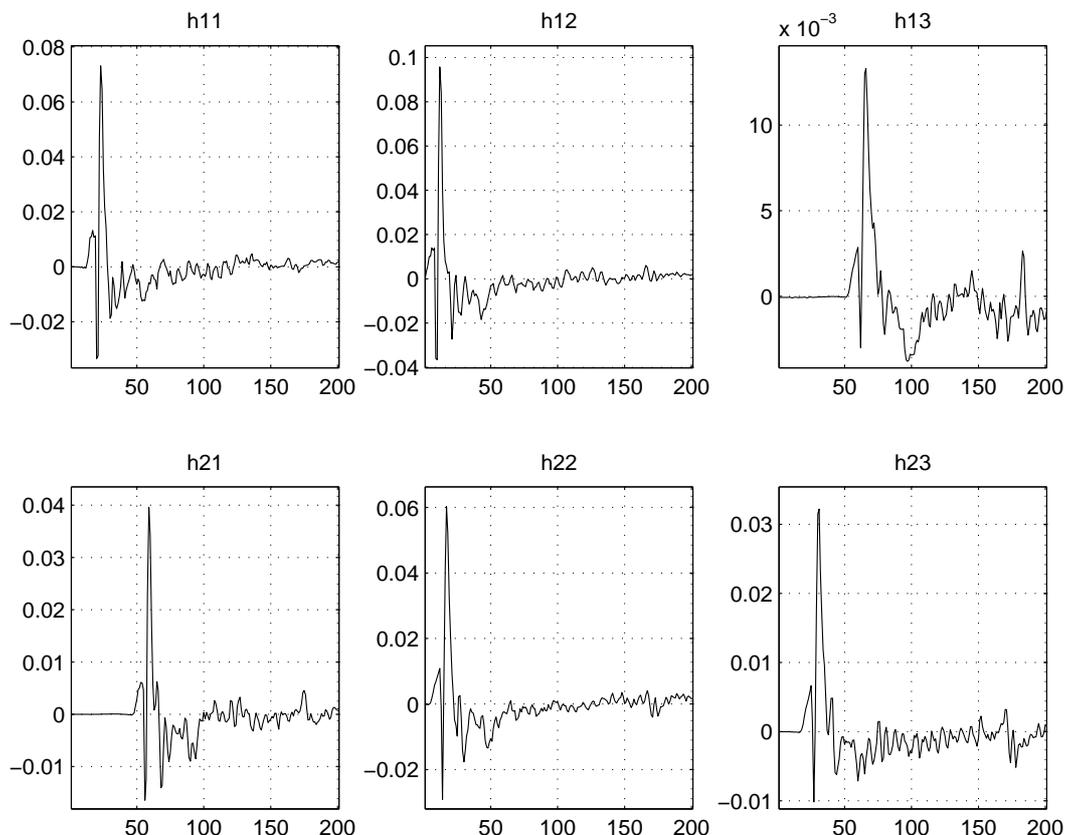


FIG. 2.7 – Visualisation de l’amplitude de la réponse impulsionnelle (unité arbitraire) des 3 transducteurs vers les 2 microphones en fonction du numéro de l’échantillon ($F_e = 44100Hz$)

Les réponses impulsionnelles étant de longueur $N = 201$ échantillons, il faut $L = 2 * 200 = 400$ coefficients pour chaque filtre inverse d’après la formule précédente. La dimension de la matrice carrée \mathbf{H} à inverser est 1200. Une fois la matrice \mathbf{H} inversée, on obtient 600 candidats possibles de réseaux de filtres inverses dépendant du retard de l’impulsion que l’on souhaite restituée. On choisit parmi ces candidats, qui réalisent tous l’inversion parfaite, celui qui nécessite le moins d’énergie (pour cela, on somme les énergie des réponses impulsionnelles de chacun des filtres du réseau de filtres inverses). Le résultat est montré sur la figure 2.8. On observe que l’énergie du réseau est minimale pour $\Delta = 281$ échantillons. Le réseau de filtres inverses correspondant à cette valeur de Δ est affiché à la figure 2.9

Le résultat du procédé transaural est de restituer le premier signal d’un enregistrement stéréophonique en un point et le deuxième signal en un deuxième point, et éliminer tous les trajets croisés, procédé connu dans la littérature sous le nom de *cross-talk cancellation*. Les résultats de ce procédé sont affichés sur la figure 2.10, où l’on voit que la restitution est visiblement parfaite. Pour mieux quantifier l’erreur, nous l’avons affiché à la figure 2.11.

Nous voyons que cette erreur est limitée en amplitude à 10^{-9} , ce qui est très faible, et seulement dû aux imprécisions du calcul numérique lors de l’inversion, qui dépend du bon ou du mauvais conditionnement de la matrice \mathbf{H} . La valeur maximale de l’énergie de ces signaux est en 10^{-16} ce qui correspond à un rapport signal/bruit de 160 dB. En pratique, cela serait inaudible, mais cette précision numérique, et ce rapport signal/bruit ‘théorique’ serait dégradé à cause

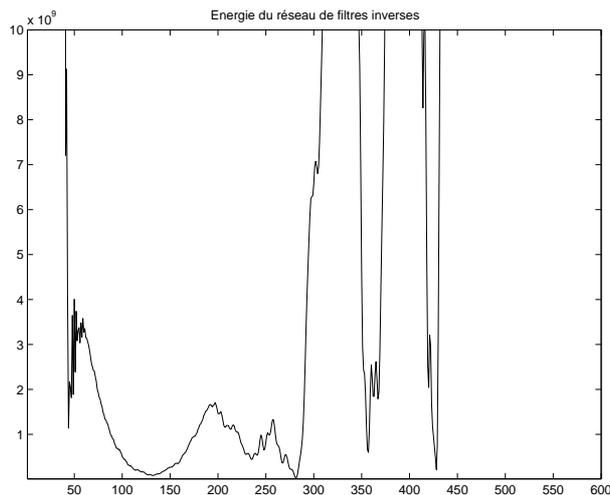


FIG. 2.8 – Visualisation de l'énergie du filtre inverse

de l'imprécision sur la mesure des réponses impulsionnelles, et notamment à cause des non-linéarités, mais aussi sur le fait que nous n'avons considéré qu'un temps très court, seulement 200 échantillons, pour modéliser une réponse impulsionnelle beaucoup plus longue.

D.3 Approche du filtrage inverse par des méthodes adaptatives : la solution optimale ?

Cette possibilité d'approche permet de calculer les coefficients de chaque filtre inverse sans avoir à inverser la matrice \mathbf{H} dont il est question jusqu'à présent. Le principe de cette méthode est le suivant :

- On commence par initialiser les coefficients de chaque filtre par une certaine valeur, choisie de manière arbitraire ou non.
- L'erreur entre le champ sonore de consigne et celui effectivement restitué est ensuite calculée
- On calcule ensuite le gradient de cette erreur vis-à-vis de chacun des coefficients de tous les filtres inverses ou toute autre fonction permettant la réactualisation des coefficients de chaque filtre inverse à l'itération suivante. De manière générale, le gradient nous renseigne sur la direction de la ligne de plus grande pente. Si on se dirige dans la direction du gradient, l'erreur va augmenter. Cela nous mène à la prochaine étape
- Les coefficients de chaque filtre inverse sont ensuite réactualisés à la prochaine étape. Dans la méthode du gradient, on réactualise chaque coefficient dans une direction opposée à celle du gradient, de sorte à minimiser l'erreur de reconstruction à la prochaine itération. C'est le but général de toute autre fonction de réactualisation, qui est caractéristique de chaque algorithme.
- On réitère le processus, jusqu'à ce que celui-ci converge vers la solution souhaitée. Les critères d'arrêt sont un problème de ce genre d'algorithmes. On peut choisir de stopper les itérations lorsque la norme de l'erreur ne change plus beaucoup après plusieurs itérations, mais si l'algorithme est instable, cela peut être source de boucles infinies, et l'on préfère alors décider de stopper l'algorithme après un certain nombre d'itérations.

Le principe décrit ici est vraiment un principe général, pour comprendre le fonctionnement de la méthode. En pratique, cela a donné naissance à une multitude d'algorithmes adaptatifs, tels que les Least Mean Squares Algorithms (LMS), les Recursive Least Squares Algorithms (RLS), les Fast-Transversal Filters Algorithms (FTF), les Multichannel Affine and Fast Affine Projec-

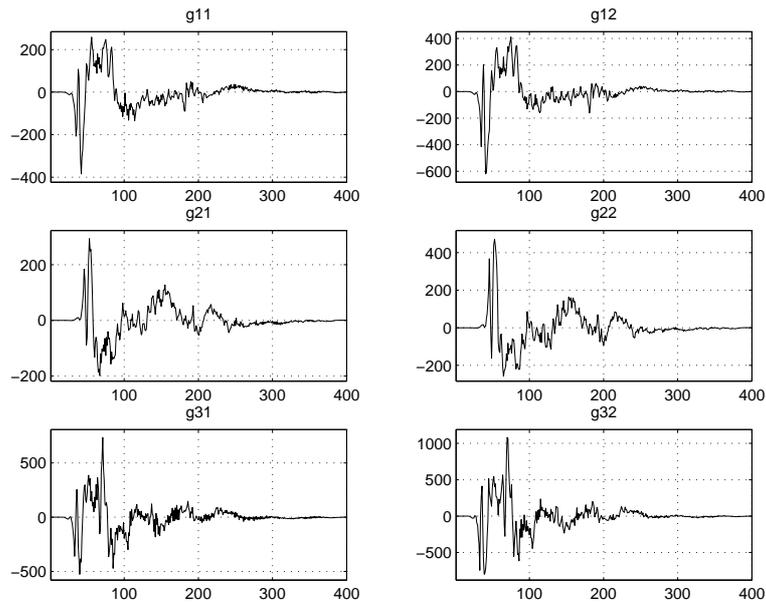


FIG. 2.9 – Réponse impulsionnelle des 6 filtres inverses

tion Algorithms (MAP et MFAP), etc. Ces algorithmes sont décrits en détails dans [35],[36],[39], et [31]. Ils ont déjà été utilisés pour l'égalisation des transducteurs dans le cadre de la WFS, comme relaté dans [37],[38]. Nous n'avons malheureusement pas eu le temps de les mettre en œuvre et de vérifier leurs performances en pratique. Pour tous les algorithmes mentionnés, il est nécessaire de comparer leurs performances, en termes de temps de calcul par itération (qui est généralement évalué en fonction du nombre de multiplications nécessaire pour chaque itération), de vitesse de convergence vers la solution, et de stabilité numérique.

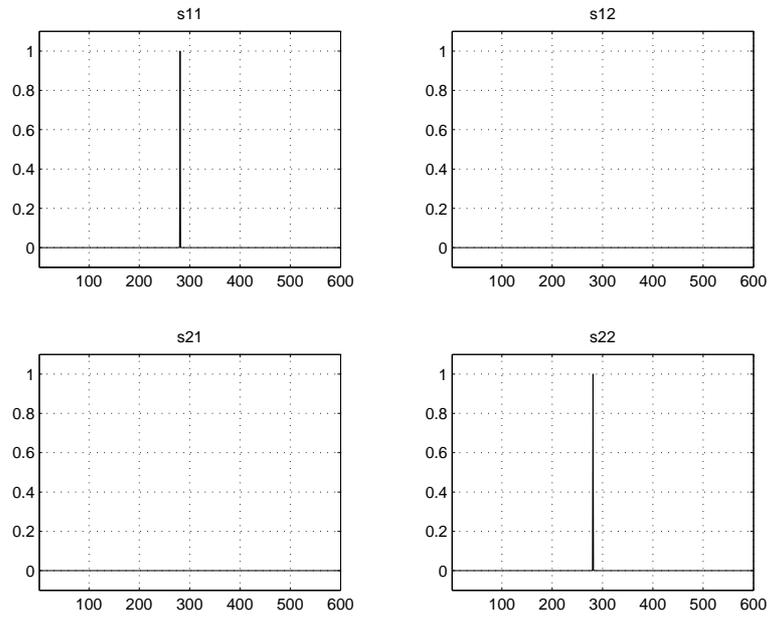


FIG. 2.10 – Résultat du *cross-talk cancellation*

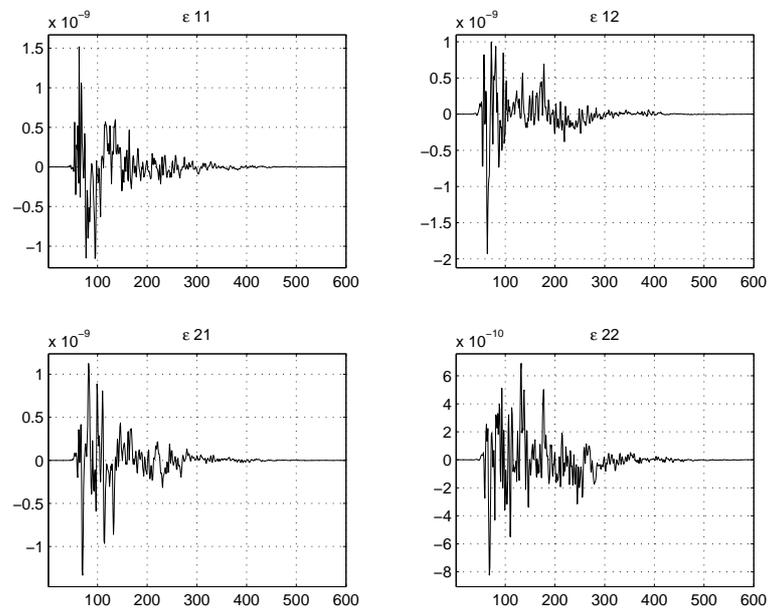


FIG. 2.11 – Visualisation de l'erreur du système d'égalisation

Chapitre 3

Conclusion

A Ce qui a été fait...

Le travail effectué au cours de ce stage a permis en premier lieu d'établir une solide base bibliographique permettant d'élucider certains points qui pouvaient sembler confus dans la littérature, notamment celui des conditions nécessaires et suffisantes permettant de caractériser avec unicité le champ sonore dans un certain volume de l'espace. De plus, cette étude a permis d'envisager, de manière théorique pour le moment, de nouveaux procédés pour la prise de son, et différentes applications de la restitution de champ sonore sur une zone étendue, à savoir la reproduction à l'identique, l'enrichissement ou la correction, ou bien la synthèse virtuelle d'un champ sonore.

Partant de cette étude bibliographique, nous avons commencé à réfléchir sur l'orientation à donner au projet sur la spatialisation des sons sur zone étendue, afin qu'il n'empiète pas trop sur les procédés en cours de développement dans d'autres centres de recherche. Pour cela, nous avons extrait quelques paramètres à prendre en considération dans tout système de reproduction sonore, et l'orientation nous est apparue de manière naturelle : aucun dispositif hormis la stéréophonie ne faisait l'hypothèse que les sources sonores pouvaient être étendues, toutes les autres méthodes visant à restituer des sources ponctuelles, au mieux à directivité variable. Il semblerait naturel pourtant que cette hypothèse soit faite, même si cela doit complexifier les traitements postérieurs. Ainsi, cela nous a permis de fixer les premiers objectifs à remplir pour le futur dispositif de restitution : savoir reproduire à l'identique le champ enregistré par un réseau de microphones dans une certaine salle dans un autre lieu d'écoute, en conservant la même géométrie pour le réseau d'enregistrement, et le réseau de points de contrôle lors de la restitution. Cette méthode d'enregistrement a l'inconvénient de coupler les sources et la salle dans une seule entité, mais le but primaire est d'abord de voir si ce dispositif est capable de donner une bonne image perceptive du relief sonore de la scène initiale, avant d'essayer de le développer de manière plus approfondie, en dissociant la modélisation des sources et celle de l'effet de salle.

Ensuite, nous avons largement décrit le problème auquel nous sommes confrontés, à savoir l'inversion d'un canal acoustique, ce qui nous a conduit à décrire les différentes méthodes possibles pour réaliser cette inversion et ainsi obtenir notre réseau de filtres inverses. Il s'agit d'un problème complexe, car il est plongé dans un contexte multivoies, qu'il est difficile de manipuler aisément. Pour l'instant, des algorithmes sont envisageables, tant que les réponses impulsionnelles à traiter ne sont pas trop longues, et que le nombre de voies n'est pas trop important, ce qui n'est pas forcément le cas dans la pratique. L'objectif est donc de trouver des méthodes permettant de réduire la complexité de design du réseau de filtres inverses, et pour cela, nous mettons tous nos espoirs sur les méthodes adaptatives, que nous n'avons malheureusement pas

encore eu le temps de tester, ce qui sera fait prochainement, dès le début de la poursuite de ce travail en thèse.

B ... Et ce qu'il reste encore à faire

Dans cette partie, nous allons faire le plan des batailles qu'il nous reste encore à livrer :

- Dans un premier temps, il faudra tester la rapidité des algorithmes adaptatifs en comparaison avec l'inversion ou la pseudo-inversion directe du canal acoustique.
- Si cette étape s'avère concluante, je pense néanmoins que le temps de traitement s'avérera quand même assez long, et il sera certainement intéressant de savoir quelle est la longueur minimale nécessaire des réponses impulsionnelles du lieu de restitution à considérer au moment de l'inversion pour obtenir des résultats au niveau perceptif. Au premier abord, il semble au moins nécessaire d'inverser une longueur supérieure au temps de réverbération de la salle, mais c'est à confirmer. Dans tous les cas, on peut douter de l'efficacité des systèmes d'égalisation n'inversant que les 1000 premiers échantillons d'une réponse impulsionnelle échantillonnée à 44.1 kHz, car cela représente une durée de 22 ms environ, et l'oreille n'est capable de dissocier nettement les réflexions d'un son direct à partir de 40 ms environ, sinon elle assimile ce son au son direct.
- Une fois ces deux problèmes réglés, si toutefois cela est possible, il faudra s'intéresser aux effets de l'égalisation du champ sonore en dehors des points de contrôle. Cela a pour but de répondre à la question : vaut-il mieux contrôler le champ de manière imparfaite sur beaucoup de points, ou le contrôler de manière précise en un nombre plus limité de points ? Nous savons que la vitesse d'évolution dans le domaine spatial du champ sonore est de l'ordre de la longueur d'onde. Ainsi, à basses fréquences, les longueurs d'onde sont grandes, et les variations spatiales du champ sont faibles lorsque l'on s'écarte des points d'égalisation. Cela n'est plus le cas à hautes fréquences, où les longueurs d'ondes sont petites, et donc les variations spatiales importantes sur de courtes distances. Ainsi, pour contrôler le champ à hautes fréquences, il est nécessaire d'augmenter la densité de points de contrôle ce qui ne sera pas forcément possible. Il reste à déterminer quelle est la plage de fréquences dans le domaine de l'audible nécessaire à contrôler pour restituer une bonne impression du relief sonore, et espérer qu'elle ne s'étende pas trop dans les hautes fréquences.
- Si tous ces problèmes arrivent à obtenir une solution, nous pourrions commencer à développer des outils permettant l'extrapolation du champ sonore uniquement à partir de la connaissance de la pression sur une surface fermée, comme nous l'avons mentionné dans la partie bibliographique de ce rapport. Le système d'enregistrement ne devra contenir qu'un unique type de microphones, et il ne sera pas nécessaire d'enregistrer le gradient de pression en plus de la pression. De plus, si ces méthodes d'extrapolation sont mises au point, il sera alors possible de commencer à dissocier la modélisation des sources sonores de celle de l'effet de salle, ce qui est un des objectifs à long terme du projet. En effet, nous serions capables d'utiliser la prise de son permettant de caractériser uniquement la source sonore (réalisée en champ proche, ou dans une chambre anéchoïque, voire les deux). Cela nous permettrait d'extrapoler cette source dans n'importe quelle autre zone de l'espace, si nous faisons l'hypothèse de champ libre. Des recherches complémentaires seront nécessaires pour obtenir une modélisation performante de l'effet de salle, ce qui permettra alors de faire évoluer n'importe quelle source dans une acoustique réelle ou virtuelle, quel que soit le lieu de restitution, et ainsi obtenir le système de restitution idéal. D'ici là, il reste encore beaucoup de chemin à parcourir...

Annexe A

Des différentes méthodes de mesure de réponses impulsionnelles en acoustique

Dans beaucoup d'applications, il est nécessaire de connaître la réponse impulsionnelle d'un système acoustique : c'est le cas par exemple lorsqu'un constructeur souhaite caractériser les performances d'un nouveau prototype d'enceinte, avant de le commercialiser, ou bien lorsqu'un acousticien souhaite caractériser certains paramètres perceptifs de l'acoustique d'un lieu d'écoute, tels que le temps de réverbération ou la densité de premières réflexions, qui sont dérivés à partir de la connaissance de la réponse impulsionnelle du canal acoustique entre un point source et un point d'observation.

Cette mesure requière généralement la plus grande précision afin d'obtenir le meilleur rapport signal sur bruit possible, ce qui est rendu nécessaire à cause de la grande plage dynamique du système auditif, d'environ 120 dB. Le codage CD-audio standard sur 16 bits permet d'atteindre des rapports signal sur bruit de l'ordre de 110 dB, il n'est donc pas présomptueux de vouloir demander que la mesure ait le même rapport signal sur bruit [44]. Une multitude de techniques se sont développées pour pouvoir mesurer la réponse impulsionnelle d'un système acoustique, et nous en ferons l'inventaire dans une première partie. Une fois ces techniques répertoriées, nous détaillerons la technique choisie pour mesurer les réponses impulsionnelles dans notre cas, ainsi que les critères qui nous ont orienté vers ce choix bien défini, c'est-à-dire, le recours à des signaux chirp —sinusoïdes balayant une certaine plage de fréquence Δf pendant une durée T —, ce qui permet à la fois de caractériser la réponse impulsionnelle du système, et de s'affranchir, voire même de prendre en compte, les non-linéarités présentes dans le système.

A Les méthodes existantes

Elles possèdent toutes le point commun d'avoir un signal d'entrée —stimulus— contenant toutes les fréquences d'intérêt du système étudié. Il en existe au moins huit, dont certaines sont assez anciennes alors que d'autres sont assez récentes, et toutes n'ont même les mêmes performances :

- L'enregistreur de niveau.
- Les techniques de *Time-Delay Spectrometry* (TDS).
- Les analyseurs FFT à deux voies.
- Excitation du système en régime harmonique.
- Excitation par des impulsions périodiques.

- Utilisation de *Maximum Length Sequences*.
- Utilisation de signaux périodiques de longueur 2^N .
- Excitation du système par des signaux chirp non périodiques.

Ces techniques vous sont présentées tour à tour. Les performances des trois dernières seront confrontées en détails dans la prochaine partie, afin de déterminer laquelle de ces trois techniques est la plus appropriée pour notre série de mesures.

A.1 L'enregistreur de niveau

Il s'agit de la méthode la plus ancienne pour caractériser des réponses impulsionnelles, utilisées principalement pour les enceintes. Le signal d'excitation utilisé par cette méthode était déjà un signal chirp logarithmique, c'est-à-dire un signal dont la fréquence instantanée croît d'un même facteur pour une échelle de temps déterminée, ce qui permet d'injecter la même énergie par octave pour le système à étudier, ce qui est déjà une certaine forme d'optimisation du rapport signal sur bruit, car le bruit ambiant pour les systèmes acoustiques présente une certaine coloration fréquentielle, et sa densité spectrale de puissance contient environ la même densité d'énergie par octave. Ce signal était généré et transmis par un générateur analogique au système d'étude. La réponse du système d'étude à ce signal était filtrée de telle sorte à n'en retenir que l'amplitude de l'enveloppe (filtrage passe-bas), et la sortie de ce filtre venait alimenter un amplificateur différentiel contrôlant la position d'un stylo sur une table traçante. En utilisant une vitesse de défilement constante pour la feuille, des graduations logarithmiques et une synchronisation correctement effectuée, il est ainsi possible d'obtenir un relevé expérimental donnant une indication sur le niveau (en dB) en fonction de la fréquence.

Ces enregistreurs de niveau sont encore fortement utilisés dans les laboratoires car ils sont robustes et permettent rapidement d'obtenir des relevés expérimentaux fiables, sans recourir à une artillerie numérique lourde. Ils sont néanmoins incapables de supprimer le bruit ambiant, ni les réflexions venant parasiter la mesure de réponse impulsionnelles en 'champ libre' pour les enceintes. Mais il est possible d'obtenir une meilleure résolution en diminuant la vitesse de déroulement, permettant ainsi de rétablir le compromis entre la résolution souhaitée, et la durée de la mesure nécessaire, tout comme pour chaque méthode moderne basée sur du traitement de signal numérique. L'inconvénient majeur de cette méthode est qu'elle ne nous fournit aucune information sur la phase du système. De plus, les relevés expérimentaux sur papier sont inconfortables pour des futures manipulations et extrapolations, et les mesures horizontales (en fréquence) ou verticales (pour le niveau en dB) sont fortement imprécises en comparaison avec les autres méthodes qui vont être développées, ce qui est dû au traitement utilisé : générateur analogique et utilisation d'un système supplémentaire permettant d'asservir la position du stylo, mais limitant les performances du système de mesure. Cette méthode renseigne uniquement sur le comportement en amplitude de la fonction de transfert du système acoustique, et il est donc impossible par cette méthode d'extraire la réponse impulsionnelle du système.

A.2 Time-Delay Spectrometry (TDS)

Les méthodes de spectrométrie à retard permettent d'obtenir la totalité de la fonction de transfert (informations sur l'amplitude et la phase), ce qui permet d'accéder à la réponse impulsionnelle, en effectuant la transformée de Fourier inverse de la fonction de transfert. Ses domaines d'utilisation concernent aussi bien la mesure de réponses impulsionnelles d'enceintes que de salles. Cette méthode est basée sur un principe analogue à celui utilisé pour la construction d'analyseurs de spectres analogiques, et des techniques de démodulation en télécommunications analogiques. Deux signaux sont utilisés en entrée : le premier est une 'sinusoïde glissante' (connu

autrement sous le nom de signal chirp linéaire) tandis que le deuxième est une ‘cosinusoïde glissante’, c’est-à-dire la même version que le premier mais dont la phase est décalée d’un angle de 90° , on utilise pour cela un premier générateur et une boucle à verrouillage de phase. Le premier signal permet d’extraire la partie réelle de la fonction de transfert, tandis que le second permet d’en extraire la partie imaginaire. Le premier signal sert d’entrée au système d’étude. La sortie du système est multipliée aux deux signaux d’entrée avant de passer par un filtre passe-bas. Mettons en évidence le mécanisme de fonctionnement de ce traitement lorsque le signal d’entrée est une sinusoïde de fréquence constante (réponse en régime harmonique). Dans ce cas, la sortie du système est une sinusoïde déformée, en amplitude et en phase.

$$\begin{cases} e_1(t) = E \sin(\omega t) \\ e_2(t) = E \cos(\omega t) \\ s(t) = S \sin(\omega t + \phi) \end{cases} \Rightarrow \begin{cases} s(t) \times e_1(t) = \frac{SE}{2} [\cos(\phi) - \cos(2\omega t + \phi)] \\ s(t) \times e_2(t) = \frac{SE}{2} [\sin(\phi) + \sin(2\omega t + \phi)] \end{cases} \quad (\text{A.1})$$

D’après l’équation précédente, nous pouvons en déduire que le module de la fonction de transfert est S/E , et que son argument est ϕ . Il suffit donc d’éliminer les termes de pulsation 2ω dans les produits entre les signaux d’entrée et de sortie, pour obtenir une image de la fonction de transfert, ce qui est réalisé par un filtrage passe-bas, laissant passer la composante continue et filtrant la composante en 2ω . Il ne reste alors plus que la partie réelle et la partie imaginaire de la fonction de transfert, au facteur E près. Ce mécanisme de fonctionnement est toujours valable lorsque le signal d’entrée est une sinusoïde ‘glissante’, mais il va y avoir quelques particularités. En effet, il existe toujours un retard dans la mesure de la fonction de transfert du système d’étude, dû au temps de propagation des ondes sonores entre la source et le point d’observation. Il se peut que la fréquence instantanée entre les signaux reçus sur le capteur et les signaux d’entrée soient assez différentes, et ne rentrent plus dans la bande passante du filtre passe-bas utilisé, c’est pour cela qu’on introduit un retard supplémentaire à chacun des deux signaux d’entrée, égal au temps de trajet des ondes de la source au capteur, avant de les multiplier au signal de sortie. Ainsi, la différence entre les fréquences instantanées des deux signaux d’entrée, et de sortie sera toujours proche de zéro.

Grâce à la présence de ce filtre, il est possible d’ajuster la vitesse de variation de la fréquence du signal initial par rapport à la fréquence de coupure du filtre passe-bas, de sorte que l’on puisse extraire uniquement la réponse impulsionnelle d’une enceinte en champ libre, et atténuer toutes les réflexions qui viennent parasiter une telle mesure. En effet, les réflexions sont séparées du son direct par des délais supplémentaires car les trajets correspondants sont plus longs. Lorsque ces réflexions sont captées par le microphone, il faut ajuster la fréquence de coupure du filtre passe-bas (ou la vitesse de variation de la fréquence du sweep), de sorte que la différence entre les fréquences instantanées du signal d’entrée et celle de la réflexion soit supérieure à la fréquence de coupure du filtre. De même, la fréquence de coupure du filtre passe-bas permet de diminuer l’influence des non-linéarités du système : en effet, la présence de non-linéarités induit l’existence d’harmoniques d’ordre supérieur au fondamental, qui sont situés au minimum au double de la fréquence fondamentale, ils sont ainsi éliminés par le filtrage. De plus la présence de ce filtre permet de diminuer l’influence d’un bruit large bande, car on ne conserve que la composante fréquentielle du bruit nécessaire à la mesure, l’autre étant éliminée par le filtre. C’est pour toutes ces raisons que les signaux d’entrées sont des signaux chirp linéaires, car ils permettent d’éliminer toujours les mêmes réflexions, sur toute la plage de fréquence balayée, ce qui ne serait pas le cas si nous avions recours à un signal chirp logarithmique. Il faudrait alors avoir recours à un filtre dont la fréquence de coupure est variable avec le temps, ce qui pose des problèmes au moment de l’implantation.

L’utilisation d’un signal chirp linéaire pour l’excitation du système est aussi le principal inconvénient de cette méthode. En effet, la densité spectrale de puissance du signal est quasiment

blanche sur toute la plage de fréquence, ce qui n'est pas le cas de la densité spectrale du bruit ambiant de la pièce, qui est plus forte aux basses fréquences qu'aux hautes fréquences. Ainsi, le rapport signal sur bruit de la méthode est pauvre aux basses fréquences, et bon aux hautes fréquences. Pour pallier cet inconvénient, il est nécessaire d'allonger la durée globale du signal, afin que les basses fréquences reçoivent une quantité d'énergie supplémentaire. Une autre méthode consisterait à utiliser deux signaux d'entrée chirps différents, l'un balayant la plage des basses fréquences, et l'autre celle des hautes fréquences. Un autre problème est l'apparition d'ondulations sur la réponse fréquentielle aux basses fréquences, qui sont dues au fait que la composante à la fréquence double du signal au moment du traitement n'est que partiellement filtrée par le filtre passe-bas. Afin de réduire ce problème une première méthode est d'allonger la durée globale du signal par un certain facteur, et de diviser la fréquence de coupure du filtre passe-bas par le même facteur. La meilleure méthode pour éliminer ces ondulations est d'effectuer une mesure supplémentaire qui consiste à échanger les rôles de la sinusoïde et de la cosinusoïde, d'ajouter les deux réponses correspondantes et d'en faire la moyenne. Dans ce cas, les termes à la fréquence double interfèrent destructivement, et il ne reste plus que les termes utiles. A la limite, le filtre passe-bas n'est plus nécessaire en sortie. Cette nouvelle méthode d'acquisition permet alors de mesurer les réponses impulsionnelles d'une salle, ce qui est notre objectif, car nous ne sommes plus obligés de placer un filtre passe-bas, qui supprime les réflexions caractéristiques du lieu d'écoute. En revanche, si l'on souhaite caractériser les performances d'une enceinte, on peut continuer à utiliser cette méthode et laisser le filtre présent afin de caractériser le comportement de l'enceinte en champ libre. Certains pourraient prétendre qu'il est nécessaire de multiplier par deux la durée d'acquisition pour une telle mesure. A titre de comparaison, il est aussi nécessaire de répéter des séquences MLS au moins deux fois pour extraire la réponse impulsionnelle, car il est nécessaire de faire la mesure dans un contexte périodique, prenant en compte le recouvrement temporel, mais, dans le cas de la TDS, cette deuxième mesure permet d'améliorer le RSB de 3dB, et les deux méthodes possèdent la même résolution.

A.3 Dual-Channel FFT Analysis

Un inventaire des méthodes de mesure de réponses impulsionnelles serait incomplet si on ne mentionnait pas cette méthode, car elle est aussi vieille que les premiers analyseurs de spectre à FFT. Le principe de base est de diviser le spectre du signal acquis par le spectre du signal de référence. La source de bruit des analyseurs à FFT étant non déterministe, il est nécessaire d'enregistrer à la fois le signal d'entrée, et le signal de sortie. C'est un premier inconvénient de cette méthode, mais ce n'est pas le pire. En effet, la source de bruit possède une densité spectrale blanche lorsque celle-ci est observée sur un horizon assez long, mais ce n'est pas le cas, lorsque l'horizon d'observation est bref, il est alors nécessaire d'effectuer une série de moyennes successives sur le signal avant d'obtenir un rapport signal sur bruit suffisant, et qui peut rester toutefois insuffisant dans certaines plages fréquentielles. La nécessité de faire une série de moyennes successives rend ce dispositif peu réactif si le système n'est pas supposé invariant. De plus, la connaissance du temps de propagation entre la source et le capteur doit être connue et prise en compte, afin que ce soit la même partie du signal qui soit analysée au moment du traitement. Le seul avantage que possède cette technique est qu'il permet de réaliser ces mesures sans obstruction du voisinage, car nous ne sommes pas obligés d'utiliser des sources de bruit comme excitation, mais tout morceau de musique peut faire l'affaire. Cependant, le contenu spectral d'un morceau de musique peut être fortement erratique, et il est souvent préférable d'utiliser des sources de bruit lorsque c'est possible.

A.4 Excitation en régime harmonique

La méthode qui requière certainement le plus de temps pour acquérir la réponse fréquentielle, en amplitude et en phase, est certainement celle qui consiste à exciter le système en régime harmonique, et de réaliser étape par étape, à chaque fréquence, la même mesure. Après chaque mesure, on élève la fréquence du signal d'un certain facteur, ce qui permet d'avoir une échelle logarithmique. Ainsi, contrairement à toutes les méthodes basées sur les FFT, nous disposons de plus de points à basses fréquences qu'à hautes fréquences. Ce n'est pas un désavantage de la méthode car les méthodes basées sur la FFT fournissent généralement trop de renseignements pour les hautes fréquences. A l'heure où les capacités de stockage des données sont grandes, l'économie sur les données n'est pas à retenir. Le principal avantage que confère cette méthode est qu'elle permet d'obtenir un rapport signal sur bruit énorme, étant donné que toute l'énergie est concentrée sur une seule fréquence. Ainsi, en dépit du temps nécessaire pour sa mise en oeuvre, cette méthode peut apparaître comme une solution possible lorsque l'on souhaite connaître avec beaucoup de précision et de certitude la réponse impulsionnelle d'un équipement acoustique.

A.5 Impulsions

La méthode la plus naturelle pour obtenir des réponses impulsionnelles est de placer des impulsions en entrée du système à étudier. Cette impulsion peut être créée de manière analogique, mais on préfère l'envoyer directement sous sa forme numérique à un convertisseur numérique-analogique. La réponse ainsi enregistrée par le capteur est la réponse impulsionnelle souhaitée, et aucun traitement supplémentaire n'est nécessaire. Toutefois, comme les transducteurs sont souvent limités par l'énergie maximale de l'impulsion —et non par l'amplitude, comme on pourrait le penser a priori— que l'on peut envoyer, il est nécessaire de répéter périodiquement ces impulsions afin d'obtenir un rapport signal sur bruit suffisant. De plus il faut faire attention à ne pas rentrer dans la zone de fonctionnement non linéaire du système, car cette méthode de mesure ne permet pas d'extraire la composante linéaire du signal. La période de cette répétition doit être supérieure à la longueur de la réponse impulsionnelle afin de se prémunir de tout phénomène de recouvrement temporel. On augmente ainsi le RSB de 3 dB lorsqu'on multiplie le nombre de moyennes par deux.

Afin d'augmenter la précision d'une telle mesure, on calcule la FFT de la réponse impulsionnelle obtenue, que l'on divise par une FFT de référence, obtenue en branchant directement l'entrée sur la sortie du système. Ainsi, on se prémunit des éventuels défauts de l'impulsion initiale. C'est ce qui est fait lorsqu'on utilise l'analyse à deux voies par FFT. Nous combinons ici les avantages des deux méthodes.

A.6 Maximum-Length Sequences (MLS)

Cette méthode a été développée dans le courant des années 70, et les restrictions n'étaient pas les mêmes qu'aujourd'hui, et certaines d'entre elles peuvent sembler superflues voire abusives. Les maximum-length sequences sont des séquences binaires qui peuvent être synthétisées très facilement par du matériel électronique extrêmement élémentaire : une porte ou exclusif et un registre à décalage d'ordre N . Ainsi, du temps où la mémoire était coûteuse, il était pratique de pouvoir générer les $2^N - 1$ —le cas "0" est éliminé— mots de N bits différents à partir d'un matériel aussi rudimentaire. Aujourd'hui cette restriction n'est plus d'actualité, et il est plus coûteux d'écrire et d'exécuter un programme en assembleur permettant de générer une telle séquence, que de la lire directement à partir d'une zone de stockage. Les séquences MLS possèdent la bonne propriété d'avoir une fonction d'autocorrélation proche du Dirac, ce qui les

rend très attractives, car leur densité spectrale est blanche. Ces séquences binaires sont passées à travers un convertisseur N/A, et il est possible de fournir beaucoup plus d'énergie lors d'une mesure que si l'on utilisait une unique impulsion, ce qui permet d'augmenter le rapport signal sur bruit.

Les signaux d'excitation à densité spectrale blanche permettent le recours à des algorithmes de *cross correlation* pour extraire la réponse impulsionnelle, ce qui peut être réalisé rapidement par l'algorithme de transformée rapide de Hadamard, sans quitter le domaine temporel. Dans le temps, cela signifiait des temps de calculs beaucoup plus courts. Aujourd'hui les temps de calcul sont analogues en utilisant la Fast Hadamard Transform (FHT) ou la FFT.

Afin d'obtenir la réponse impulsionnelle d'une enceinte, il est aussi nécessaire de prendre en compte le retard de propagation, ce qui se fait en réalisant une permutation circulaire des échantillons issus de la FHT. Les réflexions secondaires sont éliminées en multipliant le résultat par une certaine fenêtre temporelle, afin de limiter l'horizon d'observation. De même que la méthode précédente, il est possible d'améliorer grandement la précision de la mesure en prenant en compte une mesure de référence obtenue en branchant directement l'entrée du système sur la sortie, ce qui prend en compte les imperfections de l'entrée. Cette opération nécessite alors des FFT supplémentaires, mais ce n'est plus une restriction aujourd'hui.

Bien que les systèmes de mesure basés sur les MLS soient assez populaires dans le domaine de l'acoustique, ils ont néanmoins certains inconvénients si les hypothèses de linéarité et d'invariance du système ne sont plus vérifiées, mais tout ceci sera discuté plus largement dans la prochaine section, dans la partie concernant les critères nous ayant orientés vers une méthode basée sur les signaux chirp plutôt que séquences MLS ou pseudo-aléatoires, de manière générale. De plus, les MLS possèdent une densité spectrale blanche, ce qui est un autre inconvénient. On peut pallier cet inconvénient en plaçant un filtre d'égalisation en sortie de la séquence MLS. La séquence perd alors son caractère binaire, et il n'est plus possible d'utiliser un convertisseur 1-bit uniquement pour transformer la séquence numérique en un signal analogique, mais un véritable convertisseur N/A. Si on réalise cette opération d'égalisation, la réponse impulsionnelle obtenue sera égale à la réponse impulsionnelle réelle convoluée avec celle du filtre d'égalisation, ce qui nécessite une opération de déconvolution supplémentaire. Plutôt que de l'effectuer, on divise plutôt toute mesure par le spectre de référence comme mentionné précédemment.

Pour conclure, aujourd'hui, il est préférable d'utiliser la prochaine méthode si l'on veut se servir de séquences pseudo-aléatoires pour le calcul de la réponse impulsionnelle d'un système acoustique, car cela ôte toutes les restrictions permettant de développer un système rapide et peu coûteux en capacités de stockage, ce qui est devenu obsolète aujourd'hui. De plus, si l'on souhaite égaliser la séquence de départ, les calculs de FFT et de IFFT sont nécessaires, et auraient été utopiques pour les systèmes de l'époque.

A.7 Excitation par des signaux périodiques de longueur 2^N

Plutôt que d'utiliser la transformée rapide de Hadamard, il semble plus sage de comparer le signal de sortie à un signal de référence, obtenu en branchant directement l'entrée du système sur la sortie. On prend ainsi en compte toutes les caractéristiques du signal d'entrée, qu'il soit un bruit, une séquence pseudo-aléatoire, ou un extrait musical. L'essentiel est que ce soit toujours le même signal d'entrée qui soit utilisé, et que le système ait été calibré pour ce signal spécifique. Cette méthode vient généraliser celle de l'analyseur FFT à deux voies, mais permet néanmoins de s'affranchir de quelques défauts :

- Le signal d'excitation est connu à l'avance, ce qui permet de n'utiliser qu'une seule voie pour le traitement. De plus, l'utilisation de ce signal source est donc déterministe, en comparaison de l'utilisation de sources de bruits décorréliés mais non déterministes dans le

cas de l'analyseur FFT à deux voies, qui rendaient impossible l'utilisation d'un spectre de référence, et qui nécessitaient de moyenner le spectre.

- Il existait toujours des disparités entre les deux voies d'analyse dans un analyseur à deux voies initiales, malgré la bonne volonté du constructeur, ce qui venait dégrader le traitement. Maintenant, la sortie s'adapte parfaitement au signal d'entrée, étant donné que c'est la même voie qui a servi de référence.

Ainsi, il suffit de multiplier la FFT du signal de sortie par l'inverse de la FFT du signal de référence, ce qui est une multiplication complexe (ce qui permet de conserver les informations sur l'amplitude et la phase). Ensuite, il ne reste plus qu'à calculer une IFFT pour obtenir la réponse impulsionnelle du système initial.

On peut construire le signal d'excitation de toute part, afin qu'il s'adapte aux particularités du lieu d'écoute et notamment, au niveau du bruit ambiant de la pièce, de sorte à obtenir un rapport signal sur bruit constant en fonction de la fréquence. On calcule un tel signal dans le domaine temporel en prenant des phases aléatoires pour le domaine fréquentiel, et en calculant le signal d'excitation correspondant par IFFT. Si ce signal est répété de manière périodique, nous sommes sur qu'il aura le spectre désiré, et mentionné précédemment (par exemple blanc, ou rose).

Les défauts d'une telle méthode sont analogues aux défauts d'une méthode basée sur les MLS, notamment en ce qui concerne leur vulnérabilité si les hypothèses de linéarité et d'invariance du système ne sont plus vérifiées. Comme le spectre prédéfini du signal de bruit n'est valide que pour une répétition périodique, il est impossible de débiter la mesure à la mise en route du stimulus, et nous sommes obligés d'attendre que le système soit en régime permanent, ce qui est quasiment le cas lorsque une durée de l'ordre de la longueur de la réponse impulsionnelle du système à étudier s'est écoulée. Ainsi, comme la longueur de la séquence 2^N a été choisie de sorte à ce que la longueur de la réponse impulsionnelle discrétisée soit inférieure à cette limite, il est nécessaire de répéter au minimum deux fois la séquence, et de débiter l'enregistrement lors de la répétition uniquement.

A.8 Excitation du système par des signaux Chirp

Plutôt que de rendre la phase aléatoire et d'obtenir une séquence aléatoire ayant la densité spectrale escomptée pour optimiser le rapport signal sur bruit, la phase de la séquence peut aussi être ajustée de sorte à obtenir des avantages supplémentaires lors de la mesure de la réponse impulsionnelle d'un système acoustique. Le premier à essayer de tirer parti de cette constatation fut A. Farina [43], qui élaborait une nouvelle méthode de mesure caractérisant à la fois la réponse impulsionnelle du système, mais aussi la distorsion du système. L'usage de signaux Chirp procure de larges avantages par rapport à l'utilisation de séquences pseudo-aléatoires, le premier étant que le spectre d'un signal chirp est quasiment identique à celui de sa version périodique, ce qui permet de réduire la durée d'acquisition de moitié en maintenant la même résolution spectrale et le même rapport signal sur bruit.

Le principale avantage de la méthode est le fait que les composantes de la distorsion harmonique peuvent être isolées entièrement, permettant ainsi d'acquérir uniquement la réponse impulsionnelle, c'est-à-dire, correspondant à la partie linéaire du système, par définition, ce qu'était incapable de faire les méthodes basées sur les MLS ou séquences pseudo-aléatoires.

L'explication du rejet de la distorsion peut se comprendre facilement en terme de retard de groupe. Considérons un signal Chirp dont la fréquence est à 100 Hz après 100 ms de signal, et qui atteint 200 Hz au bout de 200 ms. Afin d'obtenir la réponse impulsionnelle, il est nécessaire de comprimer toutes ces composantes fréquentielles au même endroit, ce qui se fait par exemple en utilisant un filtre ayant un retard de groupe de -100 ms à 100 Hz et de -200 ms à 200 Hz.

Lorsque la fréquence instantanée du signal est à 100 Hz, le système non linéaire étudié possède des composantes dues à la distorsion harmonique à 200 Hz. Lorsqu'on filtre le signal initial par le filtre inverse, cette composante non linéaire d'ordre 2 se retrouve finalement en -100 ms. Il en est de même pour les harmoniques d'ordre plus élevé. Finalement on ne retient que la composante linéaire du système par cet artifice.

B Présentation de la méthode retenue

B.1 Critères de choix

Cette partie a pour but d'expliquer les différences de comportements entre les systèmes de mesure basés sur des séquences MLS ou pseudo-aléatoires, et ceux basés sur des signaux Chirp. Nous étudierons les différences de comportement de ces deux systèmes en ce qui concerne les critères de durée de mesure, de facteur de crête, et de ses non-linéarités.

B.1.1 Durée de la mesure

Pour toutes les mesures, il est nécessaire que la durée d'acquisition soit au moins aussi longue que la durée de la réponse impulsionnelle du système étudié. Cela semble évident dans le cas où l'on envoie une seule impulsion en entrée du système étudié, et que l'on enregistre tous les échantillons jusqu'à ce que le niveau de ceux-ci soit fondus dans le bruit ambiant. Dans le cas d'une excitation par un signal chirp non périodique, le signal contient quasiment toutes les fréquences d'une impulsion, mais elles ne sont pas réparties de la même manière dans le temps. Généralement, les signaux chirp balaient une certaine plage de fréquence des fréquences graves vers les fréquences aigues, venant exciter chacun des modes du système d'étude. Ceux-ci possèdent des durées de résonance beaucoup plus longues aux basses fréquences qu'aux hautes fréquences, si bien qu'il est nécessaire d'enregistrer en pratique sur une durée très peu supérieure à celle du système étudié (aux hautes fréquences la durée de réverbération de la salle est très brève).

Dans le cas d'excitations périodiques, par des séquences MLS, ou pseudo-aléatoires, il n'est pas forcément nécessaire d'utiliser une période supérieure à la longueur de la réponse impulsionnelle, mais dans ce cas, il y a un effet de recouvrement temporel. Au concepteur du système de mesures de voir lorsque l'erreur produite par la partie repliée devient tolérable. Pour ce genre de mesure, il est nécessaire de répéter au moins deux fois le signal d'excitation, ce qui signifie des durées d'acquisition plus longues pour la même résolution spectrale, et une dépense d'énergie inutile étant donné que 50 % de la mesure est utilisé seulement.

B.1.2 Facteur de crête

Le facteur de crête est défini comme le quotient de la valeur maximale de l'amplitude du signal sur sa valeur efficace. Afin d'obtenir les meilleures performances pour la mesure, il est nécessaire de normaliser la valeur de pic à son maximal autorisé, dépendant soit du système de mesure, soit du système étudié, selon que l'on cherche à analyser le comportement d'une salle ou d'une enceinte. Ainsi, le facteur de crête est un indicateur sur la quantité d'énergie perdue en employant tel type de signal d'excitation.

Ainsi, au premier abord, les séquences MLS bipolaires apparaissent comme le signal d'excitation idéal, étant donné que la valeur de pic est égale à la valeur efficace du système. Malheureusement, la séquence binaire passe à travers le convertisseur N/A, et un filtre anti-aliasing est nécessaire en sortie, afin de redonner une forme convenable au signal analogique. Il est nécessaire de diminuer la valeur efficace de la séquence MLS de 5 à 8 dB environ pour éviter de saturer

l'étage de sortie du convertisseur N/A. De plus, l'utilisation de MLS disposant d'une forte valeur efficace augmente le caractère non-linéaire des transducteurs et vient perturber la mesure, mais cela sera discuté par la suite. Dans le cas d'une excitation par des signaux chirp, le facteur de crête est de 3 dB, et il n'est pas nécessaire de réduire la valeur efficace maximale en entrée à cause du filtre anti-aliasing de sortie.

B.1.3 Comportement vis-à-vis de la distorsion

Pour les séquences pseudo-aléatoires, dans les endroits calmes, ce n'est pas le bruit ambiant qui vient limiter le RSB de la réponse impulsionnelle, mais la distorsion générée par l'enceinte utilisée au moment de l'acquisition. En effet, dans toute mesure utilisant du bruit comme signal d'excitation, les produits dus à la distorsion du stimulus possédant une distribution de phases aléatoire, possède lui aussi plus ou moins une distribution de phases aléatoires. Le processus de déconvolution implique lui aussi une distribution de phases aléatoires, et le spectre d'erreur finalement obtenu possède aussi une distribution de phase aléatoire, qu'aucune moyenne ne sera capable d'ôter car ce terme d'erreur est corrélé avec le signal d'excitation initial. Afin d'augmenter le rapport signal sur bruit, on peut penser augmenter la durée de la mesure d'un certain facteur, mais le résultat obtenu n'est pas forcément meilleur, car le système étudié n'est jamais généralement invariant parfaitement, et c'est la variance temporelle du système qui prédomine pour la dégradation de la mesure.

Afin de réduire l'influence des produits de distorsion, il est plus judicieux de les exclure totalement de la réponse impulsionnelle extraite finalement, en utilisant des méthodes utilisant des signaux chirp comme stimulus. C'est pourquoi les séquences pseudo-aléatoires ne permettent pas d'obtenir de bons rapports signal sur bruit, et qu'il est préférable d'utiliser des signaux chirp.

B.2 Description du protocole expérimental

Nous avons utilisé un signal chirp logarithmique balayant la plage fréquentielle de 10 Hz à 22050 Hz (fréquence d'échantillonnage à 44100 Hz) en 5 secondes, dont la formule de synthèse est la suivante :

$$e(t) = \sin \left[\frac{\omega_1 T}{\ln \left(\frac{\omega_2}{\omega_1} \right)} \cdot \left(e^{\frac{t}{T} \ln \left(\frac{\omega_2}{\omega_1} \right)} - 1 \right) \right] \quad (\text{A.2})$$

Nous pouvons visualiser sur la figure A.1 le module de la TF d'un tel signal, ainsi que le retard de groupe en fonction de la fréquence. Il faut maintenant construire le filtre inverse correspondant. Celui-ci devra avoir un retard de groupe nul pour la fréquence 22050 Hz, et un retard de groupe de 5 secondes pour la fréquence de 10 Hz. Il s'agit aussi d'un signal chirp logarithmique, dont on aurait inverser les échantillons par rapport au premier. Mais il faut aussi tenir compte de la pente dans le spectre en amplitude, et il est nécessaire d'amplifier les hautes fréquences dans le filtre inverse, afin que la convolution des deux soit la plus proche possible de l'impulsion souhaitée.

Nous pouvons visualiser sur la figure A.2 les caractéristiques du filtre inverse utilisés lors du processus de déconvolution, et nous pouvons visualiser le résultat de cette déconvolution sur la figure A.3. Comme nous pouvons le voir le résultat n'est pas parfait, étant donné que l'impulsion restitué par ce filtre inverse est une impulsion filtrée entre 10 et 22050 Hz. Néanmoins, c'est ce spectre qui servira de référence par la suite, car c'est celui obtenu en branchant directement l'entrée du système sur la sortie. On normalisera donc tous les traitements obtenus par ce spectre de référence, afin d'obtenir une meilleure précision sur nos mesures.

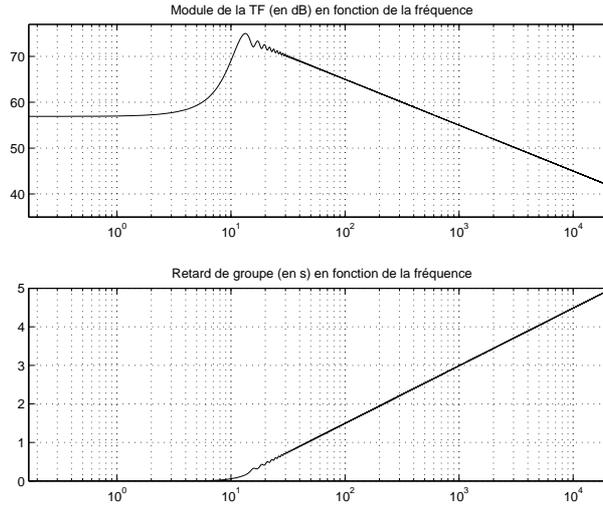


FIG. A.1 – Caractéristiques du signal chirp logarithmique

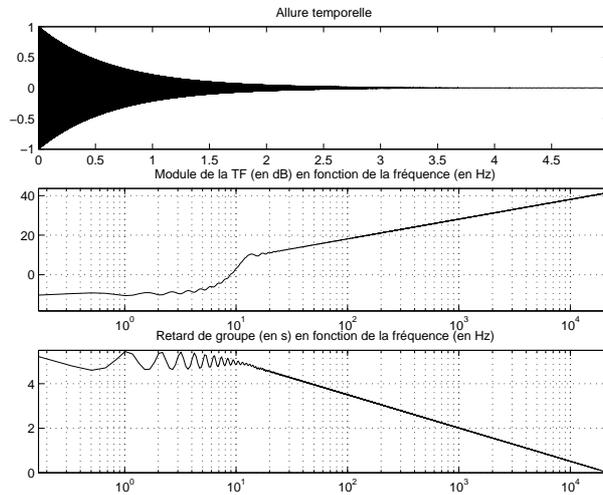


FIG. A.2 – Caractéristiques du filtre inverse

La particularité des signaux chirp logarithmiques par rapport aux signaux chirp linéaires est qu'ils permettent aussi d'identifier chacune des composantes d'un noyau de Volterra d'ordre N , si le système est supposé non linéaire, mais qu'il satisfait à une relation du type :

$$s(t) = n(t) + e(t) \star h_1(t) + e^2(t) \star h_2(t) + \dots + x^N(t) \star h_N(t) \quad (\text{A.3})$$

Si le système est supposé linéaire, on a dans ce cas $h_2(t) = \dots = h_N(t) = 0$. Si le système est supposé faiblement linéaire, ce qui est typiquement le cas pour les enceintes, la décomposition ci-dessus est valide [43]. Nous avons vu de manière générale que les produits dus à la distorsion harmonique étaient rejetés en amont de l'impulsion dans le cas du chirp linéaire. Dans le cas d'un chirp logarithmique, le délai nécessaire pour que la fréquence instantanée soit égale à N fois la fréquence actuelle est constant, et ne varie pas en fonction de la position de l'échantillon dans le signal —il a été synthétisé pour cela— ce qui permet aussi bien d'identifier la réponse impulsionnelle correspondant à la partie linéaire du système $h_1(t)$, que les différentes réponses caractérisant le noyau de Volterra. La k -ième réponse impulsionnelle $h_k(t)$ est ainsi rejetée en

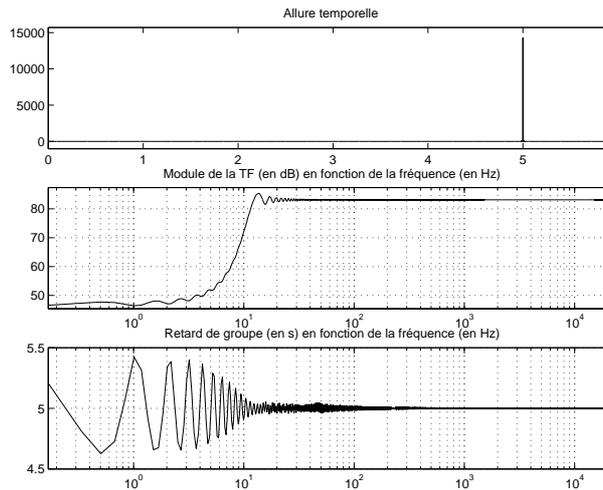


FIG. A.3 – Résultats de l'inversion

amont de l'impulsion principale d'un délai $\Delta t = T \frac{\ln k}{\ln \omega_2 / \omega_1}$. Vous pouvez voir, sur la figure A.4 un exemple de résultat obtenu par ce traitement. Il fait bien apparaître la réponse impulsionnelle principale autour de l'origine, à un délai prêt, dû au temps de latence de l'électronique qui a réalisé l'enregistrement, et au temps de propagation de l'onde sonore. Les autres composantes du noyau de Volterra d'un système non linéaire, supposé sans mémoire, apparaissent successivement en amont de cette impulsion comme indiqué sur la figure [43]. Nous pouvons observer sur la

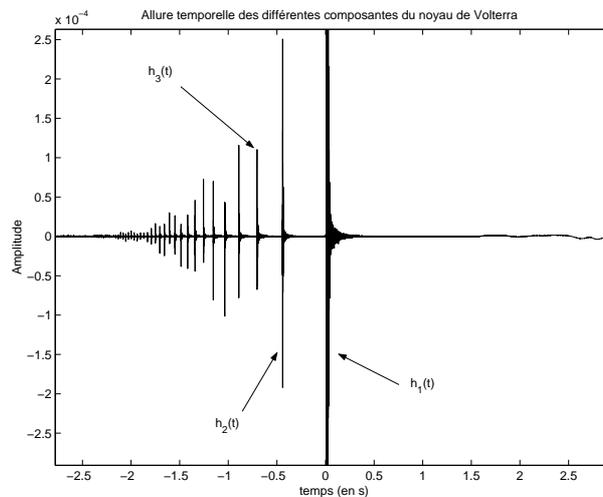


FIG. A.4 – Exemple de résultat après traitement

figure A.5 les différents spectres des trois premiers noyaux de cette décomposition. Nous pouvons en conclure que les non-linéarités sont bien présentes dans le système (25 dB de différence avec la composante linéaire en moyenne), et les négliger n'est donc qu'une première approximation.

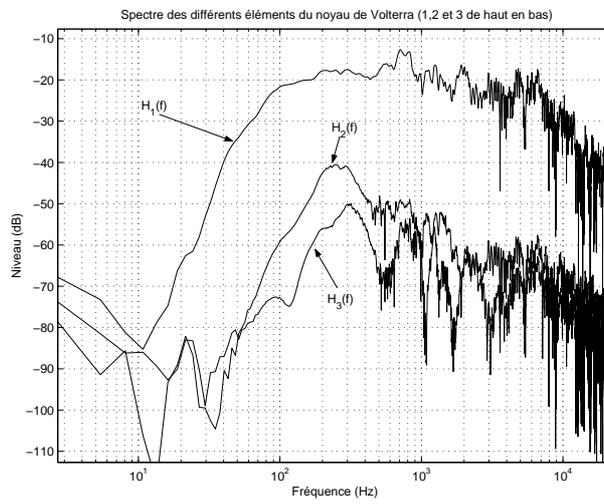


FIG. A.5 – Spectre des différentes composantes du noyau de Volterra

Bibliographie

- [1] A.J. Berkhout, “*Seismic Migration, imaging of acoustic energy by wave field extrapolation*”, Ed. Elsevier.
- [2] P.M. Morse, H. Feschbach, “*Methods of theoretical physics*”, part I, 1953.
- [3] J. Jouhaneau, “*Propagation des ondes acoustiques*”, Techniques de l’Ingénieur, Article TE 5130.
- [4] S. Ise, “*A principle of sound field control based on the Kirchoff-Helmholtz integral equation and the theory of inverse systems*”, Acta Acustica, Vol. 85, pp. 78-87, 1999.
- [5] S. Takane, Y. Suzuki and T. Sone, “*A new method for global sound field reproduction based on Kirchoff’s integral equation*”, Acta Acustica, vol. 85, pp. 250-257, 1999.
- [6] A.J. Berkhout “*A holographic Approach to acoustic control*”, Journal of Audio Engineering Society, Vol. 36, n°12, December 1988.
- [7] A.J. Berkhout, D. de Vries and P. Vogel, “*Acoustic control by Wave Field Synthesis*”, Journal of Audio Engineering Society, Vol. 93(5), May 1993.
- [8] M.M. Boone, E.N.G. Verheijen and P. Van Tol, “*Spatial sound field reproduction by Wave Field Synthesis*”, Journal of Audio Engineering Society, Vol. 43, n°12, December 1995.
- [9] D. de Vries, “*Sound reinforcement by wave field synthesis : adaptation of the synthesis operator to the loudspeaker directivity characteristics*”, Journal of Audio Engineering Society, Vol. 44, n°12, December 1996.
- [10] M. Van der Wal, E.W. Start and D. de Vries, “*Design of logarithmically spaced constant-directivity transducer arrays*”, Journal of Audio Engineering Society, Vol. 44, n° 6, June 1996.
- [11] A.J. Berkhout, D. de Vries and J.J. Sonke, “*Array technology for acoustic wave field analysis in enclosures*”, Journal of Acoustical Society of America, Vol. 102(5), November 1997.
- [12] E.W. Start, D. de Vries and A.J. Berkhout, “*Wave Field Synthesis operators for bent line arrays in a 3D space*”, Acta Acustica, Vol. 85, pp. 883-892, 1999.
- [13] D. de Vries and M.M. Boone, “*Wave Field Synthesis and Analysis using array technology*”, Proc. 1999 IEEE Workshop on Applications on Signal Processing to Audio and Acoustics, October 1999.
- [14] M.M. Boone, “*Acoustic rendering with Wave Field Synthesis*”, ACM Siggraph, Acoustics Renderings for virtual Environments, May 2001.
- [15] D. de Vries and J. Baan, “*Auralization of sound fields by Wave Field Synthesis*”.
- [16] E. Hulsebos, D. de Vries and E. Bourdillat, “*Improved microphone array configurations for auralization of sound fields by Wave Field Synthesis*”, Presented at the 110th Convention of Audio Engineering Society, Preprint 5337, 2001.
- [17] R. Nicol and M. Emerit, “*Reproducing 3D-sound for videoconferencing : a comparison between holophony and ambisonic*”, DAFX98 Proceedings, pp. 17-20, 1998.

- [18] R. Nicol, M. Emmerit, *3D-sound reproduction over an extensive listening area : a hybrid method derived from holophony and ambisonic*”, Audio Engineering Society 16th International conference, pp. 436-453, 1999.
- [19] R. Nicol, M. Emmerit and A. Gilloire, “*Mur de téléprésence pour la visioconférence, une approche holophonique*”, CORESA 98.
- [20] J.D. Maynard, E.G. Williams and Y. Lee, “*Nearfield acoustic holography : I. Theory of generalized holography and the development of NAH*”, Journal of Acoustical Society of America, vol. 78(4), 1985.
- [21] W.A. Veronesi and J.D. Maynard, “*Nearfield acoustic holography (NAH) II. Holographic reconstruction algorithms and computer implementation*”, Journal of Acoustical Society of America, vol. 81(5), 1987.
- [22] W.A. Veronesi and J.D. Maynard, “*Digital holographic reconstruction of sources with arbitrarily shaped surfaces*”, Journal of Acoustical Society of America, vol. 85(2), 1989.
- [23] A.J. Berkhout, and D. de Vries, “*Simulation of responses in enclosed spaces*”, Convention of Audio Engineering Society, Preprint : 4714, 1998.
- [24] U. Horbach et al., “*Numerical simulation of wave fields created by loudspeaker arrays*”, 107th Convention of Audio Engineering Society, September 1999.
- [25] S. Spors and R. Rabenstein, “*Virtual acoustics by numerical simulation*”, 46th International Scientific Colloquium, Ilmenau Technical University, September 2001.
- [26] M. Miyoshi, and Y. Kaneda, *Inverse control of room acoustics using multiple loudspeakers and/or microphones*, Proc. ICASSP86, Tokyo, Japan, April 1986, pp. 18A.4 1-4.
- [27] M. Miyoshi, and Y. Kaneda, *Inverse filtering of room acoustics*, IEEE trans. on Acoustics, Speech, and Signal Processing, vol. 36, n°2, February 1988.
- [28] P.A. Nelson, H. Hamada, and S.J. Elliott, *Adaptive inverse filters for stereophonic sound reproduction*, IEEE trans. on Signal Processing, vol. 40, n°7, July 1992.
- [29] J.N Mourjopoulos, *Digital equalization of room acoustics*, Journal of Audio Engineering Society, vol. 42, n°11, November 1994.
- [30] P.A. Nelson, F. Orduna-Bustamante, and H. Hamada, *Inverse filter design and equalization zones in multichannel sound reproduction*, IEEE trans. on Speech and Audio Processing, vol. 3, n°3, May 1995.
- [31] S.L. Gay and S. Tavathia, *The fast affine projection algorithm*, IEEE 1995.
- [32] P.A. Nelson, and F. Orduna-Bustamante, *Multichannel signal processing techniques in the reproduction of sound*, Journal of Audio Engineering Society, vol. 44, n°11, November 1996.
- [33] O. Kirkeby, P.A. Nelson, H. Hamada, and F. Orduna-Bustamante, *Fast deconvolution of multichannel systems using regularization*, IEEE trans. on Speech and Audio Processing, vol. 6, n°2, March 1998.
- [34] S. Weiss, G.W. Rice, and R.W. Stewart, *Multichannel equalization in subbands*, Proc. 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New York, October 1999.
- [35] M. Bouchard, and S. Quednau, *Multichannel recursive least-squares algorithms and fast-transversal-filter algorithms for active noise control and sound reproduction systems*, IEEE trans. on Speech and Audio Processing, vol. 8, n°5, September 2000.
- [36] M. Bouchard, and Y. Feng, *Inverse structure for active noise control and combined active noise control/sound reproduction systems*, IEEE trans. on Speech and Audio Processing, vol. 9, n°2, February 2001.

- [37] E. Corteel, U. Horbach, and R. Pellegrini, *Multichannel inverse filtering of multiexciter distributed mode loudspeakers for wave field synthesis*, Presented at the 112th Convention of Audio Engineering Society, Preprint 5611, May 2002.
- [38] U. Horbach, E. Corteel, and D. de Vries, *Spatial audio reproduction using distributed mode loudspeaker arrays*, presented at the 21st conference of Audio Engineering Society, St Petersburg, Russia, June 2002.
- [39] M. Bouchard, *Multichannel affine and fast affine projection algorithms for active noise control and acoustic equalization systems*, IEEE trans. on Speech and Audio Processing, vol. 11, n°1, January 2003.
- [40] D.D. Rife, and J. Vanderkooy, *Transfer-function measurement with maximum-length sequences*, Journal of the Audio Engineering Society, vol. 37, n°6, June 1989.
- [41] C. Dunn, and M.O Hawksford, *Distortion immunity of MLS-derived Impulse Response Measurements*, Journal of Audio Engineering Society, vol. 41, n°5, May 1993.
- [42] J.M Vanderkooy, *Aspects of MLS Measuring Systems*, Journal of Audio Engineering Society, vol. 42, n°4, April 1994.
- [43] A. Farina, *Simultaneous measurement of impulse response and distortion with a swept-sine technique*, Presented at the 108th Convention of Audio Engineering Society, Paris, France, February 2002.
- [44] S. Müller, and P. Massarani, *Transfer-function measurement with sweeps*, Journal of Audio Engineering Society, vol. 49, n°6, June 2001.
- [45] GB. Stan, JJ. Embrechts, and D. Archambeau, *Comparison of different impulse response measurement techniques*, Journal of Audio Engineering Society, vol. 50, n°4, April 2002.