

Université Paris 6, Université Aix-Marseille, ENST, INPG-Grenoble

Rapport de stage DEA ATIAM

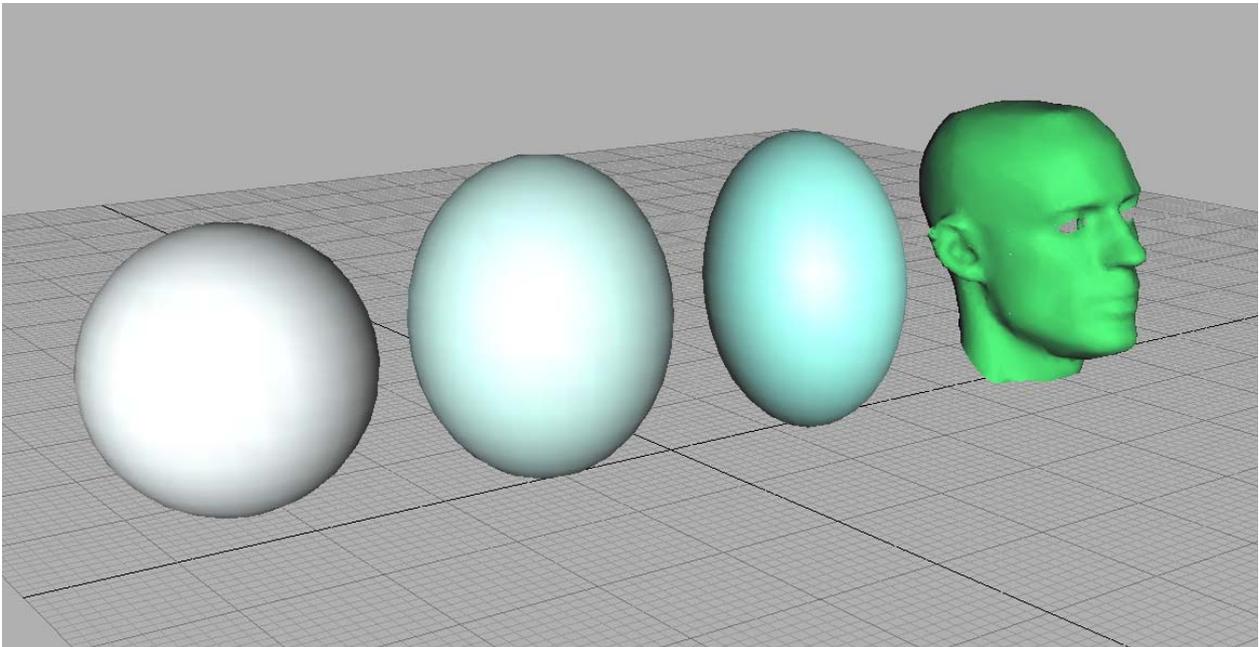
Modélisation de fonctions de transfert acoustiques de têtes humaines (HRTF) et application à l'individualisation de la synthèse binaurale

Modèles sphériques et ellipsoïdaux

Par Guillaume LE NOST

*Stage effectué dans l'équipe DIH / IPS / SON 3D
de France Telecom r&d
2, avenue Pierre Marzin
22307 LANNION CEDEX,*

d'avril à octobre 2002.



Géométrie des modèles acoustiques de tête : sphère, sphéroïde prolate, ellipsoïde

Sujet de Stage

Les techniques dites binaurales permettent de positionner des sources sonores dans l'espace en utilisant un casque d'écoute. Cependant, ces techniques nécessitent la connaissance des fonctions de transfert acoustiques de la tête de l'utilisateur (Head Related Transfer Functions en anglais ou HRTFs). Actuellement, ces fonctions sont mesurées de manière individuelle, et la procédure est longue et coûteuse. C'est pourquoi l'un de nos objectifs est de développer des modèles de HRTFs permettant d'obtenir ces fonctions par le calcul, sans recourir à la mesure. Les paramètres du modèle pouvant être individualisés pour chaque utilisateur.

Une base de données de HRTFs à haute résolution spatiale a été constituée pour 8 sujets permanents à FT R&D, et des scanners 3D de la tête et du torse de ces sujets ont été effectués. Une plate forme de tests de localisation sonore a aussi été mise en place, et un premier test de localisation utilisant la synthèse binaurale a été mené.

Une des pistes possibles pour l'individualisation des HRTFs est de modéliser le comportement acoustique de la tête humaine sous une forme simplifiée (sphère, sphéroïde prolate, ellipsoïde). Les paramètres du modèle (diamètre de la sphère, dimensions caractéristiques de l'ellipsoïde), seraient extraites de mesures anthropométriques (réalisées par exemple sur les scanners 3D des huit sujets dont nous disposons).

L'évaluation de la performance de ces modèles serait ensuite faite de manière objective par comparaison avec les mesures disponibles, puis de manière subjective par des tests de localisation qui permettront de comparer les performances obtenues avec les modèles à celles obtenues avec les HRTFs mesurées (considérées comme références).

Travail à réaliser :

Un modèle de diffraction d'une onde plane monochromatique sur une sphère rigide (solution analytique) est déjà disponible sous forme d'une fonction Matlab. Un premier travail serait de calculer les HRTFs individualisées pour les huit sujets en utilisant le modèle sphérique (avec ou sans prise en compte du décalage des oreilles), en estimant la dimension caractéristique du modèle sur les scans 3D des 8 sujets. Afin d'évaluer la validité du modèle sphérique, on procèdera ensuite à une évaluation objective (comparaison des HRTFs calculées et mesurées) et subjective (test de localisation sonore utilisant les HRTFs calculées).

Ensuite, selon l'avancement des travaux, il serait intéressant de rechercher dans la littérature et de programmer les solutions analytiques pour une sphéroïde et pour une ellipsoïde.

Si ces solutions ne peuvent pas être mises en œuvre, l'utilisation d'un programme de simulation acoustique utilisant la méthode des éléments finis de frontière (Boundary Elements Method), peut être envisagée.

Table des matières

SUJET DE STAGE.....	3
1. SPATIALISATION DES SONS PAR LES TECHNIQUES BINAURALES	6
1.1. INTRODUCTION A LA PERCEPTION SPATIALE DES SONS.....	6
1.2. PRINCIPE ET ARCHITECTURE	8
1.3. LIMITATIONS DES SYSTEMES ACTUELS.....	9
2. SIMULATION DE HRTF PAR MODELES PHYSIQUES	11
2.1. PROBLEMATIQUE	11
2.2. ETAT DE L'ART	11
2.2.1. <i>Solutions analytiques</i>	11
2.2.2. <i>Solutions numériques</i>	13
2.2.3. <i>Une structure modulaire pour la simulation de HRTF</i>	13
3. CALCULS DE HRTFS DE MODELES GEOMETRIQUES SIMPLES PAR LA METHODE DES ELEMENTS FINIS.....	15
3.1. PRINCIPES DU CALCUL PAR ELEMENTS FINIS	15
3.2. UTILISATION DU LOGICIEL VNOISE 1.11.....	17
3.3. LE PRINCIPE DE RECIPROCITE.....	18
3.3.1. <i>Enoncé</i>	18
3.3.2. <i>Définition des modèles directs et réciproques sous VNoise</i>	19
3.3.3. <i>Intérêt du raffinement local de maillage</i>	20
3.4. CONTRAINTES MATERIELLES	23
3.5. VALIDATION DU LOGICIEL VNOISE 1.11	24
4. PROTOCOLE DE CALCUL DES HRTF POUR UN MODELE SIMPLE	27
4.1. DESCRIPTION DU PROTOCOLE	27
4.2. LE MAILLAGE.....	28
4.3. LE CALCUL	28
4.3.1. <i>Positions des points de mesure</i>	28
4.3.2. <i>Choix des fréquences calculées</i>	29
4.3.3. <i>Fonctionnement du solveur</i>	33
4.4. LE TRAITEMENT DES RESULTATS	33
4.5. VALIDATION DU PROTOCOLE	33
5. CHOIX DES MODELES.....	ERREUR ! SIGNET NON DEFINI.
5.1. PROPRIETES	ERREUR ! SIGNET NON DEFINI.
5.2. CHOIX DES MESURES ANTHROPOMETRIQUES	ERREUR ! SIGNET NON DEFINI.
5.3. DEFINITION DES PARAMETRES DES MODELES.....	ERREUR ! SIGNET NON DEFINI.
6. COMPARAISON DES ITDS OBTENUS	ERREUR ! SIGNET NON DEFINI.
6.1. CHOIX DE LA METHODE DE CALCUL D'ITD	ERREUR ! SIGNET NON DEFINI.
6.2. ITD DE LA SPHERE DE RAYON "OPTIMAL"	ERREUR ! SIGNET NON DEFINI.
6.3. ITD DE L'ELLIPSOÏDE TYPE 1 (DROITE, OREILLES CENTREES)	ERREUR ! SIGNET NON DEFINI.
6.4. ELLIPSOÏDE TYPE 2 (DROITE, OREILLES DECALEES)	ERREUR ! SIGNET NON DEFINI.
6.5. ELLIPSOÏDE TYPE 3 (INCLINEE, OREILLES CENTREES)	ERREUR ! SIGNET NON DEFINI.

6.6. DISCUSSION **ERREUR ! SIGNET NON DEFINI.**
7. CONCLUSIONS ET PERSPECTIVESERREUR ! SIGNET NON DEFINI.
8. BIBLIOGRAPHIEERREUR ! SIGNET NON DEFINI.
9. ANNEXESERREUR ! SIGNET NON DEFINI.

1. Spatialisation des sons par les techniques binaurales

La recreation d'un espace sonore (enregistré ou synthétisé) en trois dimensions est de plus en plus utilisée dans des domaines variés : le cinéma avec les salles à système de diffusion multi haut-parleurs, le home cinéma, les jeux vidéos, la musique avec un certain nombre de créations contemporaines "spatialisées", ou encore les télécommunications (téléconférence, réalité virtuelle, etc.). La spatialisation sonore peut apporter une plus value appréciable à une application car elle augmente le champ des informations captées et améliore l'ergonomie et le réalisme (sensation d'immersion dans l'espace sonore).

La stéréophonie, extrêmement répandue dans les systèmes d'écoute haute fidélité ou même certains téléviseurs, constitue une première approche de la spatialisation sonore puisqu'on l'on peut déjà distinguer des sources provenant de la gauche, du centre ou de la droite du plan défini par les deux haut-parleurs. Actuellement, il existe différentes techniques de spatialisation des sons plus évolués que la stéréophonie : les systèmes dits multi haut-parleurs comme le Dolby 5.1. utilisé en home cinéma, et les système dits binauraux n'utilisant que deux enceintes ou un casque d'écoute stéréo. Nous ne nous intéressons dans cette étude qu'aux techniques binaurales.

1.1. Introduction à la perception spatiale des sons

L'être humain peut localiser des sons dans l'espace avec une précision angulaire pouvant atteindre 1° [3]. Cette capacité est due au fait qu'un son émis par une source sonore atteint les deux oreilles avec des informations différentes selon la position de la source dans l'espace : on peut raisonnablement envisager qu'un son frontal va atteindre de manière similaire les oreilles gauche et droite, puisque la source est située dans le plan médian (qui est le plan contenant les points à égale distance des deux oreilles), alors qu'un son venant de la droite va être perçu d'abord par l'oreille droite puis par l'oreille gauche (qui est plus éloignée de la source). De plus, pour une source située à la droite de l'auditeur, le trajet source-oreille droite est un trajet en champ libre, alors que le trajet source-oreille gauche est "gêné" par la tête de l'auditeur ; il en résulte des différences d'informations fréquentielles entre les signaux reçus par les deux oreilles qui jouent un rôle majeur dans la capacité de localiser des sons dans l'espace.

Au début du 20^è siècle, Lord Rayleigh [18] a été le premier à expliquer comment l'être humain localise les sons : la localisation est basée sur le fait que les trajets sont différents pour les deux oreilles, d'où des différences de temps d'arrivée ; la notion de retard interaural ou ITD (Interaural Time Delay) est née. Rayleigh explicita également le rôle de la tête comme filtre acoustique produisant ainsi des différences interaurales d'intensité ou ILD (Interaural Level Difference). Le système auditif utilise les informations d'ITD aux basses fréquences (<1,5 kHz) et d'ILD aux hautes fréquences (au-delà de 1,5 kHz).

ITD et ILD sont deux indices binauraux significatifs pour la perception spatiale des sons. A ceux-ci viennent s'ajouter des indices monauraux codant principalement la dimension haut/bas (élévation) et avant/arrière. Ils peuvent aussi jouer un rôle pour la perception de l'azimut, et enfin pour la notion de distance à la source : intensité globale du signal et rapport entre champ direct et réverbération.

ITD et ILD sont cependant insuffisants pour discriminer complètement deux sources de provenances différentes. Il existe en effet une infinité de directions dans l'espace pour lesquelles tout signal provenant de l'une de ces directions produit un ITD constant.

L'ensemble de ces directions forme un cône d'axe interaural et de sommet l'entrée des conduits auditifs de l'une des deux oreilles. Cet ensemble est appelé cône de confusion [13]. Pour des signaux ne comportant que des basses fréquences, la discrimination sur un cône de confusion est très difficile. Situés sur un même cône de confusion, un son provenant de l'avant et un son provenant de l'arrière sont peu discernables : c'est le phénomène de confusion avant arrière. Ce phénomène s'estompe lorsque l'on considère les mouvements dynamiques de tête : l'ITD d'une source frontale et l'ITD d'une source située à l'arrière varient en sens inverse lorsque l'on tourne la tête de gauche à droite [4]

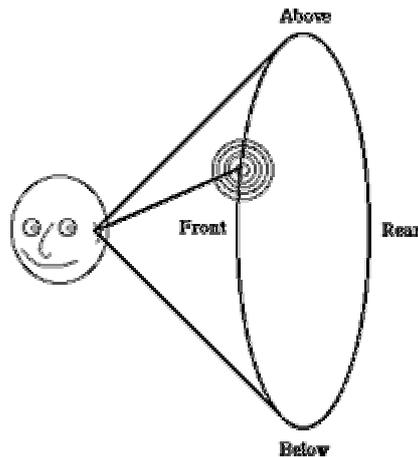


Figure 1 - Cône de confusion

Pour décrire de manière exhaustive la relation entre un son émis dans une certaine direction et les signaux captés par les deux oreilles, on définit des fonctions de transfert acoustiques entre la source sonore et l'oreille gauche, et entre la source sonore et l'oreille droite. Ces fonctions sont communément appelées HRTF (Head Related Transfer Functions). On peut ensuite calculer les indices binauraux tels qu'ILD et ITD à partir de ces HRTF. Si l'ITD et l'ILD donnent lieu à une détection latérale des événements sonores, les HRTF apportent des indices spectraux qui complètent la localisation directionnelle dans le plan médian et résolvent notamment les confusions avant arrière et affinent la position verticale.

Les HRTF peuvent varier de manière significative d'un individu à l'autre car elles sont le reflet de la morphologie de chaque individu : en effet, la diffraction des ondes acoustiques sur la tête dépend des dimensions de la boîte crânienne et de la géométrie de l'oreille externe (présence de modes propres). En ce qui concerne l'oreille externe plus communément appelée pavillon, des variations de l'ordre du millimètre sont significatives.

Nous avons introduit les indices de localisation utiles en techniques binaurales : ITD, ILD, HRTF, cône de confusion, mouvements dynamiques de tête. Voyons maintenant comment elles sont employées dans les systèmes de spatialisation binauraux.

1.2. Principe et architecture

1.2.1. Principe

Les techniques binaurales sont la conséquence d'un raisonnement simple : si nous avons deux capteurs (nos deux oreilles), pour reproduire un espace sonore il suffit de deux sources sonores (une pour chaque oreille) présentant directement aux oreilles les informations sonores qu'elles auraient reçues si le sujet avait pris place dans l'espace sonore naturel.

On utilise généralement un casque stéréophonique pour la restitution binaurale, ce qui présente l'avantage de bien faire correspondre un écouteur pour une oreille, sans interférence entre l'écouteur gauche et l'oreille droite (ce qui se présente lors d'une écoute sur enceintes).

On peut également restituer des signaux binauraux sur deux enceintes en utilisant un filtrage transaural qui a pour but d'annuler les chemins croisés HP gauche / oreille droite et HP droit / oreille gauche.

On peut obtenir des signaux binauraux par deux moyens : prise de son ou synthèse. La prise de son est extrêmement simple à effectuer puisqu'il suffit de placer une capsule microphonique à l'entrée de chaque pavillon. On peut restituer alors l'enregistrement après égalisation des réponses des micros et du système d'écoute comme le montre la figure 2.

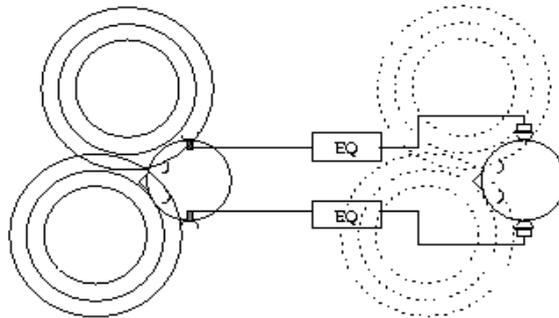


Figure 2 - Principe de restitution binaurale à partir d'un enregistrement

La synthèse binaurale s'effectue quant à elle par convolution d'un signal monophonique avec une paire de HRTF : on crée ainsi une source virtuelle de direction dépendant de la paire de HRTF.

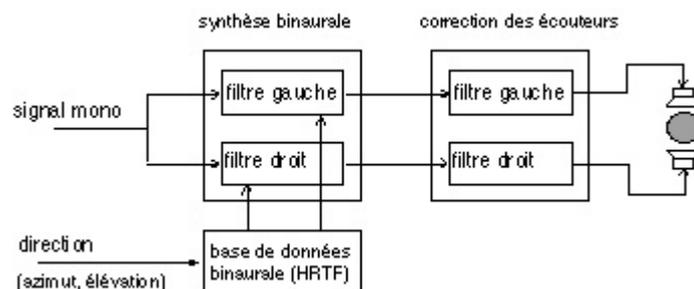


Figure 3 : principe de la synthèse binaurale : La direction de la source sonore est contrôlée à l'aide de deux filtres numériques déduits des HRTF

1.2.2. Architecture

Pour pouvoir simuler un espace sonore de composition arbitraire, il faut disposer d'un jeu de HRTF couvrant toutes les directions de l'espace. Ces HRTF sont mesurées en chambre sourde sur des têtes réelles ou artificielles (KEMAR, HEAD ACOUSTICS, etc.) sous forme de réponses impulsionnelles (HRIR ou Head Related Impulse Response).

La réalisation la plus simple consiste à convoluer chaque source monophonique avec une paire de HRIR choisie, ce qui pratiquement s'effectue de préférence dans le domaine fréquentiel par multiplication des transformées de Fourier rapides (FFT).

Pour optimiser l'architecture de la synthèse binaurale, l'idée est de séparer les informations temporelles (type ITD) et spectrales contenues dans les HRTF. Chaque HRTF est décomposée en produit d'un filtre à phase minimale $L(f)$ (contenant les informations spectrales) et d'un filtre passe-tout $\tau(f)$ (contenant les informations temporelles).

$$\begin{aligned}H_L(f) &= \tau_L(f).L(f) \\ H_R(f) &= \tau_R(f).R(f)\end{aligned}$$

Les travaux de Whitman et Kistler [11] ont montré que l'on peut modéliser les composantes passe-tout par des retards purs, indépendants de la fréquence donc peu coûteux en temps de calcul. On obtient ainsi :

$$ITD = \tau_L - \tau_R$$

L'amplitude $|H_L(f)|$ et la phase $\phi_L(f)$ d'un filtre à phase minimale sont reliées par la relation :

$$\phi_L(f) = \text{Im}(\text{Hilbert}(-\log(|L(f)|)))$$

Une paire de HRTF s'écrit donc :

$$\begin{aligned}H_L(f) &= e^{j.\omega.ITD} . |H_L(f)| . e^{j.\phi_L(f)} \\ H_R(f) &= |H_R(f)| . e^{j.\phi_R(f)}\end{aligned}$$

1.3. Limitations des systèmes actuels

Pour bien comprendre les limitations actuelles des techniques binaurales, il faut se rendre compte du caractère individuel des HRTF : imaginons qu'un sujet A enregistre des signaux binauraux à l'aide de deux microphones placés à l'entrée de ses conduits auditifs et fasse écouter cet enregistrement à un sujet B. Le sujet B aura des informations spatiales et pourra sentir les événements sonores se déplacer autour de lui, mais aura généralement du mal à placer les source devant lui (problème de perception frontale : le sujet entend à l'arrière une source qui est devant). On peut dire que les HRTF du sujet A ne sont pas adaptées au sujet B.

C'est pourquoi on utilise souvent pour les enregistrements binauraux des têtes artificielles qui sont sensées représenter une bonne moyenne des paramètres morphologiques humains. Malheureusement très peu de têtes artificielles (mis à part VALDEMAR à l'Acoustics Laboratory à Aalborg, Danemark, [5]) sont conçues selon des paramètres perceptifs tels que l'ITD.

Des erreurs sur l'ITD de l'ordre de 30 microsecondes sont perceptibles, donc un modèle générique basé sur une moyenne de l'ITD d'un grand nombre d'individus ne peut convenir à tout le monde (les amplitudes de variation de l'ITD atteignant quelques millisecondes).

Pourquoi ne pas mesurer alors les HRTF de tous les sujets afin d'avoir toujours une spatialisation individualisée ? Une série de mesure de HRTF est extrêmement longue (plus de 4h par sujet) et éprouvante, et nécessite du matériel de mesure très coûteux dont seuls quelques laboratoires de recherche disposent. Cette solution n'est pas envisageable pour des applications grand public.

De nombreuses recherches ont donc été développées dans la voie de la simulation de HRTF ou de l'ITD par des modèles géométriques simples et adaptatifs, que l'on pourrait paramétrer en fonction de critères morphologiques simples tirés de deux portraits d'un sujet donné (une de face et une de profil).

2. Simulation de HRTF par modèles physiques

2.1. Problématique

Pour éviter de mesurer les HRTF d'un sujet donné, il faut développer des modèles physiques de tête qui se rapprochent le plus de la morphologie du sujet. Le modèle sert à calculer un jeu de HRTF sensé se rapprocher du jeu de HRTF du sujet. Cette démarche soulève quelques interrogations :

Quel type de modèle choisir pour représenter une tête ?

Quelle est la précision que l'on peut atteindre avec ces modèles simples, en termes d'ITD, d'ILD et d'indices spectraux ?

Est-ce qu'un modèle simple simule les HRTF avec suffisamment de précision ?

Quelle est l'influence du pavillon, faut-il le modéliser plus finement que le reste de la tête ?

Quelle est l'influence du torse et des épaules ?

La ressemblance morphologique est-elle un critère réellement pertinent ?

2.2. Etat de l'art

Différents modèles géométriques ont été utilisés pour modéliser une tête humaine. Le plus simple et le plus employé jusqu'ici est le modèle sphérique. Néanmoins, les dimensions moyennes d'une tête montrent que des modèles moins symétriques qu'une sphère seraient plus adaptés à la morphologie humaine, et c'est pourquoi des modèles plus complexes comme les sphéroïdes oblates (sphère dont un axe est étiré) ou les ellipsoïdes, sensés pouvoir représenter plus finement une tête humaine, ont été également étudiés.

2.2.1. Solutions analytiques

Lord Rayleigh a développé à la fin du 19^e siècle les solutions analytiques de la diffraction d'une onde plane sur une sphère rigide, en décomposant les champs de pression et de vitesse sur une base d'harmoniques sphériques (se reporter à [18] pour plus de détails). Bon nombre d'études ont utilisé ses résultats pour créer des modèles sphériques de tête (Blauert [3], Martens [12], etc.).

La valeur du rayon la plus souvent employée est 87,5 mm, moyenne mesurée par Hartley et Fry en 1921, sur un panel d'individus adultes des Etats Unis. Cette valeur est susceptible d'augmenter légèrement avec la croissance de la taille moyenne des êtres humains.

Les Japonais Fukudome et Yamada ont étudié l'influence des dimensions de modèles de sphères et de sphéroïdes oblates sur la pression acoustique à l'entrée des conduits auditifs, dans le but de concevoir une tête artificielle qui soit une bonne moyenne acoustique. [7] présente les équations pour calculer la pression et l'impédance acoustique de Thévenin sur une sphère ou une sphéroïde oblate.

Toutes ces études utilisent pour le modèle sphérique ou sphéroïdique des paramètres basés sur des moyennes morphologiques, donc non individuelles. S'intéressant également aux modèles

sphériques, Woodworth [20] a développé une formule d'approximation de l'ITD d'une sphère basée sur les trajectoires des ondes sonores, et non sur la résolution de l'équation de Helmholtz comme l'a fait Rayleigh : si θ est l'azimut en radians (se référer à l'Annexe 1 pour le système d'angles et de coordonnées), c la vitesse de propagation du son et a le rayon de la sphère, alors :

$$ITD = \frac{a}{c}(\sin \theta + \theta)$$

Cette approximation dans le plan horizontal est excellente et est largement utilisée en synthèse binaurale. La formule a été étendue par Larcher et Jot [8] à tout l'espace. Si ϕ est l'élévation, alors :

$$ITD = \frac{a}{c}[\arcsin(\cos \phi \cdot \sin \theta) + \cos \phi \cdot \sin \theta]$$

La formule de Woodworth a été utilisée par Algazi et al. [2] pour estimer le rayon optimal d'un modèle sphérique en fonction de paramètres morphologiques d'un individu donné. Le rayon optimal est dérivé d'une procédure de minimisation d'erreur entre des mesures d'ITD et les valeurs données par la formule de Woodworth. Le rayon optimal est ensuite relié aux mesures anthropométriques par régression statistique. On obtient ainsi une formule reliant directement l'ITD à trois paramètres anthropométriques X_1 , X_2 et X_3 (en mm):

$$ITD = \frac{0.51X_1 + 0.019X_2 + 0.18X_3 + 32}{c}[\arcsin(\cos \phi \cdot \sin \theta) + \cos \phi \cdot \sin \theta]$$

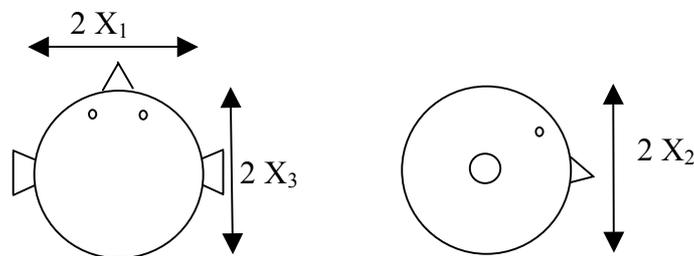


Figure 4 : paramètres anthropométriques mesurés sur deux photos

Cette dernière formule est donc une première approche d'individualisation de l'ITD en fonction de paramètres simples. Duda [6] a développé une méthode par tracé de rayons sur une ellipsoïde pour trouver une relation du même type. Malheureusement, le calcul du chemin acoustique sur l'ellipsoïde est beaucoup plus complexe.

Une autre piste a également été explorée : la résolution de l'équation de Helmholtz sur une ellipsoïde rigide. Le problème semble de même nature que le problème sphérique résolu par Rayleigh, mais pour pouvoir développer la solution il faut trouver une base de fonctions spatiales adaptées au système de coordonnées ellipsoïdale : les harmoniques ellipsoïdales. Des chercheurs russes (Abramov, Konyukhova, Levitina [1]) ont mis au point des méthodes de séparation des variables en coordonnées ellipsoïdales mais n'ont de résultats sur le champ

acoustique diffracté qu'en champ lointain et selon certaines incidences privilégiées. Pour avoir une estimation précise des solutions sur l'ellipsoïde, il faut donc recourir aux méthodes numériques.

2.2.2. Solutions numériques

L'augmentation continue de la puissance des microprocesseurs a permis l'avènement de méthodes de simulation numérique très coûteuses en calcul et en mémoire vive comme le calcul par éléments finis de frontière (Boundary Element Method ou BEM), basé sur une décomposition d'un objet physique en un ensemble d'éléments discrets. Les premiers calculs de HRTF par éléments finis ont débuté vers 1984 avec les travaux de Weinrich [19].

Les travaux de thèse de Katz [10] se sont focalisés sur la simulation numérique de HRTF par éléments finis, pour la plage de fréquence 0-8000 Hz. Différents modèles ont été utilisés : un maillage de la tête de Katz créé à partir d'un scan 3d, et deux modèles de sphère : sphères, de rayon la distance interaurale, et sphère conservant le volume de la tête. Katz a également évalué l'influence de l'impédance de la peau et des cheveux. Le logiciel utilisé était LMS Sysnoise et la taille des maillages BEM pouvait atteindre 30 000 éléments.

Katz a obtenu des résultats encourageants : la simulation BEM est une voie porteuse pour étudier l'influence de la géométrie de la tête sur les HRTF. La simulation permet d'étudier séparément l'influence de parties distinctes de la tête selon le type de maillage (avec ou sans pavillons, avec ou sans détails de visage, etc.) mais le principal inconvénient est le temps de calcul (plusieurs dizaines d'heures pour un gros modèle).

Les travaux de Kahana [9] ont exploré plus avant les possibilités de calcul de HRTF par BEM. Kahana a calculé les HRTF d'une ellipsoïde de 1 à 5000 Hz par pas de 100 Hz, selon toutes les directions de l'espace. Il a aussi calculé des HRTF de sphère, et s'est beaucoup intéressé à la question suivante : peut-on calculer avec précision le filtrage effectué par les pavillons en utilisant la simulation numérique ? Il a développé plusieurs modèles numériques de pavillons dont il a comparé les résultats aux mesures des modèles de pavillons en chambre sourde, et a obtenu des résultats très concluants. Enfin, il a apporté quelques réponses sur l'influence des épaules et du torse sur les HRTF basse fréquence (<3 kHz) dont nous parlerons dans le prochain paragraphe.

La simulation par BEM est maintenant reconnue comme un outil puissant et fiable de calcul de HRTF pour une variété importante de modèles, ce qui ouvre la voie à des tests précis sur l'influence de la ressemblance morphologique entre modèle et tête réelle, sur l'influence du torse et des pavillons sur l'ITD, et d'avoir une idée plus précise sur les possibilités pratiques d'individualisation des HRTF qui finalement restent encore assez floues.

2.2.3. Une structure modulaire pour la simulation de HRTF

L'avantage des simulations numériques est de pouvoir dissocier différentes parties du corps pour étudier séparément leur influence et ensuite reconstituer un modèle modulaire complet à partir des différents éléments séparés. L'Etat actuel des recherches est encore loin de ce but mais quelques résultats ont déjà été mis en lumière quant à l'influence du torse, des pavillons et des cheveux (cf. Katz, Kahana). Le schéma suivant les résume succinctement :

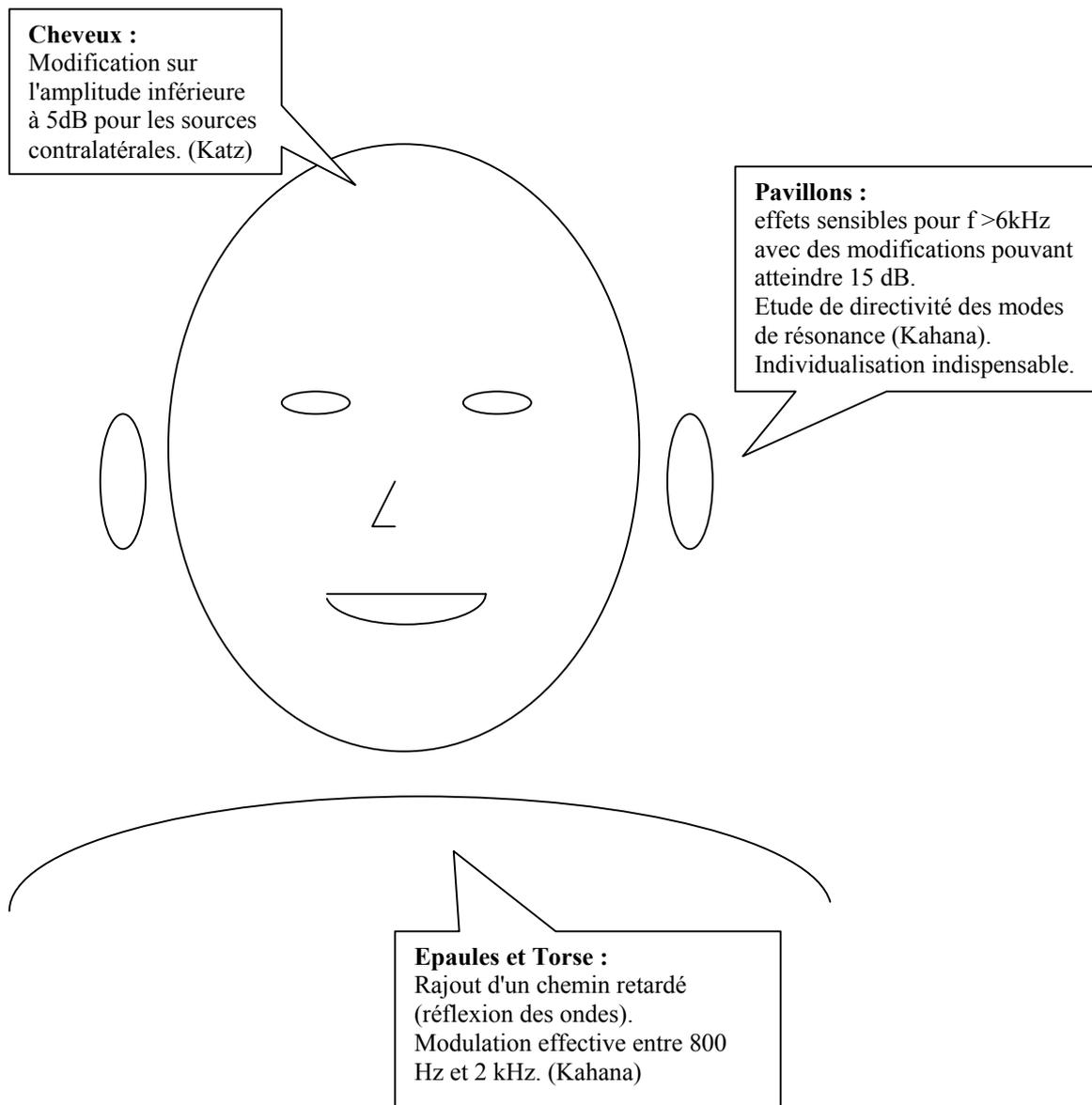


Figure 4 : structure d'étude modulaire pour la contribution d'éléments morphologiques séparés

Nous nous limiterons pour des contraintes de temps (effectif et temps de calcul) et de moyens dans le cadre de cette étude à la modélisation d'une tête sans pavillons ni torse, pour étudier l'influence de l'individualisation de cette partie du modèle structural en termes d'indices binauraux.

3. Calculs de HRTFs de modèles géométriques simples par la méthode des éléments finis

3.1. Principes du calcul par éléments finis

La plupart des problèmes de physique peuvent se formuler ainsi : trouver un champ (scalaire, vectoriel, ou tensoriel) $u(M, t)$ satisfaisant à un ensemble d'équations aux dérivées partielles et d'équations ordinaires en tout point M d'un domaine Ω et à tout instant t , et respectant des conditions aux limites (éventuellement fonction du temps) sur la frontière $\delta\Omega$ du domaine. Les conditions aux limites sont des relations ou des valeurs imposées à u et/ou à ses dérivées sur la frontière. Si le temps t n'apparaît pas comme variable, on dit que le problème est *stationnaire*.

La méthode par éléments finis consiste à rechercher une solution *approchée* de la solution exacte sous la forme d'un champ $F'(M, t)$ défini par morceaux sur des sous domaines de Ω . Les n sous domaines Ω_i doivent former une partition de Ω .

Les champs $F'_i(M, t)$, définis sur chaque sous domaine, sont des champs choisis parmi une famille *arbitraire* de champs (généralement polynomiaux). La famille de champs locaux est appelée *espace des fonctions d'interpolation de l'élément*. La famille de champs globaux $F_i(M, t)$, obtenus par juxtaposition des champs locaux est appelée *espace des fonctions d'interpolation du domaine Ω* .

Le champ dans chaque sous domaine Ω_i est déterminé par un nombre fini de valeurs du champ (ou de valeurs de ses dérivées) en des points choisis arbitrairement dans le sous domaine, et appelés nœuds. Le champ local est une *interpolation* entre les valeurs aux nœuds. Le sous domaine muni de son interpolation est appelé *élément*.

Chercher une solution par éléments finis consiste donc à déterminer quel champ local on attribue à chaque sous domaine pour que le champ global $F'(M, t)$ obtenu par juxtaposition de ces champs locaux soit proche de la solution du problème. Parmi les contraintes qu'on impose à la solution approchée cherchée, il y a souvent au moins une continuité simple (C_0) à la frontière entre les sous domaines. Les figures suivantes montrent différents types de solutions approchées : discontinues, C_0 , C_1 , etc.

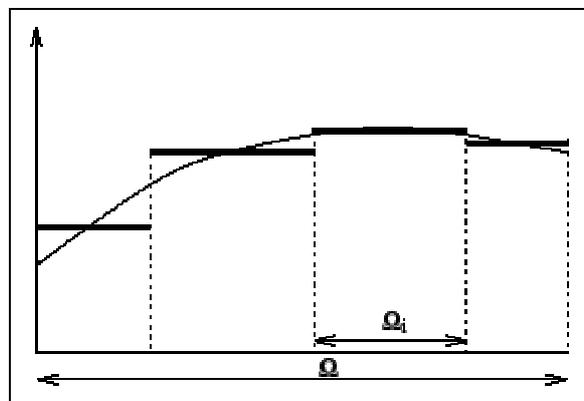


Figure 5 : solution discontinue

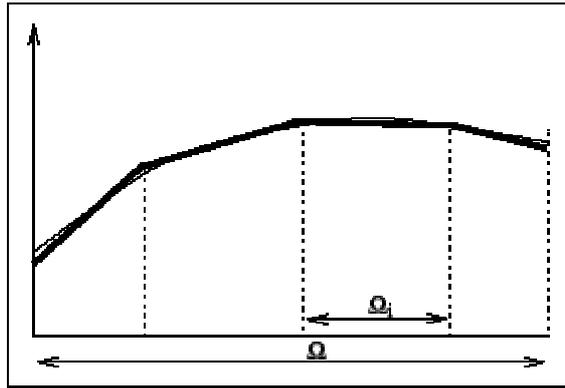


Figure 6 : solution continue C0

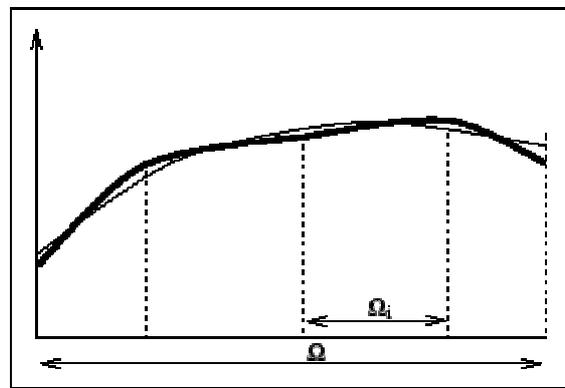


Figure 7 : solution continue C1

La qualité de la solution approchée dépend de la division en sous domaines (nombre et dimensions des sous domaines), du choix de la famille de champs locaux dans chaque sous domaine, et des conditions de continuité qu'on impose aux frontières des sous domaines (C_0 , C_1, \dots). Une fois ces choix faits, il reste à rechercher, une combinaison de champs locaux qui satisfait approximativement les équations.

Pour résoudre un problème par la méthode des éléments finis, on procède donc par étapes successives :

1. On se pose un problème physique sous la forme d'une équation différentielle ou aux dérivées partielles à satisfaire en tout point d'un domaine Ω , avec des conditions aux limites sur le bord $\delta\Omega$ nécessaires et suffisantes pour que la solution soit unique.
2. On construit une formulation intégrale du système différentiel à résoudre et de ses conditions aux limites : C'est la *formulation variationnelle du problème*.
3. On divise Ω en sous domaines : C'est le *maillage*. Les sous domaines sont appelés *mailles*.
4. On choisit la famille de champ locaux, c'est à dire à la fois la position des nœuds dans les sous domaines et les polynômes (ou autres fonctions) qui définissent le champ local en fonction des valeurs aux nœuds (et éventuellement des dérivées). La maille complétée par ces informations est alors appelée *élément*.
5. On ramène le problème à un problème discret : C'est la *discrétisation*. En effet, toute solution approchée est complètement déterminée par les valeurs aux nœuds des

éléments. Il suffit donc de trouver les valeurs à attribuer aux nœuds pour décrire une solution approchée. Le problème fondamental de la méthode des éléments finis peut se résumer en deux questions :

- (a) Comment choisir le problème discret dont la solution est <<proche>> de la solution exacte?
 - (b) Quelle signification donner au mot <<proche>> ?
6. On résout le problème discret: C'est la *résolution*
 7. On peut alors construire la solution approchée à partir des valeurs trouvées aux nœuds et en déduire d'autres grandeurs : C'est le *post-traitement*.
 8. On visualise et on exploite la solution pour juger de sa qualité numérique et juger si elle satisfait les critères du cahier des charges : C'est *l'exploitation des résultats*.

Les étapes 1,2,3,4 et 5 sont souvent rassemblées sous le nom de *pré traitement*.

Le travail de ces différentes étapes est assisté par les logiciels. Il reste que pour maîtriser leur utilisation et régler astucieusement les paramètres de calcul, il est indispensable de comprendre les fondements de la méthode.

3.2. Utilisation du logiciel VNoise 1.11

VNoise 1.11 est un logiciel de calcul par éléments finis, édité par STS Software, qui permet la simulation numérique de problèmes acoustiques. Son utilisation est simple pour qui connaît déjà un tant soit peu les logiciels de conception 3D (autoCad ou Systus...)

VNoise permet une variété de conceptions géométriques, et inclut des excitations telles que forces de contact ou bien sources acoustiques de différentes directivités. Différents points de mesure sont représentés par des microphones (les résultats donnés sont : champ de pression, champ de vitesse, et dp/dn .)

La partie calculatoire en elle-même peut s'effectuer de manière automatique (réglages par défauts) mais pour des calculs de grosse taille (maillages de plus de 5000 nœuds) elle nécessite une configuration adaptée, celle-ci s'apprenant plus par expérience...

Les caractéristiques du logiciel sont les suivantes :

Maximum number of nodes	30000
Maximum number of panels	60000
Maximum number of structural modes	50
Maximum number of acoustic sources	50
Maximum number of acoustic materials	30
Maximum number frequencies used to define acoustic materials and sources	1000

Regardons les fonctionnalités de VNoise selon les étapes décrites dans le paragraphe 3.1. : les étapes 1 et 2 sont fixées dans le logiciel puisqu'il s'agit de résoudre les équations de

propagation acoustique en tenant des conditions aux limites telles que : paroi rigide, source vibrante, source dipolaire, etc.

Les étapes 3 et 4 sont réalisées par le mailleur 3d intégré de VNoise : un éditeur graphique permet de définir simplement et rapidement des maillages simples. Pour des maillages complexes (ellipsoïdes, scans) on utilisera un mailleur externe (logiciel de design en 3 dimensions 3D Studio Max).

L'étape de discrétisation est assurée par un algorithme intégré au logiciel. Les paramètres à définir sont le nombre d'éléments par longueur d'onde à une fréquence donnée : pour atteindre une bonne précision, Katz et Kahana recommandent d'utiliser au moins 6 éléments par longueur d'onde.

L'étape 6 (résolution) est celle qui reste la moins intuitive sous VNoise. L'influence des paramètres des solveurs n'est pas évidente au premier abord et bien régler les différentes valeurs nécessite quelque peu d'expérience.

Enfin, les étapes de post traitement s'effectueront sous Matlab après export des résultats de l'étape 6.

Les deux étapes fondamentales réalisées par VNoise sont donc discrétisation et calcul. Celles-ci sont étroitement reliées car la durée du calcul dépend du carré du nombre de nœuds de la discrétisation, et la précision des résultats augmente avec le nombre d'éléments par longueur d'onde. Il faut donc trouver un compromis entre précision et temps de calcul, le nombre de nœuds dans la discrétisation est donc crucial.

3.3. Le principe de réciprocité

3.3.1. Énoncé

Habituellement, une mesure de HRTF s'effectue en plaçant un microphone dans le conduit auditif du sujet, et en envoyant un signal de mesure depuis diverses positions de l'espace (965 dans le cas des mesures effectuées par France Télécom R&D au TNO). Pour une simulation BEM, on procède de la même manière. L'inconvénient est que pour chaque position de la source, le logiciel utilisé (VNoise 1.11) effectue un calcul par fréquence, et ceci prend entre 5 et dix heures par position de la source. Pour simuler les 965 positions de mesure, le temps de calcul est énorme (5000 h de calcul minimum).

Une méthode inverse est suggérée par le principe de réciprocité : la pression acoustique p_1 produite au point B par la source A de vitesse acoustique v_1 est égale à la pression acoustique p_2 produite au point A par la source B de vitesse acoustique v_2 :

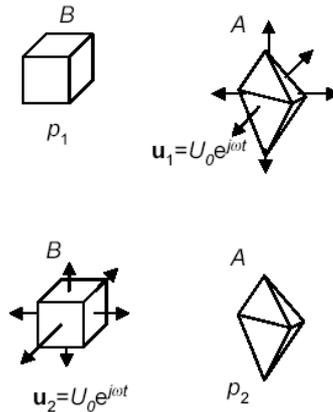


Figure 8 : le principe de réciprocité

L'application pratique du principe de réciprocité dans le logiciel VNoise semble simple (on place une source près de l'oreille, et une sphère de microphones autour du modèle à 1 mètre de distance), elle pose cependant quelques difficultés :

On ne peut pas placer la source acoustique directement sur la surface de maillage, à l'endroit de mesure précédent, car une partie de la puissance acoustique est alors perdue à l'intérieur du maillage. Il faut alors placer la source à proximité de la surface (distance de l'ordre du millimètre, ce qui correspond à une longueur d'onde acoustique pour une fréquence de l'ordre de 340kHz, négligeable devant les longueurs d'onde de la bande audible : 1,7 cm pour 20kHz).

Plusieurs essais ont été menés avec des distances source-surface différentes (2mm, 1mm, 0.1mm) et pour des fréquences inférieures à 13kHz, les différences sont négligeables. Placer une source à proximité de la surface pose cependant des problèmes de convergence intégrale dans les solveurs du logiciel. Nous allons étudier ce problème dans la prochaine partie.

Enfin, la question d'un raffinement global du maillage (augmentation du nombre d'éléments par longueur d'onde) se pose également, car il semble normal de devoir utiliser un modèle plus précis pour un calcul qui en remplace 965 !

3.3.2. Définition des modèles directs et réciproques sous VNoise

Pour chacun des modèles, on considère une sphère de rayon 0,0975 mètre, ce qui représente le rayon moyen d'une tête humaine, si on la considère comme sphérique. Les réponses amplitude-fréquence sont normalisées dans un cas comme dans l'autre par la réponse en champ libre (système identique sans le maillage représentant la tête).

Les calculs avec le modèle direct et le modèle réciproque sont tous deux comparés aux résultats théoriques de la diffraction d'une onde plane sur une sphère rigide explicités par Lord Rayleigh [18].

Un schéma explicite les normalisations effectuées pour les deux modèles, à gauche le modèle direct et à droite le modèle réciproque :

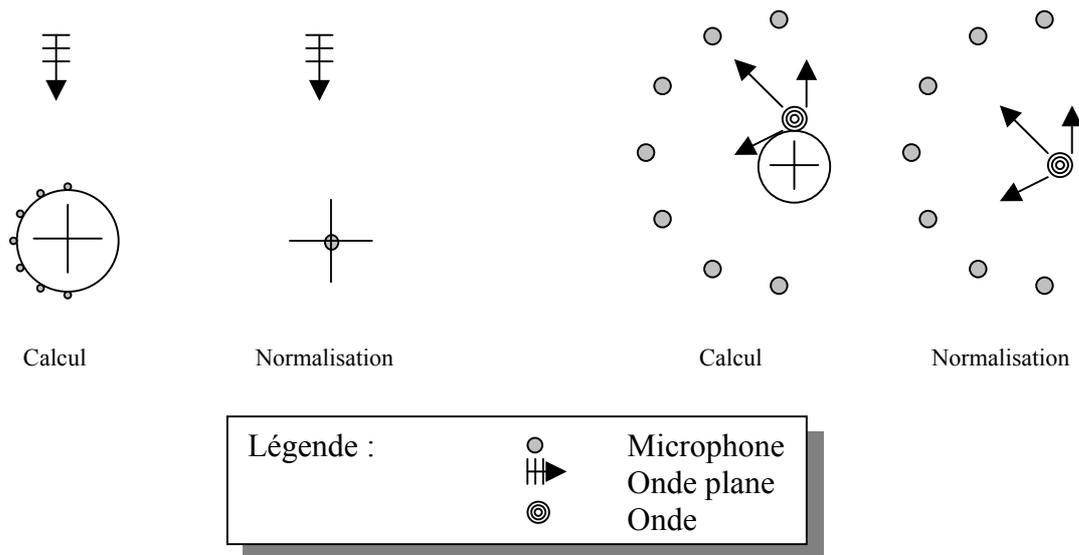


Figure 9 : schéma de normalisation en calcul direct et réciproque

Pour pouvoir comparer les résultats du modèle réciproque avec les solutions analytiques, il faut placer les microphones à une distance d'au moins 1,5 m, pour alors considérer que le système direct équivalent est une source monopolaire distante de 1,5 m de la tête et faire alors l'approximation en champ lointain (les ondes sphériques incidentes sont considérées comme planes). C'est un point à ne pas oublier lors de l'analyse des résultats. Il faut remarquer que cette approximation est effectuée lors des mesures de HRTF sur des sujets à l'aide de sources électrodynamiques.

3.3.3. Intérêt du raffinement local de maillage

Les travaux de Yuvi Kahana sur le calcul de HRTFs par modélisation à éléments finis [9] lui ont permis de dégager un certain nombre de "règles" à employer :

Il faut effectuer un maillage plus précis à proximité de la source, c'est à dire quand la distance source élément devient comparable à la longueur d'onde de la fréquence calculée, ceci pour éviter les divergences dans les différentes intégrations effectuées par le solveur.

La taille des nouveaux éléments doit être de l'ordre de grandeur de la distance normale source-surface (règle du "pouce").

Pour évaluer l'intérêt de ce nouveau maillage, deux calculs ont été effectués, l'un avec le maillage originel et l'autre avec raffinement local du maillage :

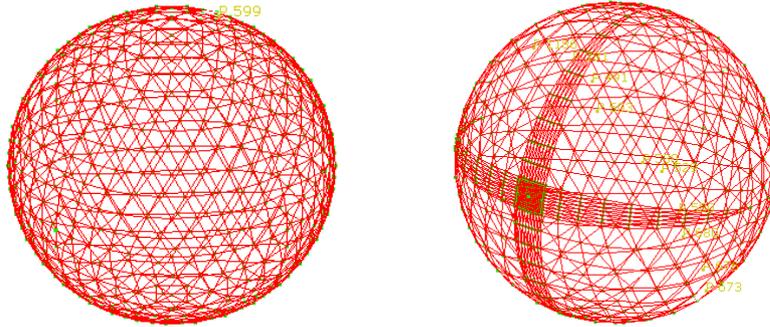


Figure 10 : maillage grossier et maillage avec raffinement local

La comparaison des deux modèles peut se résumer comme suit :

	Nombre nœuds	Temps de calcul	Fréquences calculées
Modèle grossier	578	34 min	30 fréquences de 20 à 5000 Hz
Modèle avec raffinement	783	32 min	30 fréquences de 20 à 5000 Hz

Si au niveau du temps de calcul, les deux modèles sont très comparables, les résultats obtenus sont en revanche très différents. Comme le montre les courbes ci-après, le modèle grossier ne suit absolument pas les résultats analytiques de la diffraction d'une onde acoustique sur une sphère rigide :

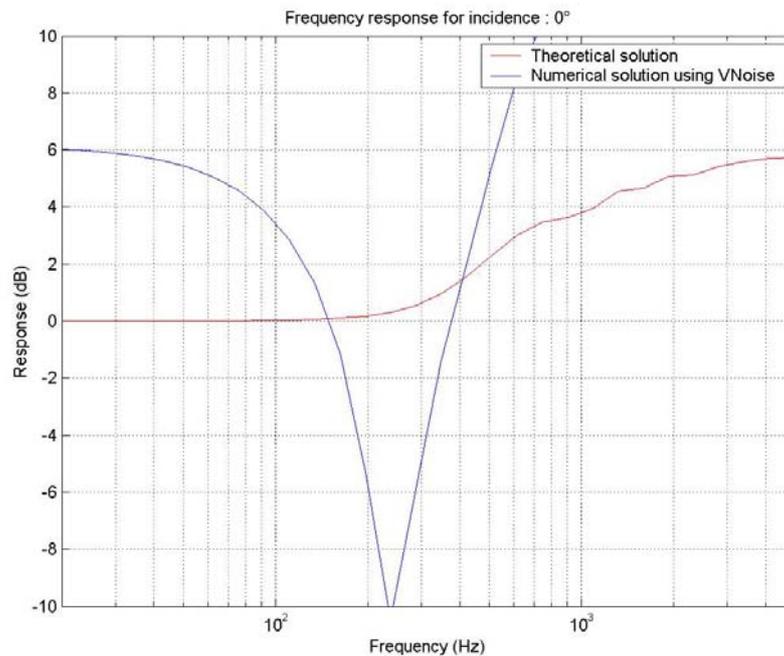


Figure 11 : réponse en fréquence pour l'incidence 0° du modèle grossier

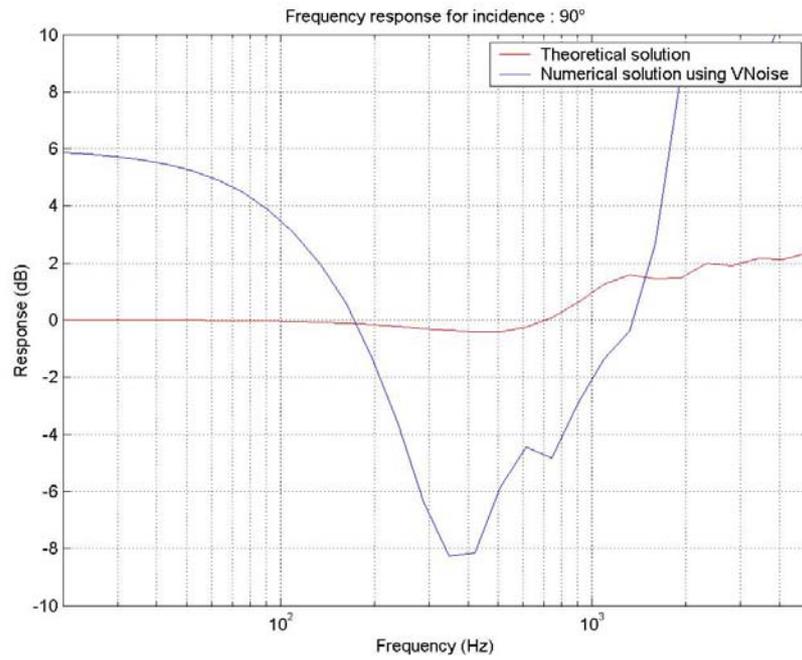


Figure 12 : réponse en fréquence pour l'incidence 90° du modèle grossier

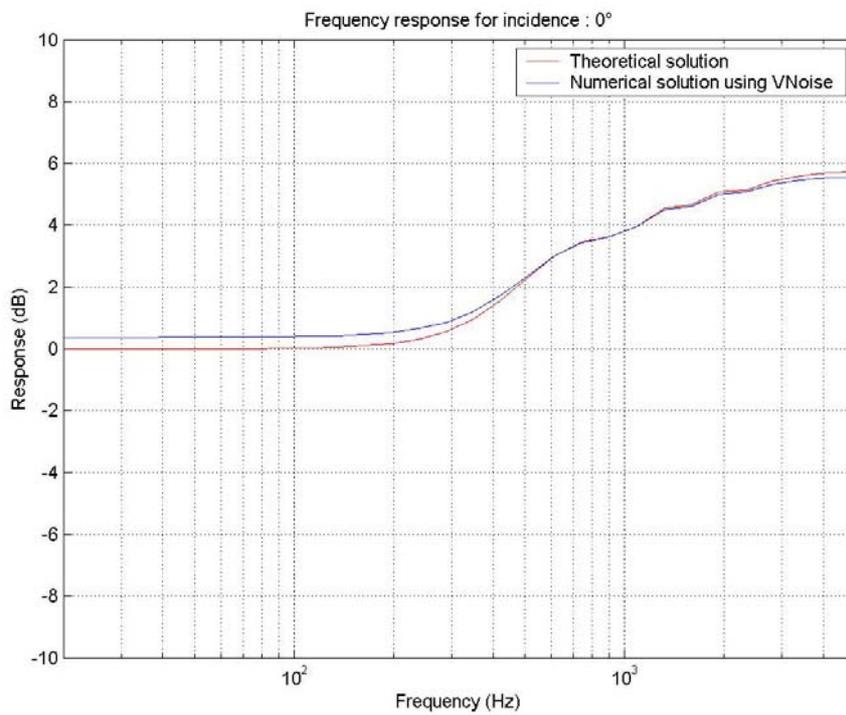


Figure 13 : modèle avec raffinement du maillage, incidence 0°

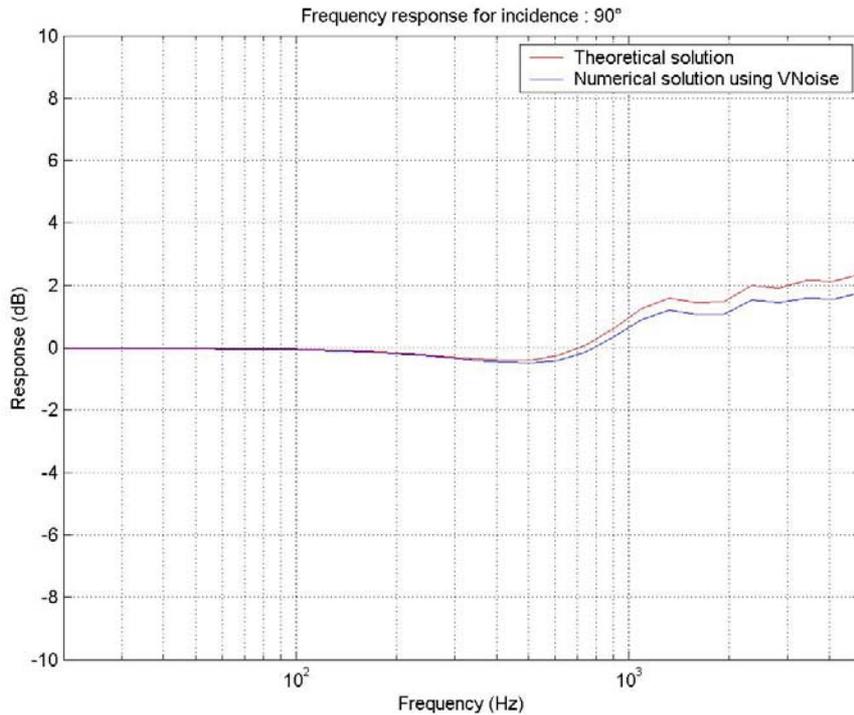


Figure 14 : modèle avec raffinement du maillage, incidence 90 °

3.4. Contraintes matérielles

Les deux principales contraintes d'utilisation d'un logiciel de calcul par éléments finis sont : la taille des maillages et temps de calcul. Les explications précédentes mettaient en lumière l'influence de la taille du maillage sur la précision et le temps de calcul. Nous allons maintenant nous placer du point de vue matériel.

Un calcul par éléments finis nécessite une quantité importante de mémoire. Cette quantité évolue approximativement comme le carré du nombre de nœuds du maillage comme le montre la figure suivante :

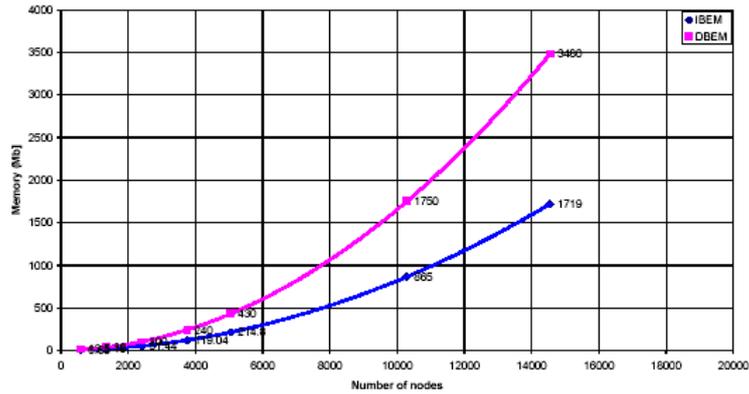


Figure 15 : Mémoire RAM requise par le solveur en fonction du nombre de nœuds (IBEM et DBEM sont deux méthodes de calcul différentes)

Si la quantité de RAM disponible est inférieure à la quantité de RAM nécessaire, le solveur passe en mode 'disk solving', c'est à dire qu'il utilise le disque dur (et non plus la RAM) pour stocker ses données. Or les temps d'accès entre un disque dur (même SCSI) et la RAM sont très différents : 10 nanosecondes environ pour la RAM, quelques microsecondes pour un disque dur SCSI. Le 'disk solving' est par conséquent beaucoup plus lent que le 'RAM solving'.

La configuration matérielle de la station de calcul est la suivante :

- Intel Xeo 1,7 GHz bi-processeur
- 1,5 Go de RAM

Il faut remarquer qu'un calcul avec VNoise n'utilise que 50% de la puissance CPU. On peut donc lancer deux calculs VNoise en même temps pour occuper la pleine puissance des processeurs. Seulement, les deux solveurs devront partager la même RAM.

Finalement, deux configurations ont été utilisées :

- Pour les maillages de faible taille (nécessitant moins de 700 Mo de RAM), deux calculs VNoise étaient lancés simultanément; la quantité de RAM étant suffisante pour chacun d'eux.
- Pour des maillages de grosse taille (nécessitant plus de 700 Mo de RAM), un unique calcul était lancé.

3.5. Validation du logiciel VNoise 1.11

Pour valider le bon fonctionnement du logiciel VNoise 1.11, qui est très récent et peut donc présenter des défauts de jeunesse, nous avons comparé les résultats analytiques (formule de Rayleigh) et numériques (BEM) de la diffraction d'une onde plane sur une sphère rigide.

Considérons le système suivant :

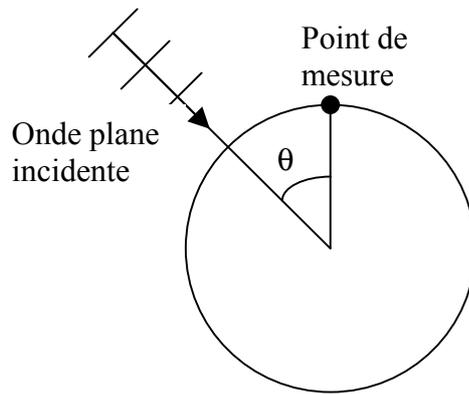


Figure 16 : description du modèle de validation VNOISE

Une onde plane arrive avec une incidence θ par rapport au point de mesure sur une sphère rigide de rayon $r = 87,5$ mm. La pression acoustique résultante au point de mesure est normalisée par la pression acoustique en champ libre au centre de la sphère.

Les incidences choisies pour la comparaison théorique / numérique vont de 0 à 180° par pas de 15° . La comparaison s'effectue sur 88 fréquences s'échelonnant de manière quasi logarithmique de 20 à 14000 Hz :

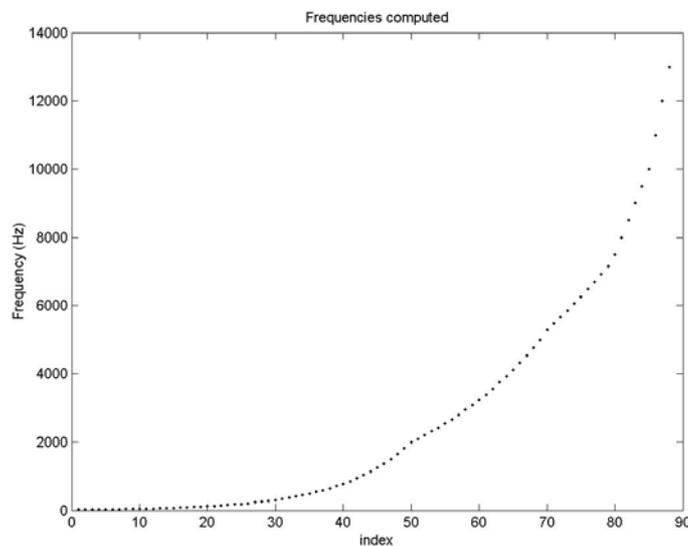


Figure 17 : fréquences calculées pour la validation VNOISE

Pour ce faire, trois discrétisations différentes ont été utilisées : une première discrétisation à 6 éléments par longueur d'onde pour $f=5000$ Hz (notation raccourcie : $f5000n6$) puis une deuxième discrétisation $f10000n6$ et une troisième $f15000n6$.

Les temps de calcul sont les suivants :

Frequency range (Hz)	Nb of computed frequencies	Nb of mesh nodes	Nb of elts per wavelength	Precision required	Time	Precision obtained
20-5000	70	578	6	1 ^e -4	39 min	1e-4
5000-10000	20	2306	6	1 ^e -4	165 min	1e-4
10000-13000	15	5186	6	1 ^e -4	16h30min	1e-1

Les colonnes *Precision required* et *Precision obtained* indiquent la précision sur la valeur du champ de pression au point de mesure (i.e. l'erreur de calcul demandée est inférieure à 10⁻⁴, et l'erreur de calcul obtenue est effectivement inférieure à 10⁻⁴).

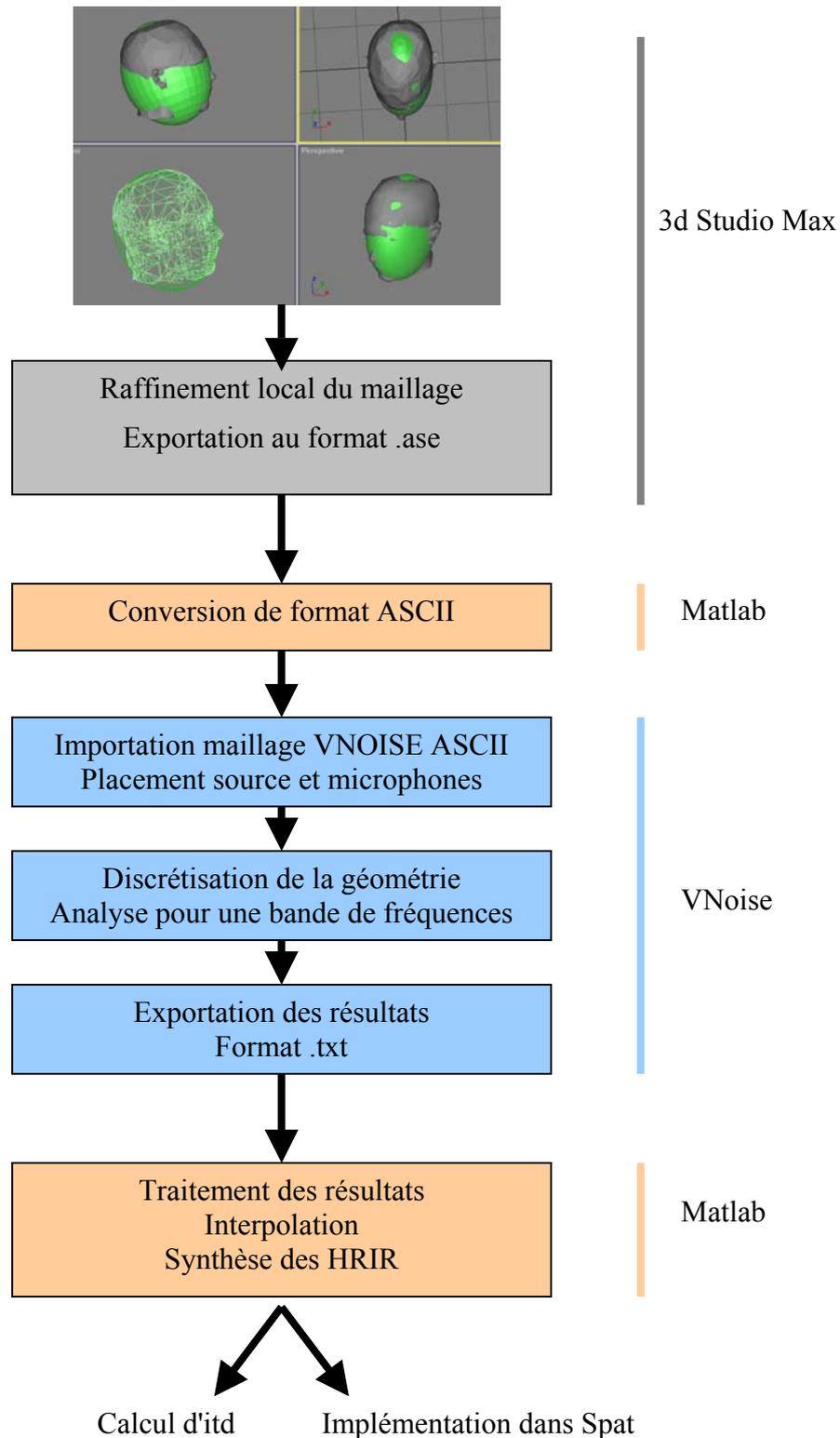
Les figures en annexe 3 montrent les résultats obtenus pour différentes incidences. Pour les fréquences inférieures à 9000 Hz, les résultats numériques sont extrêmement proches des résultats analytiques. On observe toutefois de petites différences d'environ 1 dB pour le côté opposé à la source (contralateral side), mais ceci s'explique très bien par la méthode numérique en elle-même : ce sont les nœuds les plus éloignés de la source qui auront une précision moindre sur le champ de pression résultant, car les approximations se propagent de proche en proche. Au-dessus de 9000 Hz, les différences d'amplitude s'expliquent par une mauvaise configuration des solveurs VNoise, mais ce problème a été résolu par la suite. Le choix de montrer ces erreurs est à but purement informatif. D'autre part, on remarquera que les réponses en phase théoriques et numériques sont identiques.

En conclusion, la diffraction d'une onde plane sur une sphère rigide est très bien simulée par VNoise. Ce résultat est intéressant en soi car le modèle sphérique présente un degré élevé de symétries et ceci est souvent très gênant pour les calculs par élément finis : les modes propres internes du modèle perturbent la simulation du champ externe. Pour éviter ces problèmes, on ajoute à l'intérieur de la sphère des nœuds supplémentaires pour casser les symétries et éviter les résonances. VNoise possède une fonction de collocation automatique de nœuds internes, et les résultats précédents montrent l'efficacité de l'algorithme. En comparaison, on se référera à [9] et [10] pour observer ces phénomènes de résonance interne.

STS Software a publié une partie de ces résultats sur son site web www.sts-soft.com pour montrer les performances de VNoise.

4. Protocole de calcul des HRTF pour un modèle simple

4.1. Description du protocole



4.2. Le maillage

Lorsqu'on utilise un logiciel d'éléments finis, la construction du maillage est une opération souvent fastidieuse et la plus coûteuse en heures de travail. La construction du maillage est liée à la géométrie initiale, et VNoise ne permet pas de créer des géométries à base d'ellipsoïdes. L'utilisation d'un logiciel de CAO est donc nécessaire pour notre étude. Malheureusement, le maillage en vue de l'utilisation de la méthode des éléments finis a ses exigences propres qui ne sont pas toujours compatibles avec les seules informations conservées par les logiciels de CAO. Le problème serait simplifié si un maillage plus ou moins uniforme était satisfaisant mais le recours au principe de réciprocité ne le permet pas.

Le premier essai consistait en l'importation sous VNoise d'une géométrie 3D Studio Max (contenant une centaine de nœuds définissant une ellipsoïde), et d'effectuer la discrétisation sous VNoise. Les résultats ont été très peu concluants car l'algorithme de maillage de VNoise n'est apparemment pas adapté à la définition géométrique d'objets 3D Studio Max.

La seule solution consiste alors à mailler "à la main" un objet sous 3 D Studio Max et de l'importer sous VNoise en tant que discrétisation prête au calcul. C'est une solution peu générique car ne fonctionnant que pour des modèles simples ou le contrôle du nombre de nœuds se fait facilement. Par chance, 3 D Studio Max permet de créer des objets tels que "sphère géodésique" dont le placement des nœuds (en triangle sur toute la surface de la sphère) correspond bien avec celui effectué par le mailleur VNoise. Il a fallu vérifier que le nombre d'éléments par longueur d'onde était bien respecté en mesurant la distance maximale inter-nœuds.

L'opération de maillage manuel apparaît néanmoins comme peu sûre et un protocole de validation a été mis en place pour s'assurer de la bonne marche des simulations (cf. paragraphe 4.5.)

4.3. Le calcul

4.3.1. Positions des points de mesure

France Télécom R&D a mené en 2001 une campagne de mesure de HRTF sur 965 directions spatiales pour 8 sujets du laboratoire [15], et dans le but de comparer les simulations aux mesures, nous avons choisi de placer les points de mesures aux mêmes 965 positions spatiales.

En utilisant le principe de réciprocité, les microphones de mesure sont donc placés sur une sphère de rayon 1,4 m (là où était placée la source sonore pendant la campagne de mesure) de la manière indiquée sur la figure suivante. Une calotte inférieure n'a pas été mesurée pour des raisons propres au dispositif de mesure, et n'a donc pas été simulée .

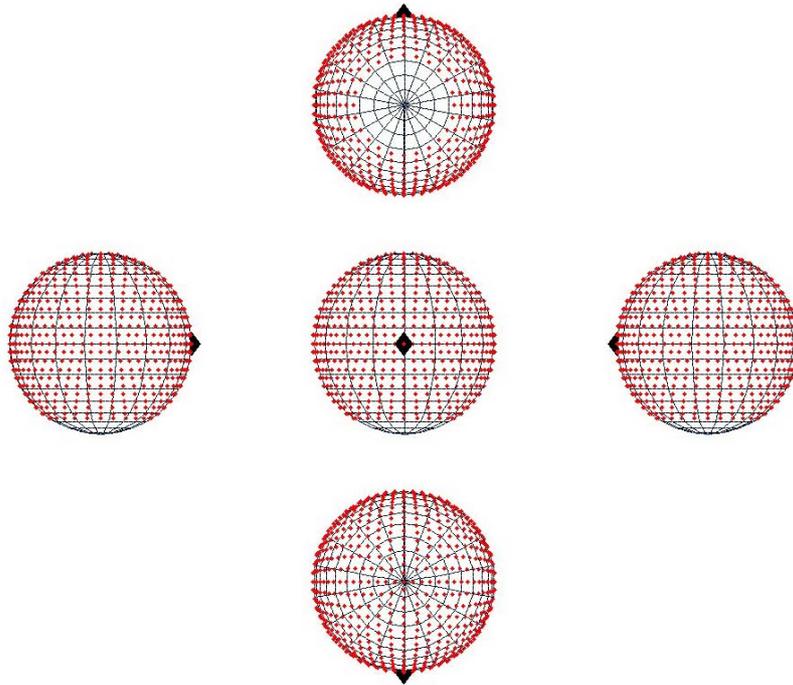


Figure 18 : De haut en bas et de gauche à droite, vues de dessous, droite, face, gauche et dessus de la sphère directionnelle. Les points rouges figurent les directions pour lesquelles les HRTFs des sujets ont été mesurées (965 directions de l'espace, pas d'élévation constant de 5.625° , 27 élévations mesurées). Le triangle noir symbolise le nez du sujet.

A remarquer que le nombre d'azimuts n'est pas constant par élévation.

Le fait d'avoir une résolution spatiale de 5° en azimut et en élévation permet de changer de mode de visualisation des mesures : par élévation, selon une courbe précise, par cônes de confusions, par interpolation des données mesurées.

4.3.2. Choix des fréquences calculées

Le choix des fréquences calculées est très important car le temps de calcul est proportionnel au nombre de fréquences, et plus les fréquences sont élevées, plus les maillages comportent un nombre de nœuds important, ce qui augmente le temps de calcul (qui évolue comme le carré du nombre de nœuds). Il est donc important d'optimiser l'échelle de fréquences choisie pour les simulations.

Si l'on se réfère aux précédents travaux de Katz et Kahana, les simulations utilisaient une échelle linéaire de fréquences allant de 0 à 8 kHz. Or il apparaît intéressant de pousser les calculs à une fréquence plus élevée pour observer les variations de spectre des HRTF qui sont importantes à hautes fréquences en raison de la diffraction plus importante de la tête et des pavillons. Pour conserver une bonne gestion du temps de calcul, augmenter la fréquence maximale implique une optimisation de la résolution fréquentielle.

Nous avons choisi d'utiliser une échelle logarithmique de fréquences, adaptée à la résolution fréquentielle du système auditif (se référer à [3]), afin de limiter le nombre de hautes fréquences qui demandent un temps de calcul plus long .

Détermination de la fréquence maximale

Afin de pouvoir comparer les résultats de simulations avec les HRTF mesurées, nous considérerons le même échantillonnage spatial. Le document [15] précise le protocole de mesure de HRTF : deux capsules microphoniques miniatures sont placées à l'entrée du conduit auditif à l'aide de bouchons en mousse.



Figure 19 : dispositif de mesure dit en "conduit bloqué"

La précision de placement du bouchon dans le conduit est de l'ordre de quelques millimètres. Une onde de longueur d'onde deux millimètres correspond à une fréquence de 170kHz donc la précision des mesures est bonne pour la plage des fréquences audibles. La fréquence maximale f_{\max} va être donc déterminée a priori, et sa valeur sera confirmée ou infirmée par des tests perceptifs avec HRTF dégradées (filtrage passe-bas de fréquence de coupure f_{\max}). Nous avons d'abord fixé $f_{\max} = 13$ kHz, pour des raisons de puissance de calcul : notre station de travail permet des calculs sur RAM pour des maillages comportant jusqu'à 10 000 nœuds, ce qui avec le critère de 6 éléments par longueur d'onde nous limite à une discrétisation donnant des résultats précis jusqu'à 13 kHz. Des tests perceptifs informels, basé sur une comparaison A/B entre deux sons, l'un convolué par les HRTF mesurées, l'autre convolué par les HRTF filtrées à 13 kHz, ont montré qu'il y avait peu de différence en termes de localisation. Nous avons donc conservé cette valeur de f_{\max} , quitte à refaire un calcul pour des fréquences supérieures avec une discrétisation plus fine si besoin est.

Détermination de la résolution fréquentielle

Le problème est de déterminer un espacement logarithmique entre 20 Hz et 13000 Hz qui soit suffisant pour, après interpolation, synthétiser des HRTF avec une précision adaptée aux vitesses de variation de l'amplitude du spectre d'une HRTF.

La première bande fréquentielle testée était une bande en tiers d'octave, c'est à dire contenant les N fréquences définies par :

$$f_0 = \text{cte}$$
$$f_i = f_0 \cdot 2^{i/3}$$

Pour savoir si une resynthèse du spectre à partir des N valeurs aux fréquences f_i est suffisamment fidèle à la réponse en fréquence originelle, ces N valeurs sont prises sur une

HRTF mesurée puis interpolées par méthode spline. La comparaison des deux réponses en amplitude est présentée sur la figure suivante :

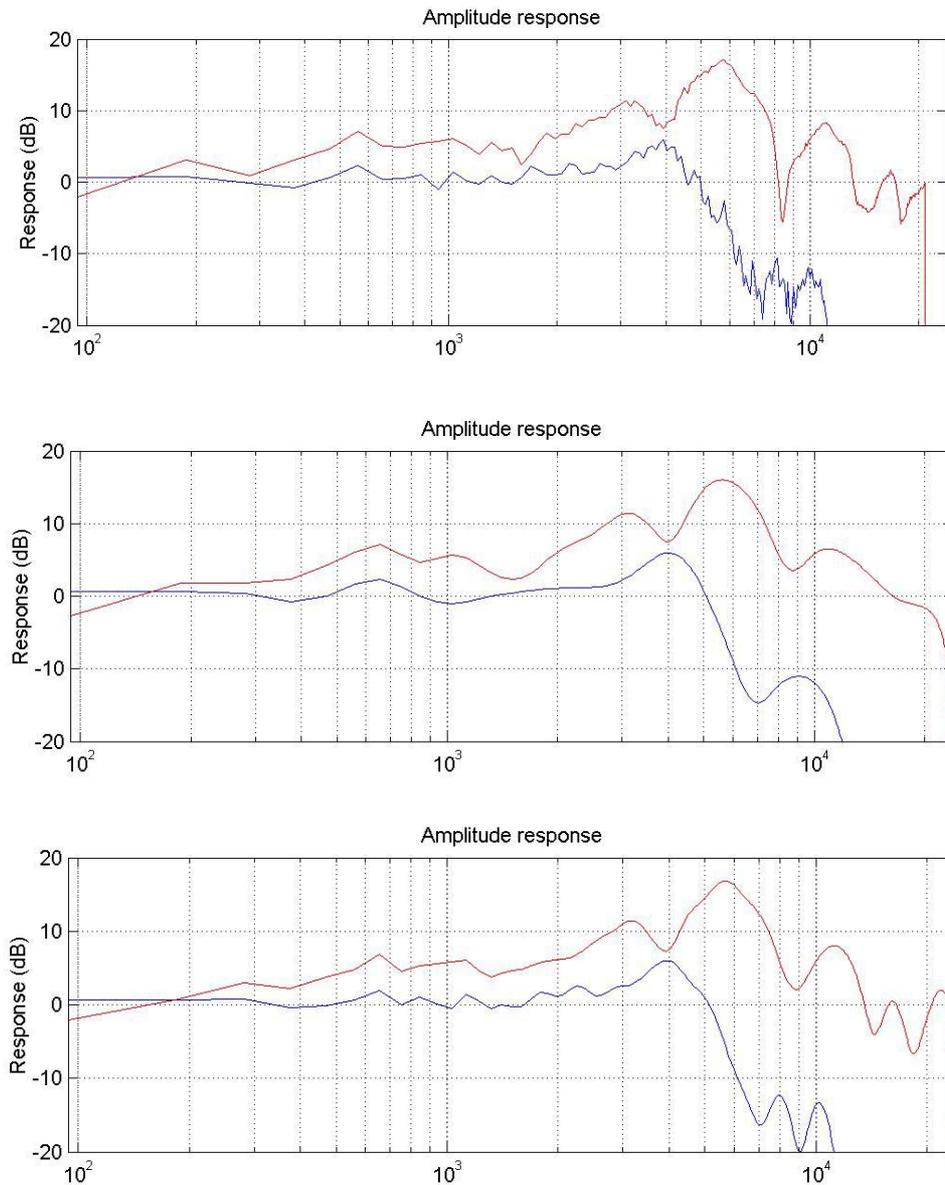


Figure 20 : comparaison entre une paire de HRTF non dégradées (haut), HRTF interpolées à partir de fréquences en 1/3 d'octave (milieu) et HRTF interpolées à partir de fréquences en 1/6 d'octave (bas)

Les sauts d'amplitudes sont trop peu marqués avec seulement N points fréquentiels à la base de l'interpolation. Un deuxième essai consiste à prendre $2N$ points (fréquences en 1/6 d'octave). La figure 19 montre que les courbes obtenues avec cette nouvelle échelle de fréquences sont nettement plus proches des courbes originelles. A remarquer que l'allure de la réponse en basse fréquence ($f < 100$ Hz) n'a pas à être prise en compte puisque les spectres des mesures commencent à $f = 100$ Hz.

Par bandes de 1/6 d'octave entre 20 et 13000 Hz, on obtient finalement un ensemble de 57 fréquences détaillées dans la figure suivante, qui sera utilisé pour les simulations.

20
22
25
28
31
35
39
44
50
56
63
70
79
88
99
11
125
140
157
177
198
223
250
281
315
354
397
445
500
561
630
707
794
891
1000
1122
1260
1414
1587
1782
2000
2245
2520
2828
3175

3564
4000
4490
5040
5657
6350
7127
8000

8980
10079
11314
12699

45 fréquences

**Figure 21 : choix des fréquences calculées (bande de 1/6 d'octave)
les fréquences sont séparées en trois groupes correspondant à trois discrétisations
distinctes.**

8 fréquences

4 fréquences

4.3.3. Fonctionnement du solveur

Le solveur de VNoise est configurable de différentes manières : si le but des algorithmes de calcul est toujours de résoudre le système linéaire $A.x = b$ où A est une matrice $n \times n$ et b un vecteur de longueur n , VNoise propose différentes méthodes de calcul.

La première est un algorithme simple qui réduit pas à pas le résidu $r_k = A-b.x_k$ mais ne force pas le système à converger. On peut donc choisir entre plusieurs autres algorithmes d'une famille appelée BICGSTAB (L) pour "BI-Conjugate Gradient " STABILisée [17]. Ce sont des algorithmes à faible coût de calcul, et à convergence assez régulière : L représente le nombre d'itérations d'une étape interne de l'algorithme, et à chaque incrémentation de ce pas interne la convergence augmente.

VNoise propose d'utiliser les solveurs BICGSTAB (2) à BICGSTAB (32), pour une variété de problèmes acoustiques. Malheureusement, il n'existe pas de règle générale pour le choix du bon algorithme, et seule l'expérience permet de choisir avec pertinence l'algorithme approprié au problème.

Cette étape d'essai erreur est indispensable car une mauvaise configuration du solveur peut faire véritablement exploser les prévisions de temps de calcul.

4.4. Le traitement des résultats

Le résultat d'un calcul est la donnée d'un triplet [fréquence, $\text{Re}(p)$, $\text{Im}(p)$] pour chaque nœud du maillage et chaque microphone de mesure placé dans la discrétisation. VNoise propose un export de ces données au format texte et tous les traitements ultérieurs sont effectués sous Matlab.

Le premier traitement consiste à normaliser les valeurs de pression par la pression en champ libre.

Ensuite, des routines de visualisation de HRTF et de calcul d'ITD ont été développées, pour permettre de comparer les mesures du TNO de Soesterberg et les simulations. Puis une routine d'implémentation dans le Spatialisateur (logiciel de spatialisation en temps réel développé par France Télécom R&D, avec module de rendu binaural) a été développée pour utiliser les simulations en tests perceptifs.

4.5. Validation du protocole

Pour s'assurer du bon fonctionnement de l'ensemble formé par les différentes étapes (qui somme toute sont assez disparates), nous avons voulu comparer nos résultats avec les résultats publiés par Yuvi Kahana sur le site web de l'ISVR à Southampton : il s'agit d'un calcul de HRTF sur une ellipsoïde représentant une tête "moyenne". Se référer à l'adresse

<http://www.isvr.soton.ac.uk/FDAG/vap/ellipsoid/info.htm>

pour plus de détails.

Nous avons repris les paramètres utilisés par Y. Kahana pour valider notre protocole :

- $r_x = 9,6$ cm, $r_y = 7.9$ cm, $r_z = 11.6$ cm
- Oreille droite à 0,86 cm de l'origine dans l'hémisphère inférieur. (i.e. position de la source = 0 0 -0.86), avec un azimuth égal à $100,22^\circ$
- Microphones disposés par pas de 5° sur une demi-sphère de rayon 1m, ce qui correspond à la mesure de 19 azimuths (de 0 à 90°) et 72 élévations (de 0 à 355°)
- 30 fréquences (de 100 à 3000 Hz par pas de 100Hz)

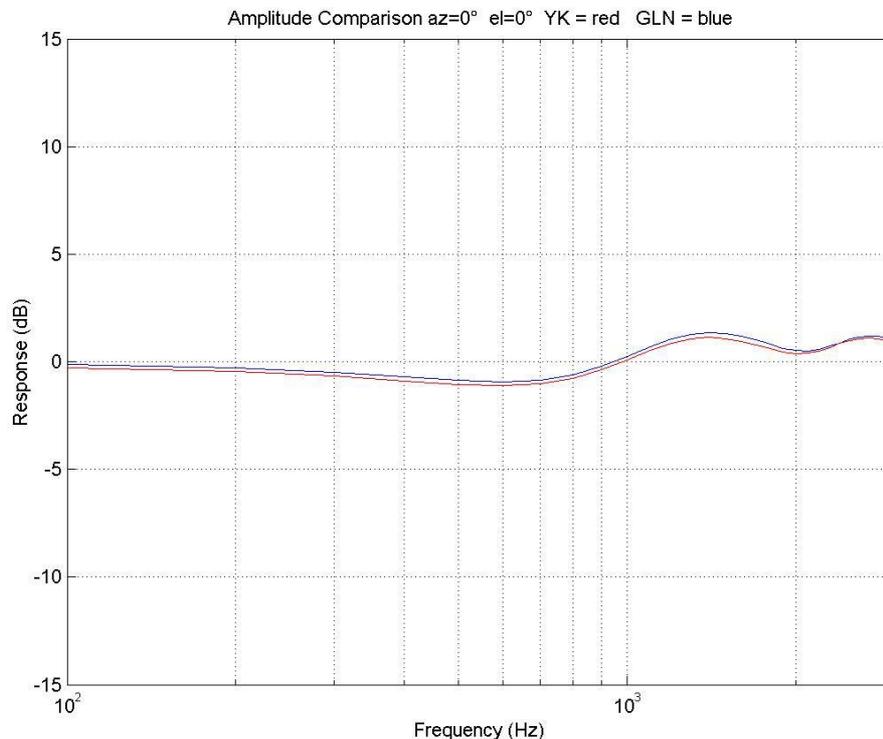
Les données concernant le maillage sont les suivantes :

- Base : sphère géodésique à 12 segments (1442 nœuds)
- Raffinement local du maillage (oreille droite) : 798 nœuds supplémentaires
- Nombre total de nœud : 2240.

La distance inter nœuds a été mesurée et est bien inférieure à 1/6 de la longueur d'onde d'une onde acoustique de fréquence 5000 Hz se propageant dans le vide. Le nombre d'éléments par longueur d'onde est donc bien suffisant pour la bande de fréquences calculée.

Nous avons calculé l'écart relatif entre les résultats de Kahana et les nôtres (différence des normes du champ de pression). Il est de 0,26 %, ce qui est extrêmement faible.

Les quelques courbes qui suivent attestent de la similarité des résultats :



**Figure 22 : comparaison des modules des HRTF des ellipsoïdes
YK = résultats de Yuvi Kahana
GLN = résultats obtenus à France Télécom R&D**

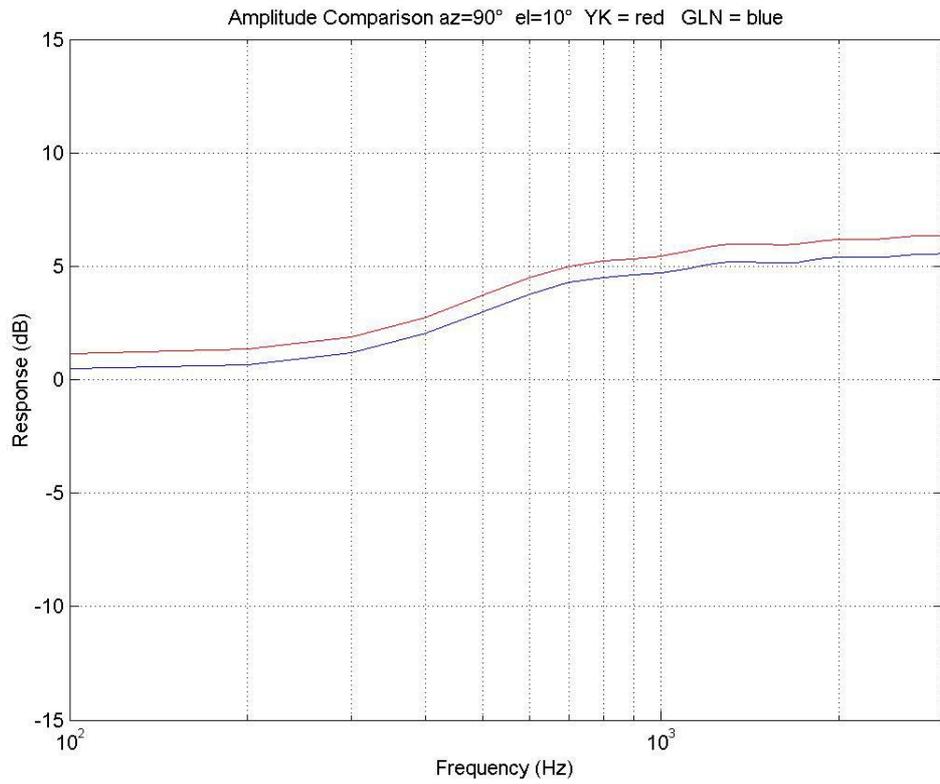
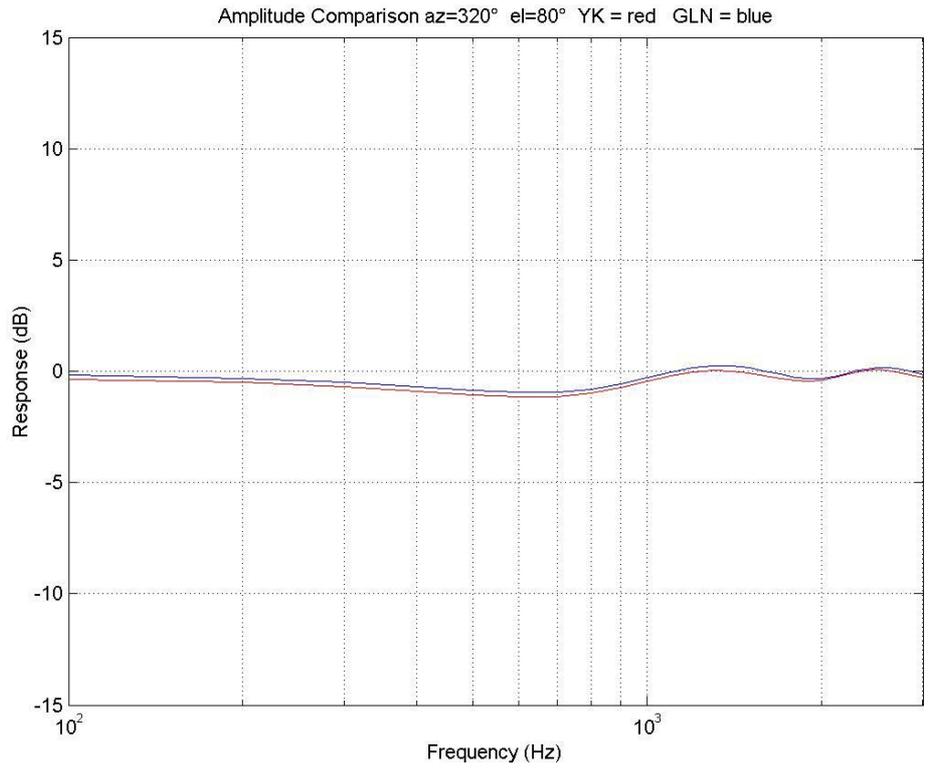


Figure 23 : comparaison des modules des HRTF des ellipsoïdes