UNIVERSITÉ JOSEPH FOURIER-GRENOBLE 1
SCIENCES ET GÉOGRAPHIE
U.F.R. D'INFORMATIQUE ET DE MATHÉMATIQUES APPLIQUÉES

THÈSE

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ JOSEPH FOURIER

Discipline : Informatique

arrêtés ministériels du 5 juillet 1984 et du 30 mars 1992

Présentée et soutenue publiquement
par

Nicolas TSINGOS

Le 4 Décembre 1998

SIMULATION DE CHAMPS SONORES DE HAUTE QUALITÉ POUR DES APPLICATIONS GRAPHIQUES INTERACTIVES

effectuée sous la direction de

Claude PUECH et Jean-Dominique GASCUEL

au sein du laboratoire *i*MAGIS/GRAVIR-IMAG/INRIA. UMR CNRS C5527.

COMPOSITION DU JURY :

Jean-Pierre	VERJUS	Président
Julie	DORSEY	Rapporteur
Murray	HODGSON	Rapporteur
Jean-Claude	PAUL	Rapporteur
Claude	PUECH	Directeur de thèse
Jean-Dominique	GASCUEL	Responsable de thèse

Sweet Echo, sweetest Nymph, that liv’st unseen
Within thy aery shell
By slow Meander’s margent green,
And in the violet imbroider’d vale
Where the love-lorn Nightingale
Nightly to thee her sad Song mourneth well.
Canst thou not tell me of a gentle Pair
That likest thy Narcissus are ?
O, if thou have
Hid them in some flowry Cave,
Tell me but where
Sweet Queen of Parly, Daughter of the Sphear,
So maist thou be translated to the skies,
And give resounding grace to all Heav’ns Harmonies.

John Milton. *Comus,* 1633.

A mon père.

Remerciements

Je tiens tout d'abord à exprimer ma profonde reconnaissance à Claude Puech pour m'avoir accueilli au sein de l'équipe *i*MAGIS pendant cinq ans et m'avoir fait confiance pour explorer pendant trois ans de thèse un sujet qui me tenait à cœur. Il a largement contribué, directement ou indirectement, à ce que ces années se déroulent à la perfection.

Je remercie dans un deuxième temps mes rapporteurs, Julie Dorsey, Jean-Claude Paul et Murray Hodgson pour avoir pris le temps de relire mon manuscript et y apporter nombre de critiques constructives. Je remercie également Jean-Pierre Verjus pour avoir accepté de présider mon jury.

Jean-Dominique Gascuel, m'a encadré pendant ces trois ans de thèse en m'apportant tout le support, tant technique que théorique, dont je pouvais avoir besoin. Je le remercie également d'avoir su me laisser conduire mes travaux de manière personnelle, me permettant d'acquérir ma propre expérience du monde de la recherche.

Je souhaite témoigner également ici ma reconnaissance aux personnes des départements «Acoustique des salles» et «Ondes acoustiques et électromagnétiques» du Centre Scientifique et Technique du Bâtiment à Grenoble pour leur aide, le temps et les moyens qu'ils ont bien voulu me consacrer. Je remercie en particulier Jacques Martin, Yannick Gabillet, et Jean-Paul Vian.

J'adresse de sincères remerciements à Murray Hodgson, qui en plus d'être rapporteur de ce travail, a bien voulu m'aider à conduire toute une série de tests comparatifs visant à valider mon approche. Son aide et son expérience m'ont été précieuses.

Au cours de ces cinq années passées dans l'équipe *i*MAGIS, j'ai apprécié la compagnie et l'expérience des autres chercheurs, des thésards et des stagiaires. Certains d'entre eux, comme d'autres qui ont à présent quitté l'équipe, vont rester pour moi plus que de simples compagnons de route, des amis... Je ne les saluerai pas tous nomément ici, ils savent qui ils sont. J'adresse néanmoins un immense merci à Frédo pour l'aide qu'il m'a apportée avant (et pendant !) ma soutenance.

La compréhension et les encouragements constants de ma famille m'ont également soutenu et rassuré durant toutes ces années. Enfin, dans la frivolité comme dans les instants de tension (il y en a toujours durant une thèse), une personne a constamment été là pour moi, pour m'écouter. Son mérite est immense et pour cause, c'est mon épouse...Agata.

Table des matières

Introduction

L'INFORMATIQUE possède cet aspect alchimiste qui permet de pouvoir présenter de manière appréhendable par un utilisateur des informations n'existant dans la machine que sous une forme de suite binaire de «bits». Pour cela, il évidemment très naturel de faire appel à nos différents sens et en particulier la vue, au travers des moniteurs informatiques trônant sur les bureaux.

Dans la recherche d'une représentation de plus en plus naturelle et avec l'augmentation de la puissance de calcul des machines, la visualisation des données informatiques est rapidement passée de bidimensionnelle à tridimensionnelle, de simpliste à hyper-«réaliste», d'une projection sur un écran plat à une immersion visuelle complète au travers de nouveaux systèmes de restitution de l'image (lunettes 3D, «CAVEs», *etc.*). L'expérience ainsi obtenue a été baptisée, un peu abusivement peut être, *réalité virtuelle*. Cet aspect ne s'applique bien évidemment pas seulement à l'ordinateur outil de calcul mais également à l'ordinateur outil de création artistique ou *medium* des loisirs interactifs. On peut très facilement s'en convaincre en allant voir au cinéma les effets spéciaux hollywoodiens dernier cri ou jouer aux jeux vidéo du moment.

Touchant notre sens principal, c'est d'abord la visualisation qui a bénéficié d'un développement privilégié. Mais qui dit réalisme et naturel, dit également multi-sensoriel. En effet, dans notre expérience quotidienne, nous faisons largement appel à nos autres sens : le toucher et l'ouïe en particulier. Ces dernières années, les interfaces *haptiques* ont, elles aussi, fait l'objet d'un intense développement permettant à un utilisateur de percevoir un contact physique avec des objets synthétiques.

Dès lors, il était tout aussi naturel d'améliorer l'aspect naturel et immersif de la représentation des données auditives. Un peu en retard sur l'«optique virtuelle» et ses images de synthèse, sa science sœur l'«acoustique virtuelle» arrive maintenant en force dans de nombreuses applications informatiques mais également pour l'écoute musicale grand-public avec le fameux son «3D» ou «surround». Ces techniques de spatialisation du son permettent aujourd'hui de simuler des sources sonores virtuelles placées arbitrairement dans l'espace autour de l'auditeur, pour un résultat final souvent impressionnant. Elles permettent dans la plupart des cas une immersion décuplée dans l'environnement virtuel tout en restant peu intrusives pour l'utilisateur.

Néanmoins, s'il est essentiel d'être capable de restituer une information sonore spatiale, il est également nécessaire de simuler de manière réaliste les phénomènes touchant la propagation du son. En effet, il est peu probable de pouvoir réaliser des enregistrements utilisables dans toute une variété de situations. On peut songer à réaliser des bandes-son pour des films de synthèse (où l'on n'a pas de possibilité de prise de son), ou sonoriser des applications interactives où l'environnement est susceptible d'être modifié à tout instant. Ce problème de la simulation réaliste des phénomènes de propagation est un des thèmes clé de la recherche en acoustique. L'une des applications principales en est l'acoustique prévisionnelle, qui permet d'évaluer la qualité d'un lieu d'écoute avant sa construction ou l'impact de telle ou telle construction sur l'environnement sonore.

Lorsque l'on s'intéresse de plus près aux différentes approches utilisées pour la simulation acoustique, on constate que la plupart d'entre-elles sont fondées sur des approximations géométriques dérivées de l'optique. Elles sont en cela très similaires aux techniques de simulation de l'éclairage en synthèse d'image. Il ne faut pas pour autant imaginer que les problèmes liés au son et à la lumière peuvent se traiter de manière unifiée. Si les phénomènes physiques peuvent être modélisés par un même formalisme ondulatoire, de nombreux facteurs ne permettent pas de transposer directement une approche d'un domaine à l'autre ; le principal étant sans nul doute la différence de vitesse de propagation des deux ondes : l'aspect temporel de la propagation sera négligé en optique, il est vital en acoustique. De plus, comme le disait Charles Garnier, si l'acoustique était, « [...] comme l'optique est, elle, une science parfaite, en ce sens que l'on sait fort bien si de tel point on verra ou on ne verra pas [...] », on pourrait utiliser directement les mêmes approches géométriques. Mais d'autres phénomènes, comme la diffraction des ondes sonore par les obstacles, contribuent à rendre les simulations acoustiques plus complexes.

Cependant, dans le cadre d'approximations géométriques, on doit pouvoir profiter du développement des techniques de rendu visuel. La plupart des machines tendent aujourd'hui à être équipées de systèmes d'affichage de données tridimensionnelles câblés. Ces cartes graphiques permettent de réaliser en temps-réel des opérations géométriques complexes. Il semble donc intéressant de mettre à profit cette puissance pour accélérer des calculs géométriques liés à une simulation sonore destinée à enrichir une expérience virtuelle interactive. De plus, ces dernières années, sont apparues des approches de simulation en synthèse d'images, permettant de simuler l'éclairage dans des environnements complexes. Les approches dites «multi-résolution», en particulier, effectuent les calculs à différents niveaux de résolution en fonction de critères de qualité ou de contraintes sur le temps de calcul. Elles permettent ainsi d'obtenir des résultats très réalistes dans des temps de calculs très intéressants. L'adaptation d'une telle approche à la simulation acoustique pourrait permettre d'en accroître les performances.

Dans cette thèse, nous nous proposons d'explorer ces deux aspects.

Nous commencerons par rappeler, dans une première partie, quelques bases d'acoustique nécessaires à la bonne compréhension du présent document. Nous examinerons ensuite les différentes approches utilisées pour la simulation en acoustique virtuelle et l'intégration du son dans des applications graphiques interactives.

Nous verrons, au Chapitre 6, comment il est possible de modéliser de manière qualitative les phénomènes liés à l'occultation des ondes sonores par des obstacles, qui sont absents de toutes les simulations acoustiques interactives actuelles. Nous proposerons une technique qui permet de réaliser les calculs nécessaires en temps-réel en mettant à profit les capacités des cartes de rendu graphique câblé. Nous verrons également comment cette approche peut être étendue pour donner des résultats plus quantitatifs.

Nous présenterons au Chapitre 7 un système permettant la modélisation et le rendu de scènes virtuelles auditives de manière cohérente avec une simulation visuelle associée. En particulier, nous nous intéresserons ici à des environnements dynamiques.

Enfin, dans une troisième partie, nous définirons une nouvelle approche géométrique, fondée sur les techniques de radiosité hiérarchique utilisées en synthèse d'image et permettant de simuler des phénomènes de propagation complexes de manière efficace. Cette approche permet de gérer des échanges d'énergie au cours du temps à différents niveaux de détail. Nous proposons également une extension permettant d'en réduire la complexité, tout en prenant en compte des phénomènes d'interférences propres à la théorie ondulatoire. Nous présenterons une application à l'évaluation de la qualité acoustique d'un lieu d'écoute, avant de conclure ce document.

Première partie

Acoustique physique et perceptive

Chapitre 1

Acoustique physique : le son objectif

Directly or indirectly, all questions connected with [the sensation of sound] must come for decision to the ear, as the organ of hearing; and from it there can be no appeal. But we are not therefore to infer that all acoustical investigations are conducted with the unassisted ear. When once we have discovered the physical phenomena which constitute the foundation of sound, our explorations are in great measure transferred to another field lying within the dominion of the principles of Mechanics. Important laws are in this way arrived at, to which the sensations of the ear cannot but conform.

Lord Rayleigh, *The theory of sound,* 1877.

L'ACOUSTIQUE peut être définie comme la science du son, incluant sa production, transmission et ses effets. Elle n'est donc pas limitée au seul phénomène responsable de la sensation d'audition. Elle se distingue de l'optique par le caractère mécanique plutôt qu'électromagnétique des ondes sonores. Dans ce chapitre, nous présentons quelques bases d'acoustique physique et de traitement du signal ; le but ici n'étant pas de dresser un compte rendu exhaustif des connaissances en acoustique mais d'introduire de la manière la plus accessible possible des notions et termes essentiels à la bonne compréhension du manuscrit, des problèmes et des choix ultérieurs. Nous introduirons tout d'abord des notions générales concernant la propagation des ondes sonores, les interférences et la différence entre régime continu et impulsionnel, la modélisation des sources et des récepteurs sonores. Nous définirons également la notion de réponse impulsionnelle d'un canal acoustique, essentielle pour la suite du document. Puis nous présenterons différents phénomènes de propagation (réflexion, diffraction, effet Doppler, etc.). Enfin nous aborderons quelques bases concernant l'étude de l'acoustique des environnements réverbérants, en particulier les espaces clos.

1.1 Caractérisation du phénomène sonore. Équation d'onde

D'un point de vue physique, le son peut être modélisé comme un mouvement ondulatoire du milieu de propagation. Il est intéressant de noter ici que la nature ondulatoire du phénomène sonore avait été envisagée dès le III^{e} siècle avant J.C. par le philosophe grec Chrysippe. Dans un milieu élastique, le passage d'une onde sonore entraîne une oscillation des molécules autour de leur position d'équilibre, due à leur inertie et aux forces élastiques présentes. Ces oscillations entraînent une série de compressions et raréfactions successives des molécules dans le milieu, provoquant une augmentation ou une diminution locale de la pression du milieu par rapport à la pression au repos (*cf.* Figure 1.1). Il faut noter ici qu'il n'y a pas de réel déplacement du milieu mais juste une fluctuation due au passage de la perturbation provoquée par la source de l'onde sonore. La présence d'un milieu de propagation est donc nécessaire à la propagation d'une onde sonore dite *mécanique*, contrairement à d'autres ondes comme la lumière ou les ondes radio qui se propagent dans le vide (ondes électromagnétiques). Ces déplacements des molécules du milieu dépendent du temps et de la position dans l'espace et suffisent à décrire complètement la propagation de l'onde sonore.

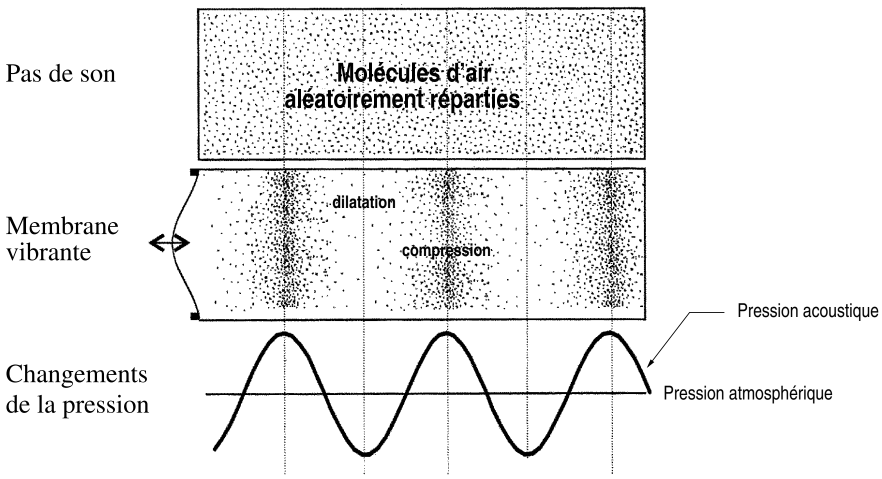


Figure 1.1 - *Le son peut être modélisé comme un mouvement ondulatoire du milieu de propagation dû à une perturbation de la pression atmosphérique (engendrée par la vibration d'une membrane de haut parleur par exemple). La pression acoustique est la différence entre la pression instantanée en un point du milieu et sa pression au repos*[1].

La *pression acoustique* (exprimée en Newtons par unité de surface[2]) est donc définie comme la différence entre la pression du milieu au repos et sa pression instantanée. C'est à cette grandeur que l'on s'intéressera particulièrement car c'est à elle que sont sensibles la plupart des récepteurs et en particulier l'oreille. D'une manière générale, on appellera *champ sonore*, la fonction spatio-temporelle correspondant aux variations de la pression acoustique dans l'environnement.

La dynamique des pressions auxquelles on s'intéresse en acoustique est extrêmement importante (*cf.* Chapitre 2). Il est donc courant d'exprimer les pressions acoustiques en utilisant un gain logarithmique, exprimé en *décibels* (dB) :

$$P \ \ (\text{dB } SPL) = 20\log_{10}\left(\frac{\tilde{p}}{\tilde{p}_0}\right),$$

où $\tilde{p}$ et $\tilde{p}_0$ sont exprimée en $N.m^{-2}$ et $\tilde{p}_0$ est une pression de référence (*cf.* Annexe A.1)[3].

La vitesse de déplacement d'une onde sonore, que nous noterons c dans le reste de ce document, dépend des paramètres physiques du milieu (température, etc.). Dans l'air, à 20^o C, la vitesse d'une onde

1. figure reproduite d'après [Hol97].
2. $N.m^{-2}$. On utilise aussi de manière indifférente les Pascals (Pa)
3. ici, $\tilde{p}$ et $\tilde{p}_0$ sont des *pressions quadratiques moyennes* i.e. la racine carrée de la moyenne temporelle du carré de la pression.

Table 1.1 - Propriétés comparées de trois différentes classes d'ondes

propriétés	**son**	**lumière**	**radio**
type	mécanique	électromagnétique	électromagnétique
célérité	340 m.s^{-1} (air) 1000 m.s^{-1} (eau) 1000 m.s^{-1} (métal)	3.10^{8} m.s^{-1} (vide) - -	3.10^{8} m.s^{-1} (vide) - -
fréquences	de 0 Hz (infrasons) à 400 kHz (ultrasons)	de 10^{11} Hz (infrarouge) à 10^{17} Hz (ultraviolet)	de 3 kHz (Very Low Frequency) à 300 GHz (Extremely High Frequency)
longueurs d'onde	de 100 m à 1 mm	de 10^{-3} m à 10^{-7} m	de 10^{5} m à 10^{-3} m
seuils perception	de 20 Hz à 20 kHz	de 7.8×10^{-7} m (rouge) à 3.9×10^{-7} m (violet)	-

sonore est d'environ 340 m.s^{-1} (cf. Annexe A.2.1). La Table 1.1 donne quelques autres éléments de comparaison quantitatifs entre ondes mécaniques et électromagnétiques.

1.1.1 Équation d'onde et grandeurs physiques pertinentes

Si l'on se place dans les hypothèses de l'acoustique linéaire (i.e. tout les phénomènes de propagations sont linéaires) et que l'on considère que le milieu de propagation est l'air (assimilé à un fluide parfait), les variables définissant l'état acoustique du fluide en un point M au temps t sont alors la pression $p(M,t)$ et la vitesse vibratoire particulaire $\mathbf{v}(M,t)$. On peut alors montrer [Kut91b] que p (de même que les composantes de $\mathbf{v}$) vérifie l'équation différentielle suivante ou *équation d'onde*[4] :

$$\Delta p - \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} = Q_0, \tag{1.1}$$

où c est la célérité du son dans l'air et $Q_0(M,t)$ est la *densité volumique de débit* crée par les sources présentes ($\Delta x = (\vec{\nabla} x)\cdot(\vec{\nabla} x)$ dénote le laplacien de la variable x).

D'autres grandeurs que la pression sont également utilisées couramment pour décrire une onde sonore. Il s'agit de l'*intensité acoustique* et de la *puissance acoustique*.

L'intensité acoustique[5] I est la moyenne dans le temps de la *puissance surfacique instantanée*, c'est-à-dire pour un point de l'espace, l'énergie acoustique transitant dans une direction donnée $\mathbf{n}$, par unité de temps, par unité de surface normale à la direction de propagation (*cf.* Figure 1.2) :

$$I_{\mathbf{n}} = \overline{p(\mathbf{v}\cdot\mathbf{n})} = \overline{pv}\cos\theta, \tag{1.2}$$

où $\mathbf{v}$ est la vitesse particulaire orientée suivant la direction de propagation. On peut montrer également, qu'en champ lointain (*i.e.* pour une distance à la source grande devant la longueur d'onde), l'intensité acoustique varie comme le carré de la pression acoustique moyenne (*cf.* Annexe A.1).

4. appelée aussi équation de D'Alembert.
5. exprimée en Watts par unité de surface (W.m^{-2}).

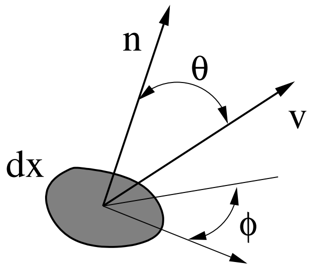


Figure 1.2 - *Notations pour la définition de l'intensité acoustique.*

On pourra donc définir l'énergie E transitant au point x à travers l'élément différentiel de surface dx pendant l'intervalle de temps dt par (*cf.* Figure 1.2) :

$$E = I(x,t) \underbrace{\cos\theta dx}_{\text{aire projetée}} dt,$$

où $I(x,t)$ est l'intensité dans la direction de propagation de l'onde.

La puissance[6], énergie par unité de temps, sera donc définie par :

$$P = I(x,t)\cos\theta\, dx.$$

On se référera à l'Annexe A.1 pour plus de détails concernant les relations entre pression acoustique, puissance et intensité.

1.1.2 Ondes planes, ondes sphériques

Considérons l'équation d'onde homogène associée à l'équation 1.1 :

$$\Delta p - \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} = 0. \tag{1.3}$$

Ondes planes

Si l'on suppose que la pression acoustique ne dépend que du temps et d'une direction de propagation **x**, on obtient l'équation régissant la propagation des ondes planes :

$$\frac{\partial^2 p}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} = 0, \tag{1.4}$$

Cette équation admet des solutions générales de la forme :

$$p(x,t) = F(t - x/c) + G(t + x/c).$$

où F représente une onde progressive et G une onde régressive. Une solution particulière à cette équation est obtenue en choisissant pour la pression, une dépendance temporelle de la forme $e^{i\omega t}$ (i est le nombre complexe tel que $i^2 = -1$). On obtient alors comme solution de l'équation d'onde des vibrations

6. exprimée en watts (W).

dites *harmoniques* de la forme (on ne garde que la partie réelle (resp. imaginaire) qui seule a un sens physique) :

$$p(x,t) = P\cos(\omega t - kx),$$

où P est l'amplitude de la pression acoustique.

Note : Nous avons introduit ici la *pulsation* ω et le *nombre d'onde* $k = \omega/c$ tel que la *période* temporelle de la vibration soit $2\pi/\omega$. Deux valeurs identiques de la vibration sont donc séparées par une distance $\lambda = 2\pi/k$ appelée *longueur d'onde* de la vibration[7]. On a donc $\lambda = \frac{2\pi c}{\omega} = \frac{c}{f}$, où $f = \omega/2\pi$ est la *fréquence* de la vibration[8].

Ondes sphériques

Si l'on suppose à présent que la pression acoustique ne dépend que de la distance r à un point d'origine (cas d'une source ponctuelle), on peut écrire l'équation d'onde en coordonnées sphériques et obtenir l'équation de propagation caractérisant une onde sphérique :

$$\frac{\partial^2 p}{\partial r^2} + \frac{2}{r}\frac{\partial p}{\partial r} - \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} = 0. \tag{1.5}$$

On obtient alors comme solution des vibrations de la forme :

$$p(r,t) = \frac{P}{r}\cos(\omega t - kr), \tag{1.6}$$

Cette dernière formulation est particulièrement importante puisque, comme nous le verrons, nous assimilerons toutes les sources sonores à des sources ponctuelles, générant des ondes sphériques. On constate que dans ce cas l'amplitude de la pression décroît de manière inversement proportionnelle à la distance à la source.

1.2 Ondes harmoniques et inharmoniques

Jusqu'à présent, nous n'avons évoqué que le cas des ondes harmoniques dont la dépendance temporelle s'exprime comme une sinusoïde et qui ne sont qu'une solution possible de l'équation d'onde. Nous allons maintenant examiner les relations entre les ondes harmoniques et les ondes plus complexes que l'on rencontre dans la réalité.

1.2.1 Phase et interférences

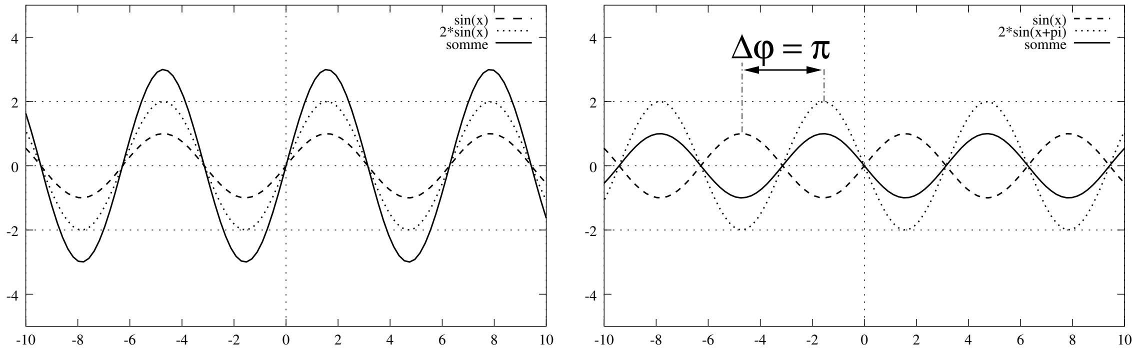
Considérons une onde harmonique plane (sans perte de généralité) $s(x,t)$. On appellera *phase* φ de l'onde le terme $\varphi = kx$ tel que $s(x,t) = S\cos(\omega t - \varphi)$, où S est l'amplitude de l'onde.

Deux ondes de même fréquence additionnées donnerons alors une onde de même fréquence mais d'amplitude différente de celle des deux ondes d'origine. Cette différence d'amplitude est liée au phénomène d'*interférence* entre les deux ondes s et s' du à leur différence de phase relative $\Delta\varphi = (\varphi - \varphi') \bmod 2\pi$ (*cf.* Figure 1.3). Dans le cas général, si $0 \leq |\Delta\varphi| \leq \pi/2$ alors l'amplitude du signal somme est

7. dans tout le reste du document λ désignera la longueur d'onde de l'onde sonore et les quantités en dépendant seront dénotés par un λ en exposant.

8. exprimée en hertz (Hz) de dimension s^{-1}.

strictement supérieure à la somme de celles des signaux d'origine ; on parle d'*interférence constructive*. Si $|\Delta\varphi| > \pi/2$, l'amplitude du signal somme est strictement inférieure à la somme de celles des signaux d'origine et on parle alors d'*interférence destructive*.



(a) (b)

Figure 1.3 - *Interférence lors de la somme de deux signaux harmoniques de même fréquence. (a) les deux signaux sont* en phase *(*$\Delta\varphi = 0$*), l'amplitude du signal somme est la somme des amplitudes des deux signaux d'origine. (b) les deux signaux sont* en opposition de phase *(*$\Delta\varphi = \pi$*) et l'amplitude du signal résultant est la différence des deux amplitudes d'origine.*

Dans le cas où plusieurs ondes de fréquences et phases relatives différentes sont additionnées, on obtient une onde inharmonique complexe. C'est ce qui est utilisé dans certains synthétiseurs pour générer les sons à partir d'une table d'ondes de référence sous le terme de *synthèse additive*. Le processus inverse où l'on «enlève» certaines fréquences (par filtrage) à des sons complexes, au contenu fréquentiel riche, a également été utilisé (*synthèse soustractive*) [dP83].

1.2.2 Vibrations harmoniques et inharmoniques. Analyse de Fourier

Les ondes harmoniques ne sont évidemment pas représentatives de celles que l'on peut rencontrer en pratique. Toutefois, elles sont particulièrement importantes puisque l'on peut montrer que n'importe quel signal non harmonique peut s'exprimer comme une somme de signaux harmoniques. Cette décomposition en signaux harmoniques de fréquence f d'un signal $s(t)$ (supposé ici réel) est obtenue par *analyse de Fourier* :

$$s(t) = \int_{-\infty}^{+\infty} \hat{S}(f) e^{i2\pi f t} df,$$

où $\hat{S}(f)$ (à valeurs dans $\mathbb{C}$) est la *transformée de Fourier* de $s(t)$.

Note : On a de manière symétrique : $\hat{S}(f) = \int_{-\infty}^{+\infty} s(t) e^{-i2\pi f t} dt$, où $s(t)$ est la *transformée de Fourier inverse* de $\hat{S}(f)$.

L'intérêt d'une telle représentation vient de l'hypothèse de linéarité des phénomènes acoustiques. Dans ce cas, la solution de l'équation d'onde pour une onde inharmonique sera obtenue par sommation des solutions pour les ondes harmoniques résultantes de son analyse de Fourier.

À partir de la transformée de Fourier d'un signal, on définit également deux fonctions importantes :

– le *spectre* du signal : c'est la fonction $P(f) = |\hat{S}(f)e^{i2\pi ft}|$, qui décrit le contenu fréquentiel du signal[9].

– la *phase* du signal : c'est la fonction $\Phi(f) = \arctan\left(-\frac{\mathrm{Im}\left(\hat{S}(f)e^{i2\pi ft}\right)}{\mathrm{Re}\left(\hat{S}(f)e^{i2\pi ft}\right)}\right)$ ($\mathrm{Re}(\hat{x})$ et $\mathrm{Im}(\hat{x})$ désignent respectivement la partie réelle et imaginaire du nombre complexe $\hat{x}$.).

Ces deux fonctions permettent de reconstruire la fonction $\hat{S}(f)$ et donc, par transformée de Fourier inverse, le signal $s(t)$. Il est donc équivalent de connaître ces deux fonctions ou le signal $s(t)$ lui-même. Pour plus de détails, on se reportera aux nombreux ouvrages de traitement du signal, parmi lesquels [OS75, Cot97, Bel95, dEV92, PTVF92].

Dans le cas d'ondes inharmoniques complexes, on pourra également définir une intensité acoustique dépendante de la longueur d'onde I^λ, telle que l'intensité émise dans la bande $[\lambda, \lambda + d\lambda]$ soit égale à $I^\lambda d\lambda$. L'intensité acoustique de l'onde est alors définie par :

$$I = \int_0^\infty I^\lambda d\lambda. \tag{1.7}$$

Le spectre d'un signal dont l'intensité est constante pour toutes les fréquences, ne sera donc pas identique à celui d'un signal dont l'intensité est constante par bandes d'octaves[10] (c'est ce qui explique la différence entre le bruit blanc et le bruit rose par exemple).

1.3 Régimes continu et impulsionnel. Réponse impulsionnelle

En *régime continu* (ou établi), les phénomènes de propagation ne dépendent plus du temps (car la dépendance temporelle de l'onde est fixe) et on ne considère plus que l'aspect spatial du phénomène. En régime harmonique établi, par exemple, on peut exprimer la pression acoustique sous la forme : $P(M,t) = \hat{P}(M)e^{i\omega t}$ où $\hat{P}$ est l'amplitude (complexe dans le cas général) de la pression acoustique et l'équation 1.3 devient alors l'*équation d'Helmholtz* homogène :

$$\Delta\hat{P} + k^2\hat{P} = 0. \tag{1.8}$$

La plupart des problèmes d'acoustique (diffraction par exemple) seront résolus sous ces hypothèses. Par contre, dans le cadre de l'acoustique virtuelle, on va s'intéresser au contraire à des phénomènes transitoires, comme l'étude de la réverbération d'une salle par exemple, car ce sont eux qui vont porter l'information subjective qui va nous permettre de distinguer un environnement d'un autre, l'«*effet de salle*». D'autre part, on s'intéresse également à des sources émettant des signaux non stationnaires comme de la musique, ou la voix d'un locuteur.

Pour décrire ces phénomènes, on va utiliser une représentation du signal utilisant une fonction, en un

9. ici, $|\hat{x}|$ désigne le module du nombre complexe $\hat{x}$.

10. on rappelle qu'une octave correspond à un doublement de la fréquence. Une bande d'octave est donc deux fois plus large que la bande d'octave précédente et deux fois moins que la suivante.

sens complémentaire d'une fonction harmonique : la *mesure de Dirac* (ou impulsion de Dirac) $\delta(t)$ définie par :

$$\begin{cases} \delta(t) = 0 \; \forall t \neq 0, \\ \int_{-\infty}^{+\infty} \delta(t) dt = 1 \end{cases}$$

On peut également la définir de manière équivalente par :

$$\int_{-\infty}^{+\infty} x(t)\delta(t)dt = x(0),$$

où s est une fonction arbitraire continue en $t = 0$.

Contrairement à un signal harmonique, dont le spectre est constitué d'une seule raie et dont la durée est infinie, ce signal a un spectre constant pour toutes les fréquences et une durée «instantanée» (*cf.* Figure 1.4).

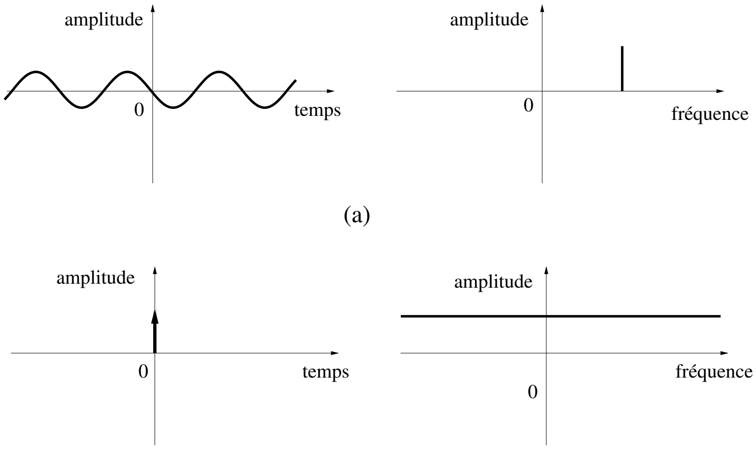


Figure 1.4 - *(a) Un signal harmonique de durée infinie a un spectre constitué d'une seule raie. (b) La mesure de Dirac a une durée instantanée et un spectre constant pour toutes les fréquences.*

On peut alors représenter n'importe quel signal s comme une suite d'impulsions de Dirac :

$$s(t) = \int_{-\infty}^{+\infty} s(\tau)\delta(t-\tau)d\tau.$$

En *régime impulsionnel*, la dépendance temporelle de l'onde n'est pas fixe et on peut exprimer l'émission élémentaire d'une source ponctuelle (par exemple) sous la forme :

$$P(M,t) = \frac{P}{r}\,\delta\left(t - \frac{r}{c}\right),$$

où $\delta(t)$ est une «impulsion» valant 1 à $t = 0$ et 0 sinon. On n'a plus dans ce cas de notion de phase de l'onde comme nous l'avons défini au paragraphe 1.2.1. Néanmoins deux signaux de pressions acoustiques additionnés provoquerons des interférences puisque les pressions (positives ou négatives) s'ajouteront.

Réponse impulsionnelle et canal acoustique

Supposons à présent qu'une source émette une impulsion de Dirac, et considérons le signal reçu par un récepteur. Au cours de leur propagation dans leur milieu (nous supposerons dans la suite de ce document qu'il s'agit de l'air), les ondes sonores sont modifiées de manière complexe (nous le verrons par la suite). Le signal reçu par le récepteur ne sera donc pas dans le cas général une impulsion de Dirac. On supposera que les phénomènes de propagation dépendent de paramètres physiques (pression atmosphérique, température) qui varient lentement avec le temps et seront donc considérés indépendants du temps. On appellera alors «*réponse impulsionnelle*» le signal (pression acoustique) reçu par le récepteur lorsque la source émet une impulsion de pression «ponctuelle» dans le temps (mesure de Dirac).

La transformation subie par un signal sonore, que nous supposerons linéaire, peut en conséquence être exprimée comme un produit de convolution dans le domaine temporel, correspondant à un *filtrage* du signal d'origine :

$$y(t) = \int_0^{+\infty} h(\tau)x(t-\tau)d\tau, \tag{1.9}$$

où $y(t)$ est le signal reçu, $x(t)$ le signal d'origine et $h(t)$ est la réponse impulsionnelle[11] de l'environnement.

Note : On peut également exprimer cette relation par un produit simple dans le domaine fréquentiel :

$$Y(f) = H(f)X(f),$$

où $Y(f)$ et $X(f)$ sont les transformées de Fourier des signaux $y(t)$ et $x(t)$ et $H(f)$ est la *fonction de transfert* de l'environnement, transformée de Fourier de la réponse impulsionnelle $h(t)$.

La réponse impulsionnelle est porteuse de toute l'information concernant les modifications du son durant sa transmission depuis la source jusqu'au récepteur que l'on référence souvent comme le *canal acoustique*. Sa connaissance permet donc de reconstruire, grâce à l'équation de filtrage 1.9, le signal reçu par un récepteur depuis une source dans un environnement donné. On notera ici que l'équation 1.9 est généralement utilisée en traitement numérique du signal sous sa forme discrète [OS75, Cot97, Bel95, dEV92, PTVF92] :

$$y(n) = \sum_m h(m)x(n-m), (n,m) \in \mathbb{N}^2.$$

1.4 Sources et récepteurs sonores

Nous abordons à présent quelques caractéristiques des sources et récepteurs sonores. Dans le cadre qui va nous intéresser dans ce document, nous supposerons que toutes les sources sonores primaires ainsi que les récepteurs sont ponctuels. Bien évidemment le simple modèle de source ponctuelle est insuffisant pour rendre compte de manière réaliste du comportement de sources réelles. À cause de leur taille et des nombreux phénomènes de propagation pouvant intervenir en *champ proche* des sources (c'est-à-dire pour des distances inférieures à la longueur d'onde du signal émis), celles-ci ont un comportement fortement variable, en fonction de la fréquence du signal émis mais aussi de la direction d'émission, qui peut être caractérisé également par une réponse impulsionnelle.

11. dans le reste du document, le terme réponse impulsionnelle désignera une réponse en pression sauf précision contraire.

1.4.1 Modèle de sources/récepteurs ponctuels

On supposera donc que pour les applications qui nous intéressent, le comportement d'une source peut être représenté par celui d'une source ponctuelle à laquelle on ajoute une réponse impulsionnelle dépendant de la direction d'émission $\Gamma(\theta,\phi,t)$. La pression acoustique est alors donnée par :

$$P(r,\theta,\phi,t) = \frac{P}{r}\int_{-\infty}^{+\infty}\Gamma(\theta,\phi,\tau)\delta\left(t-\frac{r}{c}\right)d\tau,$$

où (θ,ϕ) est la direction d'émission de l'onde sonore, r la distance à la source et P l'amplitude de la pression.

On utilisera le même formalisme pour la modélisation des récepteurs sonores, mais la fonction de directivité caractérisera ici la réception en fonction de la direction d'incidence de l'onde sonore.

En pratique, les fonctions de directivité des sources (*resp.* des récepteurs) sont souvent données sous la forme d'une fonction de gain dépendant uniquement de la direction d'incidence et de la fonction de transfert dans l'axe de la source (*resp.* du récepteur) pour différentes fréquences. Le niveau sonore de la source est généralement donné par son niveau de puissance, pression ou intensité acoustique à 1m (on se reportera à l'Annexe A.1 pour les relations entre niveaux de puissance, intensité et pressions acoustiques).

1.4.2 Sources et récepteurs en mouvement. Effet Doppler

Lorsque la source S et/ou le récepteur R sont animés d'un mouvement, l'onde sonore subit une compression dans la direction de déplacement de la source (*resp.* une dilatation dans la direction opposée) entraînant une modification de la fréquence du son reçu par rapport au son émis. Cet effet est appelé *effet Doppler*[12]. Le *décalage Doppler*, rapport entre la fréquence du son émis et celle du son reçu est donné par (*cf.*. Figure 1.5) :

$$\Delta_{Doppler} = \frac{f_R}{f_S} = \frac{1-\frac{\mathbf{n}\cdot\mathbf{v}_R}{c}}{1-\frac{\mathbf{n}\cdot\mathbf{v}_S}{c}}, \tag{1.10}$$

où $\mathbf{v}_S$ est la vitesse de la source, $\mathbf{v}_R$ celle du récepteur et $\mathbf{n} = \frac{\vec{SR}}{||\vec{SR}||}$ est la direction source-récepteur.

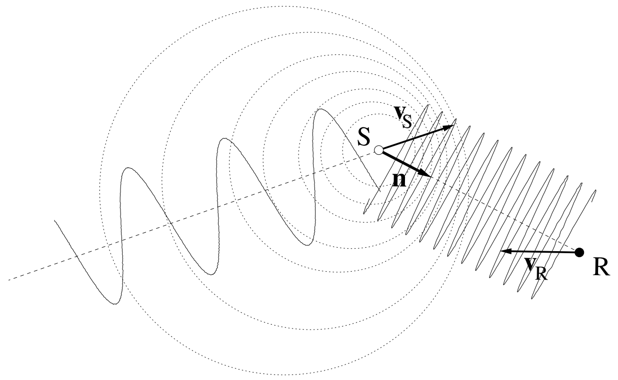


Figure 1.5 - *Notations pour le calcul du décalage Doppler*

12. mis en évidence en 1842 par Christian Johann Doppler.

Le décalage Doppler peut également être exprimé dans le domaine temporel. En effet, si l'on note $\tau(t)$ le délai de propagation variable entre la source et le récepteur alors on peut exprimer le signal arrivant à l'instant t au récepteur $r(t)$ comme :

$$r(t) = s(t - \tau(t)),$$

où $s(t)$ est le signal émis par la source au temps t. Le signal reçu peut donc être exprimé par ré-échantillonnage du signal émis. Il est nécessaire pour cela de connaître le délai $\tau(t)$. On peut l'exprimer sous la forme :

$$\tau(t) = \frac{1}{c}\|\mathbf{R}(t) - \mathbf{S}(t - \tau(t))\|,$$

où $\mathbf{R}(t)$ et $\mathbf{S}(t)$ sont les positions respectives du récepteur et de la source à t[13]. Cette équation n'est pas linéaire en τ et ne peut être, en général, résolue directement. Nous aurons l'occasion de montrer une manière efficace d'approcher son résultat dans la suite de ce document (cf. paragraphe 7.4.2).

1.5 Phénomènes de propagation

Dans les paragraphes précédents, nous nous sommes intéressés à la propagation d'une onde sonore en *champ libre*, c'est-à-dire indépendamment de la présence d'obstacles, de parois, *etc.* Or ceux-ci vont perturber la propagation de l'onde sonore à travers différents phénomènes qui vont avoir pour conséquences principales de l'atténuer (réfraction/transmission, dissipation) et de modifier sa direction de propagation (réflexion, diffraction, réfraction).

1.5.1 Réflexion

Lorsque l'onde sonore va toucher une paroi, une partie de son énergie va être en général ré-émise sous forme d'une onde émanant de la paroi, dont l'amplitude et la phase vont être différentes de celles de l'onde incidente (les modifications étant principalement dues aux phénomènes de transmission/dissipation à l'interface entre le milieu de propagation et la paroi ainsi qu'aux vibrations de celle-ci).

Réflexion spéculaire

Si la surface de la paroi est parfaitement plane et rigide, la réflexion est dite *spéculaire* et suit la loi de Snell-Descartes (*cf.* Figure 1.6 (a)). Par extension, on dira également que le matériau est spéculaire. Ce modèle de réflexion est le plus utilisé en acoustique du fait de la grandeur relative des longueurs d'ondes sonores et des aspérités des parois (si l'on songe à des aspérités de l'ordre du centimètre, elles n'interféreront *a priori* qu'avec des longueurs d'onde inférieures au centimètre, c'est-à-dire à partir de 10 kHz seulement). Dans ce cas, le changement d'amplitude et de phase est caractérisé par le coefficient de réflexion complexe de la paroi :

$$\hat{R}^{\lambda}(\theta,\phi) = R^{\lambda}(\theta,\phi)e^{i\chi^{\lambda}(\theta,\phi)},$$

où χ^{λ} traduit le déphasage induit par la réflexion et $R^{\lambda}(\theta,\phi)$ la variation de l'amplitude de la pression acoustique.

13. on peut remarquer que cette expression revient à considérer une source mobile par rapport à un microphone fixe au temps t.

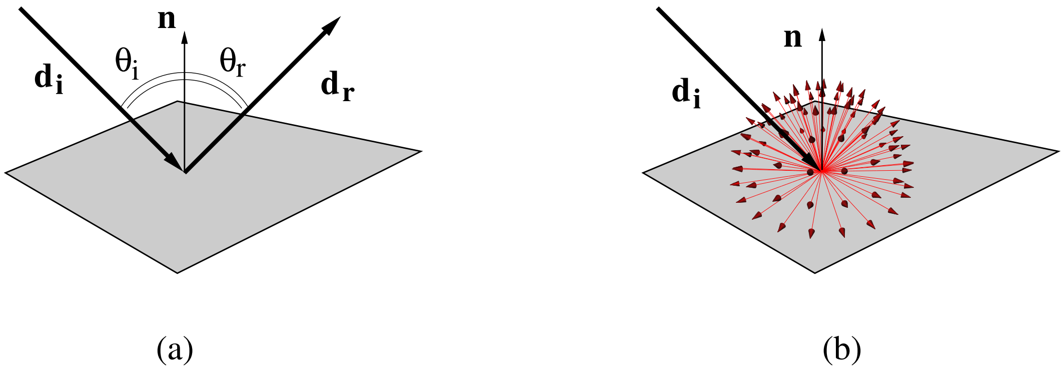


Figure 1.6 - *(a) Réflexion spéculaire selon la loi de Snell-Descartes :* $\mathbf{d}_r = \mathbf{d}_i - 2(\mathbf{n}\cdot\mathbf{d}_i)\mathbf{n}, \theta_i \ = \ \theta_r$. *(b) Réflexion diffuse. L'énergie incidente est réfléchie uniformément dans tout l'hémisphère supérieur du réflecteur quelque soit sa direction d'incidence.*

On utilisera généralement le *coefficient d'absorption* α^λ, qui mesure la proportion de l'*intensité* qui est absorbée :

$$\alpha^\lambda(\theta,\phi) = 1 - |\hat{R}^\lambda(\theta,\phi)|^2.$$

Le coefficient d'absorption est également relié à une autre grandeur, plus proche des phénomènes physiques, l'*impédance normale de la paroi* :

$$Z = \left(\frac{p}{v_n}\right)_{\text{surface}},$$

définie comme le rapport entre la pression p et la composante normale de la vitesse particulaire v_n à la surface de la paroi. Comme le coefficient d'absorption, cette grandeur est généralement complexe et dépendante de l'angle d'incidence. Toutefois, dans le cadre de matériaux rigides, elle est indépendante de la direction d'incidence (contrairement au coefficient d'absorption). On se reportera à l'Annexe A.2.3 pour plus de détails sur les relations entre impédance et coefficient d'absorption.

Étant donné la difficulté de réaliser les mesures des coefficients de réflexion, c'est généralement un coefficient d'absorption à incidence aléatoire ou *coefficient d'absorption de Sabine* qui est utilisé. On se reportera à l'Annexe A.2.3 pour plus de détails et des exemples numériques de ces coefficients.

Réflexion diffuse

Néanmoins, si la surface n'est pas plane et que ses aspérités sont d'une taille de l'ordre de (ou supérieure à) la longueur d'onde sonore considérée, on modélisera la réflexion, dite alors *diffuse*, suivant la loi de Lambert. L'onde incidente est réfléchie dans toutes les directions avec une même intensité I_r, quelle que soit la direction d'incidence (*cf.* Figure 1.6 (b)) :

$$I_r = I_i(\mathbf{d_i}\cdot\mathbf{n}),$$

où I_i est l'intensité incidente et $\mathbf{d_i}$ la direction d'incidence.

Par extension, on dira que le matériau est diffus ou lambertien. Comme nous le verrons un peu plus tard, il est également très intéressant de pouvoir modéliser des réflexions diffuses dans le cadre d'une simulation acoustique. Dans la pratique, la réflexion diffuse est caractérisée par un *coefficient de diffusion* $\beta^\lambda(\theta,\phi)$, mesurant la part de l'intensité qui est réfléchie de manière diffuse. On pourra se reporter à l'Annexe A.2.3 pour plus de détails sur ce coefficient. En général, ce modèle est moins utilisé en acoustique que le modèle de réflexion spéculaire étant donné les longueurs d'onde considérées. Néanmoins, de nombreuses études ont montré qu'il ne pouvait pas être négligé, en particulier dans le cas de hautes fréquences [Hod91].

Dans la réalité la réflexion du son ne donne pas lieu à des réflexions purement diffuses ou purement spéculaires, mais également à des réflexions «diffuses directionnelles» (*cf.* Figure 1.7) . La réflexion peut être alors caractérisée par une fonction de distribution de la réflectance bidirectionnelle (ou FDRB) qui décrit, pour une direction d'incidence donnée, la répartition de l'énergie réfléchie. Quelques essais ont été effectués pour tenter de mesurer ces fonctions de distribution [KGB97, Rin93, WH91]. Néanmoins, dans le cadre de cette étude, nous nous limiterons aux deux cas précédents.

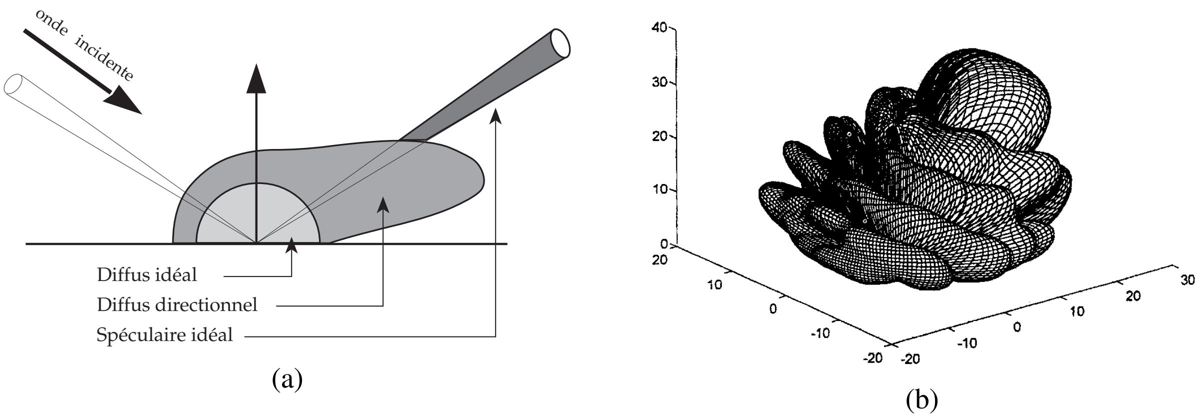


Figure 1.7 - *(a) Représentation schématique d'une fonction de réflectance générale*[11]. *(b) Exemple de la distribution directionnelle de l'intensité réfléchie pour une onde sonore à 15 kHz en incidence oblique sur un disque plan en plastique acrylique de 9 cm de diamètre environ (d'après [KGB97]).*

1.5.2 Diffraction, "visibilité sonore"

En 1665, Grimaldi publie les résultats d'une série d'expériences[12] sur la propagation de la lumière en présence d'obstacles. Il remarque la présence de franges lumineuses dans la zone d'ombre géométrique créée par l'obstacle et introduit un nouveau terme, «diffraction»[13], pour caractériser ce nouveau mode de propagation de la lumière qu'il ne considère être, ni de la réflexion, ni de la réfraction, ni de l'éclairage direct de la source.

La diffraction désigne le changement de direction des ondes sonores dû à la présence d'obstacles ou d'inhomogénéités dans le milieu de propagation et les interférences qui en résultent. Dans le cas des ondes sonores, ce sont en grande partie les phénomènes de diffraction qui expliquent qu'une source sonore puisse être audible alors qu'elle n'est pas directement visible par l'auditeur (c'est le cas dans une fosse d'orchestre par exemple). Les problèmes liés à la diffraction des ondes (que ce soit des ondes électromagnétiques ou mécaniques) restent encore aujourd'hui parmi les plus délicats à résoudre et des solutions exactes n'existent que pour un nombre très limité de configurations. Nous présenterons au

11. figure tirée de l'ouvrage «*Radiosity and Global Illumination*» par F.X. Sillion et C. Puech [SP94].
12. *Physico-Mathesis de lumine, coloribus et iride*. Bologne, 1665.
13. Newton utilisera également le terme d'*inflexion* dans son *Optique* de 1704.

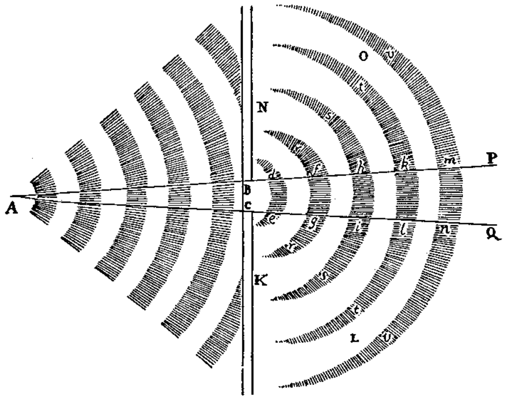


Figure 1.8 - *Schéma tiré du* Principia *de Newton (1686) illustrant le passage d'ondes à travers une ouverture. Le point* A *est la source ; l'ouverture est décrite par les points* B *et* C *;* d-e, f-g, h-i, *représentent les "sommets de plusieurs ondes séparés par autant de vallées ou creux "*[14].

Chapitre 3 quelques solutions classiques pour étudier la diffraction des ondes sonores et au Chapitre 6 une nouvelle approche pour caractériser la «visibilité» sonore.

1.5.3 Autres phénomènes mis en jeu

D'autres phénomènes vont également modifier la direction de propagation ou l'amplitude d'une onde sonore, en particulier les phénomènes de réfraction et transmission.

La réfraction désigne un changement de direction de propagation de l'onde sonore due à une variation de sa vitesse de propagation. Dans le cas des ondes sonores cela se produit à l'interface entre deux milieux de propagation mais également lors de variation de la température ou des conditions météorologiques (vent) qui vont influer sur la vitesse de propagation (dans le cas d'acoustique en milieu extérieur) [Ing53, Def96].

La transmission, elle, caractérise le passage d'une partie de l'onde sonore «à travers» une paroi. Dans le cadre de cette étude, nous ne nous intéresseront toutefois que très peu à ces phénomènes.

Un dernier phénomène significatif est l'*absorption atmosphérique* qui dépend de la fréquence et de la distance parcourue par l'onde. Elle est importante d'un point de vue perceptif pour l'estimation de la distance à la source (cf. Annexe A.2.2).

1.6 Acoustique des espaces clos. Réverbération

Dans le cas d'espaces clos, ou d'environnements comportant des parois, la pression acoustique est solution d'un système formé de l'équation d'onde à laquelle on ajoute des conditions aux limites dues à la présence de parois :

$$\begin{cases} \Delta p - \dfrac{1}{c^2}\dfrac{\partial^2 p}{\partial t^2} = P_0 & \text{dans le volume,} \\ Z\dfrac{\partial p}{\partial n} + i\omega\rho_0 p = 0 & \text{sur les parois}^{15}\ , \end{cases}$$

14. adapté du *Principia* de Sir Isaac Newton, 4ème édition, 1726, ré-édition de 1871 par MacLehose, Glasgow.

où $\frac{\partial p}{\partial n}$ dénote la dérivée de la pression normale à la paroi définie par $\frac{\partial p}{\partial n} = \vec{n} \cdot \vec{\nabla} p$ ($\vec{n}$ étant la normale à la paroi), ρ_0 est la densité du milieu au repos, Z est l'impédance de la paroi et P_0 est la pression créée par les sources. Nous examinerons au chapitre 3 diverses solutions pour résoudre ce système.

Dans ce cas, les ondes sonores provenant des sources subissent de multiples réflexions au contact des parois. Ces réflexions entraînent l'existence de trajets multiples de longueurs différentes entre une source et un récepteur. Le signal reçu est donc constitué de la superposition de plusieurs répétitions (que l'on nomme communément «échos») du signal émis. Ce phénomène est appelé *réverbération*.

1.6.1 Structure temporelle de la réverbération

Il est intéressant d'examiner la structure temporelle et fréquentielle de la réponse impulsionnelle d'un environnement réverbérant, en particulier dans le cas d'un volume clos puisque cela va traduire l'effet de salle. Généralement il est plus simple de l'étudier en examinant un *échogramme*, représentant la puissance instantanée dans la réponse impulsionnelle en fonction du temps. La structure temporelle de la réponse impulsionnelle d'un environnement réverbérant peut être généralement divisée en trois parties, clairement visibles sur l'échogramme (*cf.* Figure 1.9 (a)) :

– le *son direct* arrivant directement depuis la source jusqu'au récepteur ;

– les *réflexions précoces*, contributions de l'onde sonore ayant subi un faible nombre de réflexions (de l'ordre de 1 à 5 en moyenne) avant d'arriver au récepteur et qui sont temporellement séparables ;

– la *réverbération tardive*, dans laquelle de très nombreuses réflexions d'ordre élevés se superposent, formant un continuum et ne pouvant plus être individuellement séparées.

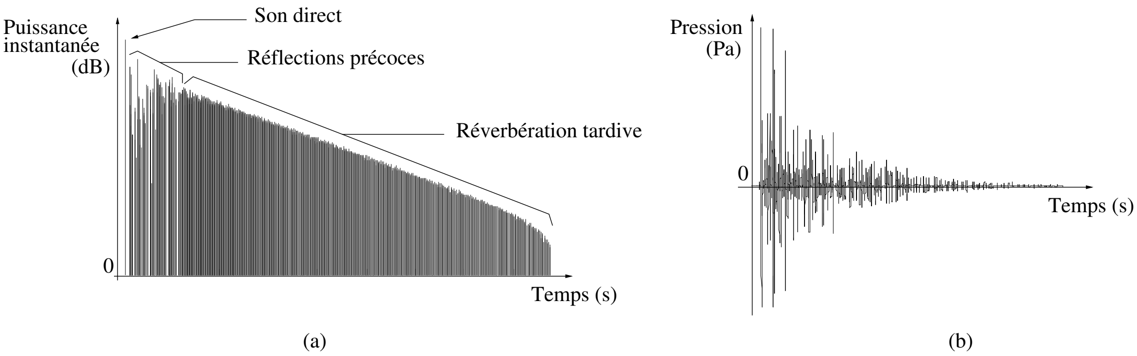


Figure 1.9 - *(a) Échogramme et structure temporelle de la réponse impulsionnelle. (b) Réponse impulsionnelle en pression.*

On peut montrer, dans le cas de volumes clos, qu'à partir d'une certaine fréquence appelée *fréquence de Schroeder* et d'un certain temps, le *temps de mélange*, la réponse impulsionnelle peut être définie par un processus stochastique indépendant de la position de la source et du récepteur [Kut91b]. C'est ce qui justifie physiquement la séparation entre réflexions précoces et réverbération tardive. Nous en reparlerons au paragraphe 3.5. En espace ouvert, la structure de la réverbération tardive pourra bien évidemment être

15. équation de la *condition aux limites de Robin* valable dans beaucoup de cas pratiques pour des parois absorbantes de type revêtements muraux. D'autres possibilités peuvent être utilisées, en particulier la *condition aux limites de Neumann*, valable dans le cas de parois parfaitement rigides dont nous reparlerons au paragraphe 3.3.1.

très différente de celle que nous avons décrite puisque les surfaces généreront moins de réflexions ou du moins des réflexions dont les directions d'incidence au point de réception seront plus corrélées.

1.6.2 Réverbération tardive et notion de champ diffus. Temps de réverbération

La réverbération tardive dans les volumes clos est très souvent associée à la notion de *champ diffus* car de très nombreuses réflexions du son atteignent les surfaces à chaque instant et leurs directions d'incidences sont supposées suffisamment décorrélées. Par définition, un champ diffus est un champ sonore caractérisé par les deux conditions suivantes :

– en n'importe quelle position dans le volume, les ondes sonores incidentes proviennent de toutes les directions, avec la même intensité et des phases relatives aléatoires,

– le champ réverbéré est le même quel que soit la position dans le volume.

Ces conditions, très fortes, sont à la base de tous les travaux d'acoustique statistiques, permettant de calculer de manière rapide la réverbération tardive d'un volume. Elles ne sont pourtant pas toujours réalisées en pratique et dépendent de différents facteurs, principalement les dimensions du volume, la diffusivité et l'absorption des parois. On trouvera dans [Hod96] une discussion de la validité des hypothèses de champ diffus. Il faut bien noter ici que «champ diffus» n'est pas équivalent à «surfaces diffusantes», même si cette dernière condition est très souvent nécessaire. Les hypothèses de champ diffus restent néanmoins plus fortes et des parois diffusantes peuvent ne pas suffire à l'établissement d'un champ diffus dans l'environnement. Les hypothèses de champ diffus permettent d'établir des bilans énergétiques d'où l'on peut tirer des lois gouvernant le phénomène de réverbération. On peut montrer en particulier que, sous ces hypothèses, la décroissance de l'intensité sonore suit une loi exponentielle décroissante en fonction du temps (linéaire si on l'exprime en décibels, *cf.* Figure 1.9 et Annexe A.1) [Kut91b].

Champ ergodique. Temps de réverbération

Une hypothèse un peu moins contraignante que celle de champ diffu est souvent utilisée dans les problèmes d'acoustique : celle de *champ ergodique*. Cette hypothèse est basée sur une formulation géométrique du problème. On considère que le front d'onde va se déplacer le long de rayons [Pol93] dans un volume fermé en subissant des réflexions spéculaires sur les parois. Considèrons le mouvement d'une masse ponctuelle le long de ces rayons (un «phonon» [Joy75]). Le volume est dit ergodique si le long de n'importe quelle trajectoire, la particule passe autant de temps au voisinage de n'importe quel couple (position,orientation). Comme dans le cas du champ diffu, la validité de cette hypothèse dépend de la géométrie et de l'absorption des surfaces de l'environnement considéré. L'hypothèse de champ ergodique permet de donner une approximation d'une grandeur importante caractérisant l'acoustique d'un espace clos : le *temps de réverbération*. Il est communément défini par le temps nécessaire pour que l'intensité acoustique dans l'environnement décroisse d'un facteur 10^{-6} par rapport au niveau initialement émis par la source (équivalent à un gain de -60 dB) après l'arrêt de toute émission sonore. Un modèle de cette valeur (pour des salles de grand volume) est le temps de réverbération *de Eyring* (exprimé en secondes[16]) est donnée par [Kut91b] :

$$T_r^{\lambda} = -0.163 \frac{V}{S \ln\left(1 - \bar{\alpha}^{\lambda}\right) + 4(1-\mu^{\lambda})V},$$

16. l'homogénéité de la formule est assurée par les unités de la constante multiplicative.

où V est le volume de l'espace considéré (en m^3), $S = \sum_i S_i$ la surface totale des parois (en m^2), $\bar{\alpha}^\lambda = 1/S\sum_i S_i\alpha_i^\lambda$ leur absorption moyenne et μ^λ le coefficient d'atténuation atmosphérique par mètre (décrit en Annexe A.2.2). Bien que limitée (hypothèses de champ ergodique), elle reste la meilleure approximation du temps de réverbération dans le cas où l'on ne dispose pas de la géométrie exacte de la salle. Une autre formule très utilisée dans le domaine de l'acoustique des salles est celle du temps de réverbération *de Sabine*. Elle correspond *dans l'hypothèse de valeurs du coefficient d'absorption petites devant 1*, à effectuer un développement limité du logarithme présent dans la formule d'Eyring et ne conserver que le terme de premier ordre [Kut91b, Joy75] :

$$T_r^\lambda = 0.163\frac{V}{S\bar{\alpha}^\lambda + 4(1-\mu^\lambda)V},$$

avec $\bar{\alpha}^\lambda << 1$.
Ces formules permettent d'évaluer également l'absorption moyenne des surfaces (coefficient d'absorption de Eyring ou de Sabine).

1.7 En bref

Dans ce chapitre, nous avons rappelé quelques bases d'acoustique physique, nécessaires *a priori* pour la compréhension du manuscrit. Pour plus d'information on pourra se reporter aux références [Kut91b, Eve94, DW83]. Parmi ces notions clés, on retiendra particulièrement :

– que le son est une *onde mécanique*, nécessitant un milieu de propagation (ici l'air) ;

– que sa propagation peut être caractérisée par les *variations de la pression* atmosphérique dues à une source de perturbation ;

– que la grandeur pertinente pour l'étude des phénomènes acoustiques est la *pression acoustique* définie comme la différence entre la pression instantanée en un point du milieu et la pression atmosphérique de ce point au repos ;

– qu'une onde sonore est soumise à de multiples *phénomènes de propagation* dépendant de la *fréquence* de l'onde (réflexion, diffraction, transmission, effet Doppler, *etc.*) qui vont modifier ses caractéristiques de manière complexe ;

– que dans l'hypothèse où tout ces phénomènes sont *linéaires*, toute l'information de propagation d'un signal depuis une source jusqu'à un point de réception peut être représenté par un *filtre* temporel appelé *réponse impulsionnelle* ;

– qu'on peut l'assimiler au signal reçu par le récepteur lorsque la source émet une «impulsion» de pression instantanée ;

– que ce filtre peut s'exprimer de manière équivalente dans le domaine fréquentiel (on le nomme alors *fonction de transfert*) , les deux représentations étant liées par *transformée de Fourier* ;

– qu'il permet, en outre, par convolution avec un signal quelconque émis par la source d'obtenir le signal au point de réception ;

– qu'enfin, dans le cas d'environnements clos ou comportant des parois, les réflexions de l'onde sonore sont à l'origine du phénomène de *réverbération* dont l'étude est particulièrement importante pour le modélisation des réponses impulsionnelles.

Nous allons à présent voir que, comme le remarquait Lord Rayleigh, il est particulièrement important de considérer dans l'étude des phénomènes acoustiques, la *perception* que notre système auditif nous en offre.

Chapitre 2

Acoustique perceptive : le son subjectif

UNE large part de la recherche en acoustique s'intéresse au son comme responsable de la sensation d'audition qui est un phénomène extrêmement complexe. L'aspect perceptif est bien évidemment un facteur de premier plan dans les applications d'acoustique virtuelle comme celles qui nous concernent dans ce document, puisqu'elles s'adressent en premier lieu à des auditeurs humains. Bien que ce travail n'apporte pas de contribution au domaine de la psychoacoustique, il en utilise indirectement certains résultats dont nous présentons ici quelques aspects. Nous commencerons par présenter comment l'oreille perçoit l'intensité et la fréquence d'un son. Puis nous nous intéresseront particulièrement à l'aspect spatial de la perception sonore qui a donné lieu au développement des techniques de «son 3D». Nous évoquerons en dernier lieu les aspects perceptifs liés aux environnements réverbérants.

2.1 L'oreille, récepteur du son

L'oreille est un organe qui enregistre les variation du champ de pression en deux points : les deux *tympans*. Ce sont les micro-déplacements des tympans dus à l'incidence d'une onde sonore qui vont être à l'origine de la sensation d'audition. Ces micro-déplacements entraînent le mouvement de trois osselets, qui vont à leur tour transmettre des variations de pression (avec un rapport 15 pour 1) à un liquide contenu dans la *cochlée*, cavité en forme de spirale de l'oreille interne. Enfin, les variations de pression de ce liquide vont être enregistrées le long d'une membrane séparant la cochlée en deux parties : la *membrane basilaire*, comprenant environ 25000 terminaisons nerveuses du nerf auditif principal.

Comme nous allons le voir à présent, l'oreille est un récepteur exceptionnel qui n'a rien de commun avec d'autres récepteurs artificiels comme les microphones. En effet, elle nous permet d'analyser de manière extrêmement fine tant du point de vue fréquentiel, que directionnel les ondes sonores qui les atteignent, et ce sur une plage dynamique extrêmement étendue. Nous invitons également le lecteur intéressé à consulter les ouvrages suivants [Ber86, Bla83, ZF81] pour plus de détails.

2.1.1 Perception de l'intensité d'une onde sonore

L'oreille humaine perçoit des signaux dont la dynamique d'intensité sonore atteint 10^{12} W.m^{-2}, le son le plus faible perceptible ayant une intensité de 10^{-12} W.m^{-2} (seuil d'audition), le plus fort de 1 W.m^{-2} (seuil de douleur). Cela correspond à une amplitude de la pression acoustique de 2.10^{-5} à 20 Pa[1]. Néanmoins l'oreille perçoit les différences d'intensité acoustique de manière logarithmique,

1. on peut noter ici que les plus faibles pressions acoustiques auxquelles notre oreille est sensible entraînent des déplacements du tympan de l'ordre de $10^{-9}cm$, moins du dixième du diamètre d'une molécule d'hydrogène !

la sensation d'audition variant comme le logarithme de l'excitation. Ces deux aspects justifient bien évidemment l'utilisation d'une unité de mesure logarithmique comme le décibel. On notera également que l'intensité perçue ou «*sonie*» n'est pas uniquement liée à l'amplitude de la pression acoustique mesurée. Elle dépend en particulier de la fréquence. En particulier, à niveaux de pression acoustique égaux, les sons de fréquence basse paraîtrons moins «forts» que les sons de fréquences plus élevées.

2.1.2 Perception fréquentielle du son. Bandes d'octaves

On peut également s'intéresser à la manière dont l'oreille est sensible aux différentes fréquences. La sensibilité en fréquence varie beaucoup suivant les individus. Chez l'enfant, le seuil supérieur d'audition se situe autour de 20 kHz. Chez l'adulte, il tombe rapidement au dessous des 15 kHz. La limite d'audition basse se situe généralement vers 16 Hz. La bande passante de l'oreille est donc approximativement de 20 kHz, ce qui justifie, entre autre, les fréquences d'échantillonnages communément utilisées pour la représentation numérique des signaux audio (44.1 kHz pour le CD, 48 kHz pour le DAT, le DVD ou les formats professionnels type ADAT[2]). On peut remarquer également que l'oreille est particulièrement sensible aux variations de fréquence dans une zone entre 1 kHz et 3 kHz, qui correspond aux fréquences rencontrées dans la voix humaine. Dans cette zone l'oreille peut percevoir des différences de l'ordre de 3 Hz. (environ 10 Hz dans les basses fréquences au dessous de 70Hz et de l'ordre de 30 Hz dans les hautes fréquences). Sur toute la bande de fréquence audible, un auditeur moyen peut distinguer environ 2000 changements de hauteur fréquentielle. On notera également que la sensation de «*hauteur*» subjective d'un son (ou «*tonie*») évolue en fonction de l'intensité de l'onde sonore. La hauteur perçue n'est donc pas directement reliée à la fréquence de l'onde sonore [Ber86, ZF81]. Jusqu'à 3000 Hz environ, la hauteur perçue décroît lorsque l'intensité de l'onde augmente ; pour les fréquences plus élevées, le phénomène s'inverse. Une conséquence directe de ce fait est qu'il faut ajuster les niveaux de deux signaux pour pouvoir les comparer subjectivement.

Lorsque des sons au contenu fréquentiel riche sont perçus, on assiste à des phénomènes de masquage , une partie du signal devenant inaudible. L'étude de ces phénomènes a montré qu'un son ne pouvait en masquer un autre que s'il possédait des fréquences comprises dans une bande autour de la fréquence du son à masquer. Ce résultat prouve que le système auditif (au niveau de la membrane basilaire) analyse les signaux dans différentes plages de fréquences appelées *bandes critiques* . Bien qu'elles se recouvrent continûment, on les divise communément en 24 bandes adjacentes. La largeur de chaque bande critique est de l'ordre du tiers d'octave. On considère alors fréquemment qu'il n'est pas nécessaire d'utiliser une résolution fréquentielle inférieure. Cela justifie la pratique courante de mesures en tiers d'octaves par exemple. Ce comportement de l'oreille, agissant comme un banc de filtres est également à l'origine des méthodes de simulation acoustique en bandes d'octaves (*cf.* paragraphe 4.1.3).

2.1.3 Perception temporelle du son. Influence de la phase

L'oreille a également une certaine résolution temporelle. De nombreuses études ont montré que l'oreille effectue une intégration temporelle des signaux reçus. Un certain temps est donc nécessaire pour que le niveau d'un son bref ne s'établisse.

2. on rappelle que, suivant le *théorème de Shannon*, il est nécessaire d'échantillonner un signal analogique à une fréquence double de la plus haute fréquence qu'il contient pour pouvoir le reconstruire parfaitement à partir des valeurs numériques (ou échantillons). Cette fréquence d'échantillonnage minimale est appelée *fréquence de Nyquist* [Cot97, OS75, Bel95, dEV92].

Effet d'antériorité

Un autre phénomène lié à la perception temporelle du son est le masquage intervenant lors de l'arrivée à des temps voisins de deux sons cohérents. Le son arrivé le premier masque alors le second. C'est l'*effet d'antériorité* (ou «loi du premier front d'onde»)[3] mis en évidence dès 1849. On le remarque très facilement en se rapprochant d'une des deux enceintes d'un système stéréophonique ; aucun son ne semble alors provenir de l'enceinte la plus éloignée. Le masquage dépend bien entendu du niveau relatif des deux signaux. Les études effectuées sur l'effet d'antériorité ont également des conséquences sur la perception des réflexions du son dans les environnements réverbérants.

Influence de la phase

S'il est clairement admis que l'oreille est particulièrement sensible au variations d'amplitude et de fréquence d'un signal et donc en particulier à son spectre, on peut s'interroger sur l'importance de sa phase sur la perception du son. Des études ont montré que les différences dans la phase de signaux était perceptibles [PS69]. Le changement de timbre impliqué est plus important à basse fréquence et est indépendant de l'intensité subjective du signal.

2.2 Perception spatiale du son. Filtrage directif de la tête

Notre système auditif ne nous permet pas seulement de percevoir la sensation auditive du signal provenant d'une source sonore mais il nous donne également des informations spatiales relative à la localisation spatiale de cette source. Comme notre système visuel, l'information qu'il restitue est tridimensionnelle.

2.2.1 Perception de la distance à la source

La perception de la distance à la source est *a priori* guidée par deux facteurs : la décroissance de l'amplitude de la pression sonore en fonction de la distance à la source (sans doute le plus important) et l'atténuation des hautes fréquences due à l'absorption atmosphérique [Col68]. D'autres facteurs peuvent venir s'y ajouter dans le cas d'environnements réverbérants (rapport entre l'intensité directe et l'intensité réfléchie par exemple). Un son lointain aura donc une intensité plus faible et verra ses hautes fréquences atténuées. Néanmoins, ces deux critères, bien que nécessaires, peuvent être insuffisants pour donner une impression correcte de la distance à la source, qui reste très variable suivant les auditeurs.

2.2.2 Perception directionnelle. Filtrage directif de la tête

L'un des points les plus important de la perception auditive est la perception directionnelle des sources sonores. Tout d'abord, la différence de temps d'arrivée à chaque oreille va varier en fonction de la position de la source [Beg94]. De plus, lorsque l'onde sonore rencontre la tête d'un auditeur, plusieurs phénomènes complexes (diffraction, réflexions et atténuation causées par les épaules, la tête, le pavillon et le conduit auditif externe de l'oreille) vont modifier de manière spécifique à la direction d'incidence, les signaux reçus à chaque oreille. Ces modifications peuvent être représentée par une paire de filtres (un pour chaque oreille) directifs appelés *fonctions de transfert de la tête*. La connaissance de ces filtres permet donc de reproduire la perception spatiale d'une source sonore virtuelle. C'est pourquoi la

3. appelé aussi parfois *effet Hass* suite à ses expériences de 1951 sur le sujet.

recherche concernant la mesure [GM94], la représentation [Eme95] et l'utilisation [Beg94] de tels filtres fait l'objet d'intenses recherches dans la communauté acoustique. Si les résultats obtenus sont déjà extrêmement convaincants, des problèmes subsistent néanmoins. En effet, ces filtres s'avèrent être très spécifiques d'un auditeur donné ce qui impose en théorie de mesurer ses propres fonctions de transfert pour un effet optimal. Généralement ce sont des fonctions de transfert mesurées[4] sur plusieurs auditeurs (et moyennées) ou sur des «têtes artificielles» [GM94], voire complètement synthétisées [Bur92b] qui sont utilisées, ce qui peut diminuer la qualité de l'effet produit. En particulier, de fréquentes confusions avant/arrière se produisent alors, en particulier quand la source est n'est pas visible. Nous reviendrons sur l'utilisation des fonctions de transfert de la tête dans le contexte de réalité virtuelle auditive dans le paragraphe 4.1.2. Des informations complémentaires sur la perception spatiale du son et l'utilisation des fonctions de transfert de la tête pourront être trouvées dans [Beg94, Bla83].

2.2.3 Perception des sources en mouvement

Dans le cas de sources en mouvement, la fréquence du son perçu varie continûment en raison de l'effet Doppler (*cf.* paragraphe 1.4.2). Ce changement dans la fréquence, que l'on observe couramment dans la vie quotidienne avec les véhicules en mouvement, est donc particulièrement perceptible et est essentiel pour le réalisme d'une simulation avec sources (et/ou récepteurs) mobiles [Str98, Cho71].

2.2.4 Localisation dans le cas de sources multiples

Les phénomènes d'intégration temporelle et de masquage que nous avons évoqués plus haut ont bien évidemment des conséquences directes sur la localisation d'un événement sonore dans le cas de sources multiples. Considérons encore une fois le cas de deux sources émettant des signaux cohérents. On distingue alors trois effets suivant le retard relatif et le niveau des deux sources :

– lorsque les retards et les niveaux des deux sources sont très proches, un événement unique est perçu et sa position dépend de celle des deux sources ;

– lorsque le retard est compris entre 630 μs et 1 ms, il y a effet d'antériorité : un seul événement sonore est perçu et sa position est celle de l'une des deux sources ;

– enfin deux événements distincts à deux positions distinctes peuvent apparaître, en particulier lorsque le retard et la différence de niveau relative sont plus importants.

2.3 Perception dans un milieu réverbérant

De nombreuses études psychoacoustiques ont été effectuée pour mesurer l'influence d'un milieu réverbérant sur la perception des sons. On s'intéressera particulièrement dans ce cadre aux espaces clos.

Réflexions précoces et réverbération tardive

Comme nous l'avons vu au chapitre précédent la réverbération peut être découpée en différentes parties : le son direct, les premières réflexions et la réverbération tardive. Ce découpage structurel est également justifié perceptiblement. Les réflexions précoces sont très importantes et vont modifier le son perçu en fonction de la position des sources et des auditeurs. Ce sont elles qui vont conditionner en partie la qualité acoustique d'un lieu d'écoute [Bar71, BM81, Ber96]. La perception des différents échos est

4. on place pour cela deux micros dans les conduits auditifs.

liée à la perception de sources multiples, telle que nous l'avons présenté plus haut. Les réflexions du son peuvent entraîner un décalage de la position perçue d'une source sonore [Har83] Le niveau du son réverbéré par rapport au son direct va également donner des indices auditifs sur la distance à la source.

La réverbération tardive est tout aussi importante d'un point de vue perceptif mais est plus une caractéristique de l'environnement. C'est-à-dire qu'elle varie peu lorsque l'on modifie la position de la source ou que l'auditeur se déplace. Ce dernier point est intéressant puisqu'il implique que les réflexions tardives peuvent être recalculées moins fréquemment que les réflexions précoces. Des progrès récents en psychoacoustique ont permis de déterminer différents critères permettant d'apprécier la qualité *a priori* d'un lieu découte. Nous en reparlerons plus en détail au Chapitre 10.

Importance de la phase

Comme nous l'avons vu, la phase peut avoir des effets perceptibles sur le timbre d'un signal. Toutefois, Kuttruff a montré [Kut91a] que dans le cas d'un milieu réverbérant, toute relation de phase était détruite à partir d'une certaine distance à la source qu'il appelle distance de réverbération $d_r\ (\mathrm{m}) = 0.06\sqrt{\frac{V}{T_r}}$ (où V (m^3) est le volume de l'espace considéré et T_r (s) le temps de réverbération). Cette constatation apporte une certaine validité aux approches qui négligent la phase du signal pour la simulation de la propagation en lieux clos.

2.4 En bref

Dans ce chapitre nous avons présenté quelques bases importantes d'acoustique perceptive qui sont sous-jacentes à tout système de simulation acoustique. Parmi les points importants, on retiendra en particulier :

- que l'oreille est un récepteur extrêmement performant de la pression acoustique ;
- qu'elle perçoit une dynamique d'intensités acoustiques de 12 ordres de grandeurs :
- qu'elle est également sensible à une plage fréquentielle couvrant 3 ordres de grandeur ;
- que le système auditif agit comme un banc de filtres, analysant le signal reçu en séparant ses composantes fréquentielles en différentes bandes : les *bandes critiques*. Ces bandes critiques ont une largeur de l'ordre du tiers d'octave ;
- que l'oreille est particulièrement sensible aux variations de fréquence et d'intensité d'un signal et dans une moindre mesure sensible aux différences de phase ;
- que des phénomènes de masquage sont susceptibles d'apparaître dans le cas de sons au contenu fréquentiel proche et atteignant l'oreille à des temps voisins ;
- que la réverbération influence grandement la perception du son. On peut distinguer, là encore, les différentes composantes de la réverbération. L'influence des réflexions précoces est variable en fonction de la position de la source ou de l'auditeur dans le lieu d'écoute. Celle de la réverbération tardive est plus caractéristique du lieu lui-même.

❒ ❒ ❒

Dans ces premiers chapitres, nous avons présenté quelques bases d'acoustique physique et perceptive, prérequis nécessaires à une bonne compréhension de la suite du document. Nous avons, en particulier, vu que toute l'information nécessaire pour représenter la propagation d'une onde sonore entre une source et un récepteur peut être représentée par un signal temporel : la réponse impulsionnelle de l'environnement. De plus, le signal reçu peut être obtenu à partir du signal émis par une opération de filtrage. Simuler la réponse impulsionnelle d'un environnement de synthèse est donc la clef de toute application d'acoustique virtuelle car c'est elle qui permettra de rendre audible le champ sonore synthétique. Enfin, il faut noter que ces simulations acoustiques s'adressent principalement à des auditeurs humains. De nombreux facteurs perceptifs pourront donc intervenir au cours des simulations pour en conserver la qualité tout en diminuant leur complexité algorithmique. Nous allons nous intéresser à présent aux méthodes classiques de simulation mises en œuvre en acoustique pour la prédiction de réponse impulsionnelle et comment elles peuvent être utilisées dans le contexte de réalité virtuelle auditive.

Deuxième partie

Simulation acoustique. Acoustique virtuelle

Chapitre 3

Modèles et simulations acoustiques

Dans ce chapitre, nous présentons une vue d'ensemble des différents modèles et techniques utilisés pour la simulation acoustique. Nous insisterons plus particulièrement sur les modèles dits géométriques qui sont bien adaptés aux simulations de réponse impulsionnelle pour l'acoustique virtuelle. Nous présenterons également quelques modèles de diffraction, dont la théorie de la diffraction de Fresnel-Kirchhoff que nous utiliserons au Chapitre 6. Nous évoquerons en dernier lieu les techniques de simulation par échanges radiatifs en acoustique ainsi que les techniques statistiques, permettant de simuler efficacement les réflexions diffuses et la réverbération tardive.

3.1 Solutions analytiques à l'équation d'onde

La première solution au problème de calcul de réponse impulsionnelle est bien sur la recherche d'une solution analytique à l'équation d'onde gouvernant le comportement de la pression acoustique. On rappelle que dans le cas général en présence de parois, la pression acoustique est solution d'un système formé de l'équation d'onde à laquelle on ajoute des conditions aux limites (*cf.* paragraphe 1.6) :

$$\begin{cases} \Delta p - \dfrac{1}{c^2}\dfrac{\partial^2 p}{\partial t^2} = Q_0 & \text{dans le volume,} \\ Z\dfrac{\partial p}{\partial n} + i\omega\rho_0 p = 0 & \text{sur les parois.} \end{cases} \tag{3.1}$$

La plupart des solutions analytiques à ce système sont valables en régime continu et sont obtenues par des développements en série [Fil94]. Nous ne les détaillerons pas ici.

3.1.1 Acoustique modale

La théorie de l'acoustique modale permet d'expliciter n'importe quelle solution du système d'équations 3.1 (donc la réponse impulsionnelle de n'importe quel canal acoustique dans l'environnement) à partir d'une base de fonctions, solutions du système homogène associé : les *modes de résonnance* (ou modes propres) de l'environnement.

On peut montrer que ces solutions correspondent à différentes valeurs du nombre d'onde $\hat{k}_n$ (en général complexes) appelés nombres d'ondes propres du système. À ces nombres $\hat{k}_n$ correspondent les *fréquences de résonnance* (ou fréquences propres) f_n du système.

On peut donc définir la réponse impulsionnelle d'un canal acoustique quelconque dans l'environnement par :

$$h(t) = \sum_n \hat{A}_n e^{\hat{s}_n t},$$

où seuls les coefficients $\hat{A}_j$ dépendent de la position de la source et du récepteur. Les coefficients $\hat{s}_j = 2\pi i f_j - \sigma_j = ic\hat{k}_j$, appelés *pôles* de la solution, dépendent uniquement de l'environnement (σ_j est un coefficient d'amortissement associé au mode de résonnance considéré) [Kut91b]. La réponse impulsionnelle s'exprime donc comme une somme (pondérée) de résonnateurs amortis.

Dans le cas d'une salle parallélépipédique (de dimensions L_x, L_y, L_z) avec des *conditions aux limites de Neumann* caractérisant des parois parfaitement rigides :

$$\frac{\partial p}{\partial n} = 0 \text{ sur les parois,}$$

le problème est séparable selon les trois coordonnées et on obtient les fréquences de résonnance classiques (en coordonnées cartésiennes) [Kut91b] :

$$f_{n_x n_y n_z} = \frac{c}{2}\sqrt{\left(\frac{n_x}{L_x}\right)^2 + \left(\frac{n_y}{L_y}\right)^2 + \left(\frac{n_z}{L_z}\right)^2},\ (n_x, n_y, n_z) \in \mathbb{N}^3.$$

Les modes propres, représentant une onde stationnaire tridimensionnelle, sont alors de la forme :

$$p_{n_x n_y n_z}(x,y,z) = C \cos\left(\frac{n_x \pi x}{L_x}\right) \cos\left(\frac{n_y \pi y}{L_y}\right) \cos\left(\frac{n_z \pi z}{L_z}\right),$$

où C est une constante arbitraire.

On peut évaluer dans ce cas le nombre de modes inférieurs à une fréquence donnée f, qui varie en $O(f^3)$ [Kut91b], ce qui implique de calculer de l'ordre de 10^9 modes pour couvrir l'ensemble des fréquences audibles ($f \leq 20$ kHz) !!! Si cette méthode permet d'obtenir une solution exacte de l'équation d'onde pour une fréquence fixée, elle reste donc inapplicable en pratique pour le calcul de réponse impulsionnelle (sauf à basse fréquence dans des environnements très simples où elle est parfois utilisée []).

3.1.2 Solutions intégrales. Théorie de la diffraction de Fresnel-Kirchhoff

Les solutions intégrales à l'équation d'onde avec conditions aux limites sont fondées sur une «projection» du problème tridimensionnel en un problème bidimensionnel. On peut en effet montrer que le champ acoustique en un point quelconque M peut être exprimé en fonction du champ acoustique et de la vitesse particulaire sur une surface fermée arbitraire S entourant M [Kut91b, Fil94] (en pratique, le choix d'une surface d'intégration judicieuse permet de ramener le problème à une intégration sur les parois de l'environnement) ; c'est le théorème intégral d'Helmholtz-Kirchhoff (*cf.* Figure 3.1 (a)) :

$$\hat{P}(M) = \hat{P}_0(M) + \left[\iint_S \hat{P}(U)\nabla\left(-\frac{e^{ikr}}{4\pi r}\right) \cdot \mathbf{dS} - \iint_S -\frac{e^{ikr}}{4\pi r}\nabla\hat{\mathbf{P}}(U) \cdot \mathbf{dS}\right], \tag{3.2}$$

où $\mathbf{dS} = \mathbf{n}dS$ ($\mathbf{n}$ unitaire), $\hat{P}_0$ est la pression de la source «en espace indéfini» (solution particulière de l'équation d'Helmholtz avec second membre) et $-\frac{e^{ikr}}{4\pi r}$ est la fonction de Green dite *de l'espace libre* représentant la propagation d'une onde sphérique en champ libre à partir d'un point arbitraire.
L'intégrale de surface représente une solution de l'équation homogène associée satisfaisant les conditions aux limites, qu'on appelle parfois *champ diffracté*.

Nous ne détaillerons pas ici les méthodes permettant d'obtenir des solutions analytiques de cette équation intégrale (des détails peuvent être trouvés dans [Fil94]) qui utilisent des développements asymptotiques ou des développements en série. Ce théorème est également à la base des approches de résolution numériques par éléments finis de frontière que nous évoquerons au paragraphe 3.2.

Une solution approchée aux problèmes de propagation en présence d'obstacles peut être également obtenue avec la théorie de Fresnel-Kirchhoff qui dérive également d'une formulation intégrale de l'équation d'onde.

Théorie de Fresnel-Kirchhoff

La théorie de Fresnel-Kirchhoff développée en optique est fondée sur une combinaison du principe de Huygens et de la théorie des interférences apportée par Fresnel : chaque point non obstrué d'un front d'onde peut être considéré à un instant donné comme une source de fronts d'ondes secondaires de même fréquence. L'amplitude du champ *optique* en n'importe quel point au delà du front d'onde est égale à la somme de toutes les contributions de ces ondes secondaires *en tenant compte de leurs amplitudes et phases relatives*.

Kirchhoff a, par la suite, développé une théorie plus rigoureuse utilisant directement une formulation intégrale de l'équation d'onde permettant, entre autre, d'exprimer la contribution élémentaire en tant que source secondaire d'un élément différentiel de surface du front d'onde.

Considérons une source ponctuelle harmonique (en régime stationnaire). La pression en champ libre au point de réception situé à la distance ρ de la source est donnée par :

$$\hat{P}(t,\rho) = \frac{P_o}{\rho} e^{i(k\rho - \omega t)} = \hat{P}_u(\rho) e^{-i\omega t},$$

où P_o est l'amplitude de la pression émise par la source.

La résolution de l'équation d'Helmholtz homogène à l'aide du théorème de Green, associée à un choix judicieux de la surface d'intégration [Hec87], permet alors d'exprimer la contribution en pression acoustique d'un élément différentiel de surface dS du front d'onde à la pression totale en champ libre[1]resnel avait précédemment obtenu une approximation de cette expression en utilisant la décomposition de l'espace en *zones de Fresnel* :

$$d\hat{P}(M) = -\frac{P_o e^{ik(\rho+r)}}{4\pi\rho r} \left(ik(1 + \mathbf{v} \cdot \mathbf{r}) - \frac{1}{\rho} - \frac{\mathbf{v} \cdot \mathbf{r}}{r} \right), \tag{3.3}$$

où $\mathbf{r}$ et $\mathbf{v}$ sont des vecteurs unitaires (voir également les notations de la Figure 3.1 (b)).

Des détails concernant cette expression sont donnés dans l'Annexe A.4 de ce document.

Cette expression a été utilisée pour calculer le champ diffracté par un écran mince par exemple (en intégrant le terme 3.3 sur la partie du plan de l'écran non obstruée[2]) [Hec87, Fil94]. Elle est fondée sur des approximations valables dans le cas de l'optique (seule la portion éclairée de l'écran est prise en compte et les termes d'ordre deux en $1/r$ et $1/\rho$ sont négligés dans le terme 3.3)) mais qui conduisent à un résultat qui n'est toutefois pas solution de l'équation d'Helmholtz. Néanmoins, ces approximations ont également été utilisées avec succès en acoustique dans le cas de hautes fréquences.

1. F
2. on peut également intégrer ce terme sur la portion de plan obstruée, ce qui conduit au même résultat (théorème de Babinet) [Hec87].

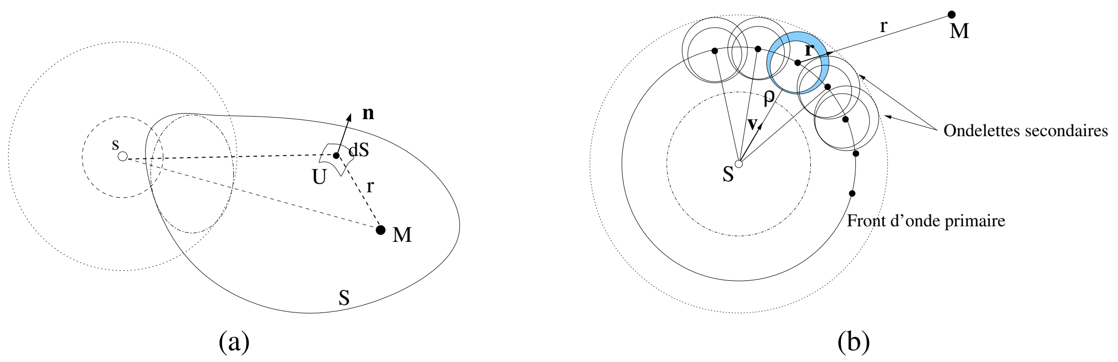


Figure 3.1 - *(a) Le théorème intégral de d'Helmholtz-Kirchhoff : utilisation du théorème de Green pour résoudre l'équation d'Helmholtz. (b) Le principe d'Huygens-Fresnel : une onde sonore peut être représentée comme une somme d'ondelettes secondaires de même fréquence en tenant compte de leurs amplitudes et de leurs phases relatives.*

3.2 Éléments finis. Éléments finis de frontière

Les techniques d'éléments finis utilisent une résolution numérique de l'équation d'onde par subdivision des quatre dimensions spatiales et temporelles en «éléments» [Wri95]. L'équation d'onde est alors exprimée de manière discrète pour chacun de ces éléments, ramenant le problème à la résolution d'un système linéaire. Généralement, à cause du coût en mémoire et en temps de calcul impliqué, ces techniques sont utilisés en régime continu, où la dimension temporelle est omise.

Éléments finis de frontière

Dans le cas des éléments finis de frontières, on résout numériquement la forme intégrale de l'équation d'onde par maillage des surfaces de l'environnement uniquement, ce qui permet d'alléger le coût des calculs [BdV98, KL95, Klu91, HCWH91]. La pression et la vitesse particulaire étant liées sur les parois, certaines approches, dites *variationnelles*, choisissent de calculer la vitesse particulaire [Jea98].

La résolution est effectuée en supposant que la pression est une combinaison linéaire d'un nombre fini de fonctions de base. On cherche alors les coefficients de combinaison donnant la «meilleure» solution. Plusieurs méthodes permettent soit d'imposer que l'équation soit vérifiée exactement en certains points fixés (méthode de collocation) ou assurer un critère global de minimisation (méthode de Galerkin) [Fil94].

L'un des problèmes de ces approches est qu'il faut utiliser un maillage dont la taille est de l'ordre de la longueur d'onde traitée pour avoir des interférences correctes. Ceci tend à limiter leur utilisation dans le cas des hautes fréquences.

Les méthodes d'éléments finis restent les méthodes de référence pour la résolution de problèmes de propagation d'ondes acoustiques, puisque elles permettent de modéliser de manière exacte tous les phénomènes de propagation. Néanmoins, elles demeurent très coûteuses en temps de calcul et en espace mémoire, et deviennent inutilisables dès que les environnements traités deviennent complexes ou qu'elles sont utilisées pour des calculs en régime impulsionnel [BdV98].

3.3 Théorie des rayons sonores : acoustique géométrique

Comme nous l'avons vu, appliquer les principes de la théorie des ondes devient toutefois très rapidement impraticable quand la complexité de l'environnement augmente ou que les calculs ne sont pas réalisés en régime établi. De plus, en supposant que de tels calculs soient réalisables, ils apporteraient bien plus d'information que nécessaire d'un point de vue perceptif. Si l'on se place dans le cas limite des hautes fréquences, on peut utiliser un moyen de description beaucoup plus simple des phénomènes de propagation. Comme en optique, on se base sur le concept de rayons le long desquels le son va se propager de manière rectiligne. Le champ sonore en un point est alors obtenu par sommation des contributions de tous les rayons (direct, réfléchis, réfractés) incidents à ce point. On supposera également que la visibilité est géométrique c'est-à-dire qu'un rayon intersectant un obstacle ne contribuera pas au champ sonore.

3.3.1 Traitement géométrique de la réflexion

Nous présentons ici un aperçu des diverses techniques fondées sur la théorie des rayons sonores permettant de traiter le problème des réflexions du son, et donc de calculer de manière approchée la réponse impulsionnelle d'un milieu réverbérant.

Sources-images

Les méthodes de sources-images [AB79], dérivent d'une solution exacte particulière à l'équation d'onde avec conditions aux limites de Neumann dans le cas d'un réflecteur plan. En effet, si l'on suppose le réflecteur infini, on peut alors montrer que le champ créé au point de réception est la superposition du champ créé par la source sonore réelle et de celui d'une source virtuelle, image miroir de la source réelle par rapport au plan de réflexion. Cette source virtuelle est appelée «*source-image*» (*cf.* Figure 3.2 (a)).

Ce raisonnement a été étendu [Bor84, KOK93, Hei93], dans le cadre de l'acoustique géométrique, à des réflecteurs multiples et aux réflexions d'ordres supérieurs en créant des images miroir des sources virtuelles. Cela constitue alors une solution approchée dans le cas général [Fil94]. La pression sonore peut alors s'exprimer comme :

$$P(M,t) = P_0 \left[\frac{\delta\left(t - \frac{r}{c}\right)}{r} + \sum_{i=1}^{+\infty} \left(\prod_{k=1}^{n} (1 - \alpha_k^{\lambda}) \right) \frac{\delta\left(t - \frac{r_i}{c}\right)}{r_i} \right], \tag{3.4}$$

où r est la distance de la source au point M, r_i est la distance de la source-image i (d'ordre n) à M, α_k^{λ} est le coefficient d'absorption du matériau correspondant à la $k^{ème}$ réflexion et P_0 est l'amplitude de la pression acoustique de la source.

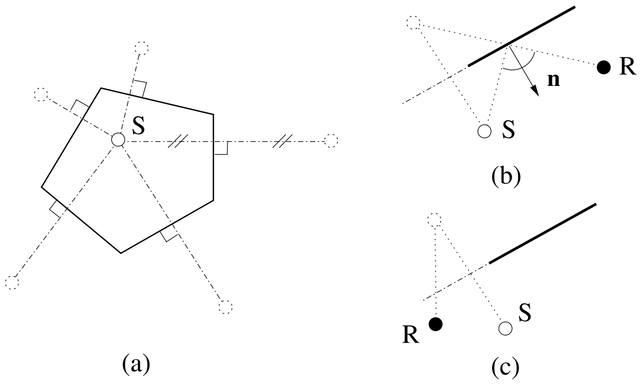


Figure 3.2 - *Méthode des sources-images en 2D. La Figure (a) montre une source sonore (S) et ses sources images de premier ordre pour un contour pentagonal. La Figure (b) représente une source image valide pour une position de récepteur (R) ; la Figure (c) représente une configuration invalide car le chemin réfléchi entre la source virtuelle et le récepteur n'intersecte pas le réflecteur.*

Les sources-images peuvent être très facilement construites. Leur principal inconvénient est la croissance exponentielle du nombre de sources en fonction de l'ordre de réflexion. En effet, pour N surfaces ($N \geq 1$), on a $N(N-1)^{i-1}$ sources images d'ordre i ($i \geq 1$). Le nombre total de sources images jusqu'à l'ordre i (inclus) est donc $\frac{N(N-1)^i-1}{N-2}$ (pour $N > 2$).

Si l'on suppose les plans de réflexion de taille finie, toutes les sources images ne contribuent pas au champ au point de réception, ce qui implique un test de validité géométrique supplémentaire pour chaque source (*cf.* Figures 3.2 (b) et (c)). On ne considère alors plus les sources images invalides à un ordre de réflexion donné pour le calcul des sources images valides à un ordre supérieur.

Un test de visibilité est également nécessaire dans le cas où des obstacles sont présents le long du chemin de propagation. En général, on utilisera un test géométrique en lançant un rayon depuis la source-image jusqu'au récepteur. Il faut bien noter ici que, par contre, les sources-images "invisibles" (depuis le récepteur) à un ordre donné peuvent donner naissance à des source-images *valides et visibles* à un ordre supérieur. On ne peut donc pas les éliminer du processus.

Ceci contribue à rendre la méthode assez rapidement inutilisable telle qu'elle dans le cadre d'environnements complexes et d'ordres de réflexion élevés. Toutefois, la méthode des sources-images, bien que limitée à des réflexions spéculaires sur des surfaces planes, permet dans le cadre de l'acoustique géométrique d'obtenir la réponse impulsionnelle «exacte» de l'environnement (dans la limite des hypothèses effectuées et de l'ordre de réflexion atteint bien sûr). De plus, la méthode reste très simple à implémenter, bien adaptée à des environnements interactifs et le traitement à effectuer pour chaque source-image est générique, permettant une implémentation parallèle, comme nous le verrons au Chapitre 7.

Lancer de rayons. Lancer de cônes

La deuxième grande famille d'approches pour résoudre le problème des réflexions du son sont les techniques de lancer de rayons ou de cônes [Dal96, Eme95, Lew93, MvMV93, Nay93, SK95, vMM93, Vor89]. Les deux approches, très similaires dans le principe, sont probablement les plus répandues pour le calcul des réflexions spéculaires. Nous verrons toutefois qu'elles peuvent être étendues pour prendre en compte de modèles de réflexion plus généraux.

Dans les deux cas (rayons et cônes), un nombre donné de rayons (resp. cônes) est émis dans toutes les directions depuis la source. Les rayons sont propagés dans la scène et un rayon réfléchi est construit suivant la loi de Snell-Descartes chaque fois qu'un rayon intersecte une paroi (*cf.* Figure 1.6 (a)).

Lorsque le rayon intersecte un volume de réception (*resp.* le cône contient le point de réception), sa contribution (son intensité acoustique par exemple) est stockée dans un échogramme (ou *réflectogramme*) sous la forme d'une impulsion de Dirac en fonction de son temps d'arrivée (cet échogramme est donc ici équivalent à une réponse impulsionnelle en intensité). Le rayon est ensuite prolongé. Le processus continue jusqu'á ce que l'énergie portée par le rayon soit trop faible, que l'on dépasse un certain ordre de réflexion, ou bien que l'écartement entre deux rayons soit trop grand (dans ce cas on rate le récepteur). Dans la plupart des cas, l'atténuation lors de la réflexion dépend de la fréquence. Différents échogrammes sont donc construits pour différentes bandes de fréquences (nous verrons dans le paragraphe 4.1.3 comment la réponse impulsionnelle en pression est reconstruite à partir de ces échogrammes).

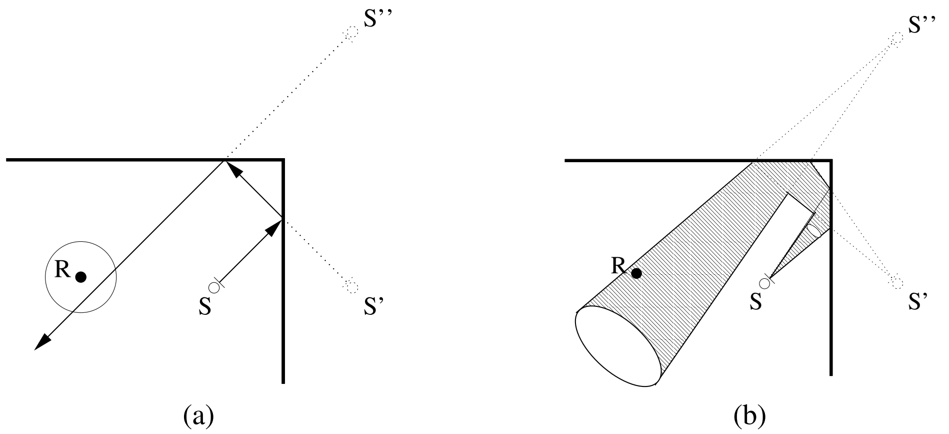


Figure 3.3 - *(a) Calcul des sources-images par lancer de rayons : Des rayons sont tirés depuis la source et réfléchis spéculairement jusqu'à ce qu'ils atteignent un volume de réception où leurs contributions sont comptabilisées. Les rayons sont ensuite prolongés jusqu'á ce que leur énergie soit trop faible. (b) Par lancer de cônes : des cônes sont tirés depuis la source et réfléchis jusqu'à qu'ils contiennent le récepteur.*

Connaissant l'historique des réflexions, on peut également séparer l'aspect propagatif de l'aspect énergétique et une fois l'arbre des chemins réfléchis construit on peut modifier les paramètres des sources et des matériaux et recalculer la réponse de manière efficace [MOD96]. Dans ce cas, le calcul de la réponse impulsionnelle peut également être effectué de manière plus précise (nous y reviendrons dans le paragraphe 4.1.3) [MvMV93].

Les deux méthodes de lancer de rayons ou de cônes sont équivalentes dans le principe à un calcul de sources-images mais permettent de déterminer les sources-images valides contribuant au champ sonore au point de réception en évitant une construction systématique de toutes les images et donc un coût exponentiel. Toutefois, le nombre de rayons (*resp.* cônes) étant fixé au départ, seul un nombre limité de sources-images pourra être identifié.

La principale limitation de ces approches est le problème d'échantillonnage spatial [Lew93, MOD96, Leh93] lié à la discrétisation de l'espace des directions depuis la source et dans le cas du lancer de cône, à la non subdivision des cônes dans le cas où ceux-ci intersectent plusieurs surfaces (généralement le cône est assimilé à son axe et les intersections sont calculées comme dans le cas d'un lancer de rayon). Ces problèmes peuvent conduire à la prise en compte de «fausses» réflexions ou au contraire à l'oubli de réflexions effectives (*cf.* Figure 3.4). Il est donc nécessaire de recourir à des tests de rétropropagation depuis le récepteur pour vérifier que le chemin est effectivement valide et de sur-échantillonner l'espace des directions au départ de la source.

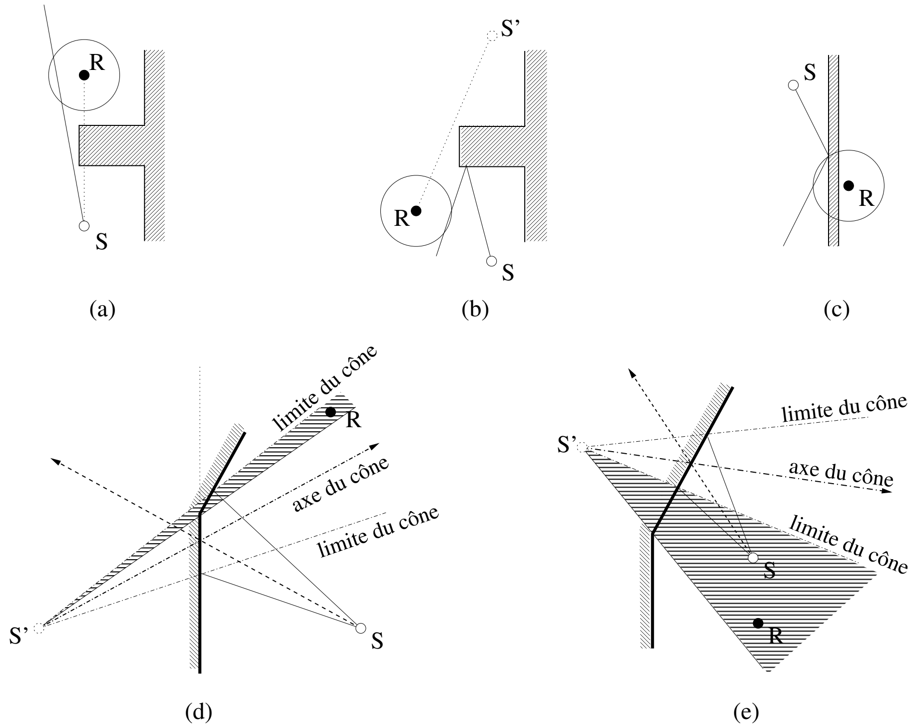


Figure 3.4 - *Problèmes d'échantillonnage spatial dans les méthodes de lancer de rayons (a,b,c) et de lancer de cônes (d,e)*[2]. *(a) Contribution prise en compte alors que le récepteur n'est pas géométriquement visible depuis la source. (b) Réflexion prise en compte alors que la source-image n'est pas valide. (c) Réflexion prise en compte alors que le récepteur est de l'autre côté de la paroi. (d) Réflexion prise en compte alors que la source-image n'est pas valide. (e) Réflexion non prise en compte alors que la source-image est valide.*

Néanmoins, ces méthodes sont très efficaces pour le calcul des réflexions spéculaires et donnent de très bon résultats au moins pour les premiers ordres de réflexion (inférieurs à 5). De plus elles ne sont pas limitées aux surfaces polygonales. Toutefois, pour de grands ordres de réflexions, de nombreuses contributions sont «oubliées» ; le nombre de rayons (ou cônes) utilisés doit donc être augmenté au détriment de l'efficacité du processus.

Lancer de faisceaux adaptatif

Une réponse au problème d'échantillonnage spatial est apporté par les méthodes de lancer de faisceaux adaptatifs très proches du lancer de cônes mais dans lesquelles les cônes (appelés ici faisceaux) sont supportés par les polygones de la scène [MOD96, FCE$^+$98, HH84]. Ici, chaque polygone devient une «fenêtre» que l'on va utiliser pour limiter le faisceau réfléchi (*cf.* Figure 3.5). De plus, un nouveau faisceau est généré pour chaque portion de surface intersectant le faisceau incident. Le nombre de faisceaux dépend donc de la géométrie et n'est pas limité arbitrairement au début du processus, garantissant la possibilité d'obtenir toutes les sources-images. Le nombre de faisceaux générés est également «optimal» en ce sens que les faisceaux couvrent exactement la portion de l'espace pour laquelle la source-image correspondante est valide. Enfin, les erreurs de détection sont entièrement éliminées.

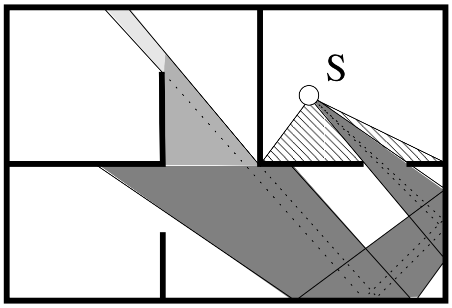


Figure 3.5 - *Réflexion et fenêtrage des faisceaux générés depuis la source (S)*[4].

Cette méthode est donc la méthode la plus efficace pour déterminer de manière exacte l'ensemble des sources-images valides. La principale difficulté reste néanmoins de réaliser les opérations géométriques d'intersection, réflexion et fenêtrage («clipping») des faisceaux par les polygones de la scène (et les éventuels obstacles) de manière efficace. En outre, la méthode est *a priori* limitée aux surfaces polygonales.

Méthodes de Monte-Carlo

Les méthodes dites «de Monte-Carlo» sont également du type lancer de rayons, mais dans ce cas, à chaque réflexion, plusieurs nouveaux chemins sont générés, leur direction étant choisie aléatoirement en utilisant une loi de probabilité correspondant à la distribution de la radiance émise par les surfaces [Ste89, Nay93, Vor89, VG97]. L'avantage de ces méthodes est de pouvoir prendre en compte des modèles de réflexion arbitrairement complexes mais généralement limités dans le cadre de l'acoustique à une composante spéculaire et une composante lambertienne. Néanmoins, elles souffrent des mêmes artefacts que nous avons décrit précédemment. En outre, elles nécessitent d'utiliser un très grand nombre de rayons, en particulier dans le cas de parois diffusantes et sont donc rapidement limitées en efficacité. Enfin, ces approches comptabilisent les contributions des rayons sous forme énergétique (intensités) et intégrées sur des petits intervalles de temps constants[3] . En effet, le nombre de contributions atteignant le récepteur devient très vite ingérable (cas des réflexions diffuses tout particulièrement). Elles négligent donc de ce fait les phénomènes d'interférence. On notera ici que certaines approches hybrides combinent sources-images pour les réflexions spéculaires et lancer de rayons stochastique pour les réflexions diffuses [Vor89].

3.3.2 Théorie géométrique de la diffraction

Les hypothèses de déplacement rectiligne de l'onde sonore et de «visibilité» géométrique ne permettent pas a priori de simuler de manière satisfaisante les phénomènes de propagation complexes des ondes en présence d'obstacles. C'est la raison pour laquelle la diffraction est généralement négligée dans les approches géométriques, dans lesquelles on va supposer qu'un point qui n'est pas «visible» depuis une source ne reçoit pas d'énergie. Toutefois, en 1962, Keller développe une extension à la théorie des rayons en optique géométrique afin de traiter les phénomènes de diffraction des ondes électromagnétiques [Kel62]. Pour cela, il introduit de nouveaux rayons diffractés (*cf.* Figure 3.6) et

2. d'après [Leh93] et [Lew93].
3. petits devant le temps de réverbération. De l'ordre de 5 ou 10 ms d'après [Kut93].

exprime le champ diffracté comme une somme de la contribution de chaque rayon, de manière semblable à ce qui était fait pour les autres phénomènes de propagation (réflexion, réfraction). Toutefois, calculer l'énergie d'un rayon diffracté impose de connaître un *coefficient de diffraction* dont la théorie géométrique seule ne permet pas de calculer la valeur. On ne peut l'obtenir que par comparaison avec des situations canoniques pour lesquelles on dispose d'une solution analytique et procéder par identification des termes. Ces situations de référence sont bien entendu assez rares et c'est l'un des principaux inconvénients de la théorie. Le champ diffracté au point M peut être alors exprimée par :

$$\hat{P}_{diff}(M) = \sum_{i=1}^{N} \hat{D}_i \frac{e^{ik\rho_i}}{\rho_i} \frac{e^{ikr_i}}{\sqrt{r_i}},$$

où $\hat{D}_i$ est le coefficient de diffraction, ρ_i est la distance de la source au $i^{\text{ème}}$ point de diffraction et r_i est la distance du point de diffraction au point de réception M.

Deux types de situations sont rencontrées en pratique : si l'obstacle comporte des arêtes ou des coins, on considère qu'ils sont à l'origine des rayons diffractés ; si la surface de l'obstacle est régulière, la diffraction est due à une onde de surface donnant naissance aux rayons diffractés [Kel62, KP74, LS97].

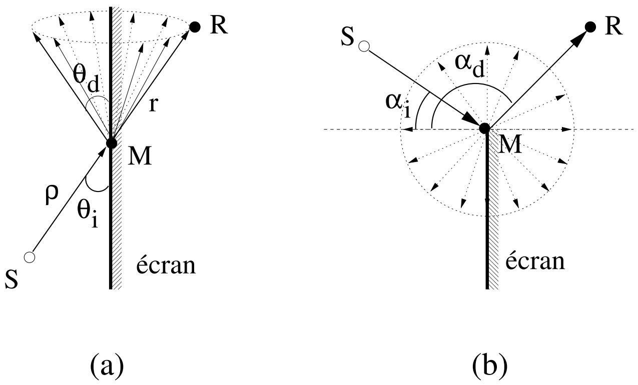


Figure 3.6 - *Rayons diffractés par une simple arête. (a) Vue de dessus. (b) Vue de coté. En incidence oblique (*cf. *vue de dessus), un rayon incident sur l'arête donne naissance à un cône de rayons diffractés tel que l'angle d'ouverture du cône θ_d soit égal à l'angle d'incidence du rayon avec le plan de l'arête θ_i. En incidence normale (*cf. *vue de profil), un rayon incident donne naissance à des rayons diffractés coplanaires omnidirectionnels dans le plan orthogonal au plan de l'arête. Le coefficient de diffraction est donné dans le cas d'ondes électromagnétiques par $D = -\frac{e^{i\frac{\pi}{4}}}{2(2k\pi)^{1/2}} \sin\theta_i \left[\sec\left(\frac{1}{2}(\alpha_d - \alpha_i)\right) \pm \csc\left(\frac{1}{2}(\alpha_d + \alpha_i)\right)\right]$ (*sec *et* csc *désignent respectivement les fonctions* $1/\sin$ *et* $1/\cos$*, le signe* $\pm$ *dépend de la polarisation de l'onde).*

La théorie géométrique de la diffraction nécessite donc de déterminer les zones diffractantes ce qui peut être délicat [FCE+98] et de lancer de nouveaux rayons, augmentant de fait la complexité de l'algorithme. De plus, les coefficients de diffraction sont rapidement difficiles à calculer.

3.3.3 Validité et limitations des modèles géométriques

Les modèles que nous avons présentés dans cette section sont tous fondés sur les hypothèses de l'optique géométrique et sont donc valables dans le cas des hautes fréquences acoustiques (longueurs

4. d'après [FCE+98].

d'onde centimétriques) contrairement aux modèles ondulatoires qui eux sont valides quel que soit la fréquence considérée. Malgré l'existence de modèles géométriques de la diffraction, celle-ci n'est pratiquement jamais prise en compte. De plus, comme nous l'avons vu, les modèles géométriques deviennent assez vite coûteux en calcul lorsque l'ordre de réflexion augmente. La validité de leur utilisation voire leur utilité du point de vue perceptif pour la simulation de grands ordres de réflexion sont également discutables [Kut91b], du moins dans le cadre d'environnements très réverbérants. Ils sont également limités à des modèles de réflexion simple, purement spéculaire en général. Pourtant, le phénomène de réflexion diffuse est également essentiel à considérer. D'une part, à haute fréquence (le cadre de validité des modèles géométriques) les surfaces ne peuvent plus toujours être considérées comme parfaitement planes vis à vis de la longueur d'onde. D'autre part, étant donnée la complexité d'un processus de simulation acoustique, il est très fréquent que de larges parties de la scène utilisée (les sièges dans le cas d'une salle de concert) soit remplacées par de simples surfaces planes mais munies de propriétés diffusantes. La complexité géométrique est alors transférée au niveau radiométrique. Cet genre d'approche est également très utilisé en synthèse d'image («clustering» en radiosité, visualisation par «voxels» []).

C'est pourquoi sont apparues deux alternatives permettant de traiter de grands ordres de réflexion de manière plus efficace ainsi que les réflexions diffuses : les méthodes d'échanges radiatifs et les méthodes fondées sur l'acoustique statistique.

3.4 Échanges radiatifs

Développées dans les années 50 pour la simulation de transfert thermique, ces approches géométriques particulières, expriment un bilan des échanges énergétiques entre surfaces. Elles ont ensuite été utilisées dans le milieu des années 70 pour la simulation d'échanges diffus en acoustique [Kut91b] mais également en synthèse d'image dès le début des années 80 [GTGB84]. Elles ont d'ailleurs, dans ce dernier cadre, subi de nombreuses améliorations (nous présenterons en détail les plus importantes au Chapitre 8). Leur efficacité à simuler des réflexions diffuses leur ouvre deux possibilités d'application en acoustique que nous détaillerons : simuler des réflexions diffuses ou bien simuler des échanges en champ diffus. Nous présentons ici quelques bases de ces techniques, de plus amples détails pouvant être trouvés dans les ouvrages [Kut91b, CW93, SP94].

3.4.1 L'équation de radiance

La grandeur physique pertinente pour établir un bilan énergétique est la *radiance* notée L. La radiance représente la quantité d'énergie émise (resp. reçue) en un point dans une direction donnée, par unité de temps, par unité d'aire perpendiculaire à la direction de propagation, par unité d'angle solide. L'expression générale de l'équilibre des échanges d'énergie pour un ensemble de surfaces peut s'écrire comme (*cf.* Figure 3.7) :

$$\underbrace{L^{\lambda}(x,t,\theta_0,\phi_0)}_{\text{radiance totale}} = \underbrace{L_e^{\lambda}(x,t,\theta_0,\phi_0)}_{\text{radiance émise}} + \underbrace{\int_{\Omega} \rho_{bd}^{\lambda}(x,t,\theta_0,\phi_0,\theta,\phi) L_i^{\lambda}(x,t,\theta,\phi) cos\theta d\omega}_{\text{radiance réfléchie}}, \qquad (3.5)$$

où

$L^{\lambda}(x,t,\theta_0,\phi_0)$	est la radiance totale quittant le point x à l'instant t, dans la direction (θ_0,ϕ_0),
$L_e^{\lambda}(x,t,\theta_0,\phi_0)$	est la radiance émise par x à t dans la direction (θ_0,ϕ_0), propriété du point x,
$\rho_{bd}^{\lambda}(x,t,\theta_0,\phi_0,\theta,\phi)$	est la Fonction de Distribution de la Réflectance Bidirectionnelle, caractérisant les propriétés de réflexion de la surface au point x,
Ω	est l'ensemble des directions (θ,ϕ) dans l'hémisphère supérieur au point x,
$L_i^{\lambda}(x,t,\theta,\phi)$	est la radiance incidente en x à t depuis la direction (θ,ϕ).

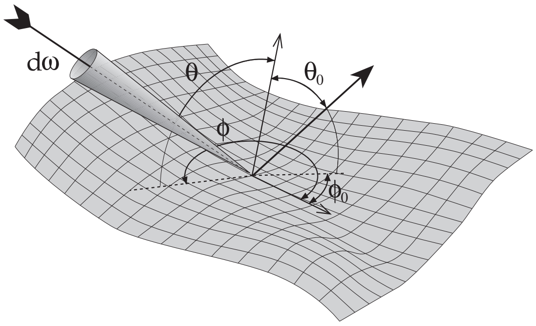


Figure 3.7 - *Notations pour l'équation de radiance*[3].

Nous allons supposer ici que les surfaces réfléchissantes de la scène sont lambertiennes et que la fonction de réflectance ne dépend pas du temps. On a donc

$$\rho_{bd}^{\lambda}(x,t,\theta_0,\phi_0,\theta,\phi) = \frac{\gamma^{\lambda}(x)}{\pi}.$$

d'où :

$$L^{\lambda}(x,t) = L_e^{\lambda}(x,t) + \frac{\gamma^{\lambda}(x)}{\pi}\int_{\Omega} L_i^{\lambda}(x,t,\theta,\phi) cos\theta d\omega,$$

où

$L^{\lambda}(x,t)$	est la radiance totale quittant le point x à l'instant t,
$L_e^{\lambda}(x,t)$	est la radiance émise par x à t, propriété du point x,
$\gamma^{\lambda}(x)$	est la réflectance diffuse,
Ω	est l'ensemble des directions (θ,ϕ) dans l'hémisphère supérieur au point x,
$L_i^{\lambda}(x,t,\theta,\phi)$	est la radiance incidente en x à t depuis la direction (θ,ϕ).

Exprimons également la puissance par unité de surface quittant l'élément de surface dx au temps t :

$$\frac{dP}{dx} = \int_{\Omega} \underbrace{L^{\lambda}(x,t,\theta,\phi)\cos\theta d\omega}_{I^{\lambda}(x,t,\theta,\phi)} = I^{\lambda}(x,t), \qquad (3.6)$$

où Ω est l'hémisphère supérieur à la face et $I^{\lambda}(x,t,\theta,\phi)$ est l'intensité acoustique quittant la face dans la direction (θ,ϕ). On remarque que l'intensité totale quittant la face dans l'hémisphère supérieur est donc l'équivalent de la *radiosité* ou de l'*irradiance* (dans le cas d'énergie incidente) qui sont les dénominations récentes utilisées en optique. On utilisera donc indifféremment les termes «intensité» et «radiosité» dans la suite du document.

3. Figure tirée de l'ouvrage «*Radiosity and Global Illumination*» par F.X. Sillion et C. Puech [SP94].

Dans le cas lambertien, l'intensité totale quittant la face $I^{\lambda}(x,t)$ est proportionnelle à la radiance $L^{\lambda}(x,t)$, alors indépendante de la direction (on peut donc sortir le terme de radiance de l'intégrale 3.6) :

$$L^{\lambda}(x,t) = \pi I^{\lambda}(x,t), \tag{3.7}$$

On peut alors utiliser l'intensité à la place de la radiance pour décrire les échanges énergétiques. On peut donc ré-écrire l'équation 3.5 en multipliant chaque terme par π et en exprimant l'intégrale sur les surfaces S de l'environnement (*cf.* Figure 3.8) :

$$I^{\lambda}(x,t) = Ie^{\lambda}(x,t) + \gamma^{\lambda}(x) \int_{y \in S} \kappa^{\lambda}(x,y) I^{\lambda}\left(y, t - \frac{r}{c}\right) dy, \tag{3.8}$$

où

$I^{\lambda}(x,t)$	est l'intensité totale quittant le point x à l'instant t,
$I_e^{\lambda}(x,t)$	est l'intensité émise par x à t, propriété du point x,
$\gamma^{\lambda}(x)$	est la réflectance diffuse,
S	est l'ensemble des surfaces de la scène,
r	est la distance du point x au point y,
c	est la célérité de l'onde sonore.

On a également (*cf.* Figure 3.8) :

$$\kappa^{\lambda}(x,y) = \frac{\cos\theta \cos\theta'}{\pi r^2} V(x,y) \mu^{\lambda}(r), \tag{3.9}$$

où $V(x,y)$ est la *fonction de visibilité* binaire entre x et y[4] et $\mu^{\lambda}(r)$ est l'atténuation atmosphérique sur la distance r (*cf.* Annexe A.2.2).

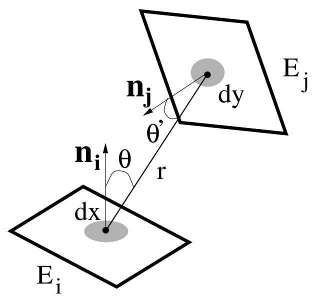


Figure 3.8 - *Notations pour les échanges radiatifs entre surfaces et le facteur de forme.*

Dans ce cas, comme dans le cas des solutions intégrales de l'équation d'onde (*cf.* paragraphe 3.1.2), on obtient une équation intégrale que l'on va résoudre par éléments finis d'une manière similaire aux éléments finis de frontière présentés au paragraphe 3.2. Il faut bien noter ici qu'à la différence des éléments finis de frontière, cette approche reste purement géométrique et énergétique. La solution obtenue n'est donc pas une solution de l'équation d'onde. On ne pourra obtenir au mieux qu'une approximation de la «réponse impulsionnelle énergétique» du canal acoustique simulé. (une répartition de l'intensité et non pas de la pression elle même au cours du temps puisque l'on ne conserve pas d'information sur la phase)

4. $V(x,y) = 1$ si et seulement si x est visible depuis y et $V(x,y) = 0$ sinon.

3.4.2 Résolution par éléments finis : technique des radiosités

Les surfaces de la scène sont donc subdivisées en facettes ou *éléments*. Dans la méthode originale, on suppose l'intensité *constante sur les éléments* (ce qui revient à choisir des fonctions de base constantes dans une formulation d'éléments finis). On assimile également les facettes à leur centre, ce qui implique un transfert «instantané» de facette à facette.

On obtient alors l'*équation de radiosité temporelle* qui caractérise l'énergie émise (ou reçue) par chaque facette i :

$$I_i^\lambda(t) = Ie_i^\lambda(t) + \gamma_i^\lambda \sum_{j=1}^{N} F_{ij}^\lambda I_j\,(t - T_{ij})\,, \tag{3.10}$$

où T_{ij} est le temps de propagation de centre à centre des éléments. Le terme F_{ij}^λ est appelée *facteur de forme* entre les éléments i et j. Il mesure la proportion d'énergie quittant i qui arrive sur j, qu'on supposera constante au cours de l'échange. Il est défini par :

$$F_{ij}^\lambda = \frac{1}{A_i}\int_{x\in E_i}\int_{y\in E_j}\frac{\cos\theta\cos\theta'}{\pi r^2}V(x,y)\mu^\lambda(r)\,dxdy, \tag{3.11}$$

où l'on a repris les notations de l'équation 3.9 (voir également la Figure 3.8). Cette quantité ne peut être évaluée analytiquement que dans le cas de polygones en situation de visibilité mutuelle totale. Dans la pratique, on effectuera un calcul approché de l'intégrale par échantillonnage des deux éléments [SP94].

Résolution du système. Complexité. Convergence

Supposons le problème complètement indépendant du temps (ondes lumineuses, ondes sonores en régime établi). Si l'on note $\mathbf{I} = [\mathbf{I_i}]$ le vecteur des intensités incidentes sur les éléments, $\mathbf{E} = [\mathbf{Ie_i}]$ le vecteur des intensités émises par les éléments et $\mathbf{F} = [\mathbf{F_{ij}}]$ la matrice des facteurs de forme, on peut exprimer le système formé des équations discrétisées 3.10 sous forme matricielle :

$$\mathbf{I} = \mathbf{E} + \mathbf{F}\mathbf{I}.$$

On peut résoudre le système :

$$\mathbf{I} = (\mathbf{Id} - \mathbf{F})^{-1}\mathbf{E} = \sum_{i=0}^{+\infty}\mathbf{F}^i\mathbf{E}.$$

Pour chaque paire d'éléments i et j, on va commencer par calculer la matrice $[F_{ij}^\lambda]$. Chaque terme implique un calcul de visibilité qui ne peut être fait qu'en $O(N^2\log m)$, où N est le nombre d'éléments et m le nombre d'obstacles. C'est la partie la plus coûteuse de la résolution. Ensuite, chaque itération nécessaire au calcul de la matrice $\mathbf{F}^i$ est en $O(N^2)$. Chaque itération représente une réflexion des ondes sur les surfaces de l'environnement. La convergence de la méthode provient ici du fait que toutes les réflectances des matériaux sont plus petites que 1 ainsi que la somme des facteurs de forme pour un élément donné (on ne crée donc pas d'énergie au moment de la réflexion). La plupart des algorithmes de radiosité (du moins dans le cadre de la synthèse d'image) ne calculent pas toute la matrice F lors de la résolution mais n'en évaluent qu'une colonne ou une ligne à la fois. Cela correspond respectivement à propager l'intensité au départ d'un seul élément vers tout les autres («shooting») ou au contraire de collecter l'intensité arrivant de tout les autres éléments sur un élément en particulier («gathering») [SP94]. Si l'on utilise la première approche en propageant à chaque étape l'intensité de l'élément qui en a le plus à ré-émettre, on accélère de façon significative la convergence du problème ; ces approches sont désignées sous le nom de *radiosité progressive* [SP94].

Dans le cas où l'on doit prendre en compte un temps de propagation, comme pour les ondes sonores en régime impulsionnel, les valeurs d'intensité dépendent du temps et ne sont donc plus scalaires. En pratique, dans les applications d'acoustique, on considère comme critère de convergence le moment où l'énergie à émettre a décru d'un facteur 1.10^{-6} (60 dB).

3.4.3 Application aux réflexions diffuses et au calcul de la réverbération tardive

Ces méthodes d'échanges radiatifs ont été appliquées en acoustique pour le calcul des réflexions diffuses et de la réverbération tardive [Kut91b, Lew93, Mal86, SZEG93, SK95]. Dans la plupart des cas, elles constituent une deuxième passe après une phase de lancer de rayons ou cônes. L'énergie des rayons incidents aux surfaces est stockée dans des échogrammes attachés aux éléments en fonction du temps d'arrivée. Là encore, comme dans les méthodes de lancer de rayon stochastique, on intègre l'énergie sur de petits pas de temps. Dans la deuxième phase, cette énergie est échangée entre tous les éléments par itérations successives en fonction des facteurs de forme. Un temps de propagation doit alors être associé à l'échange. Généralement on choisit le temps de transfert de centre à centre pour chaque paire d'éléments considérée. Lewers [Lew93] propose le tirage aléatoire d'un temps à chaque itération suivant une loi de probabilité qui dépend de la configuration géométrique entre les éléments.

Un point discutable de ces approches dans le cas de l'acoustique est leur aspect purement énergétique. En effet, si pour la réflexion tardive on peut supposer le champ diffus (*cf.* paragraphe 1.6.2), il n'en est pas de même pour les premières réflexions, où les phénomènes d'interférence peuvent ne pas être négligeables. En outre, dans le cadre de la modélisation de la réflexion diffuse, on peut imaginer que si la taille des éléments est supérieure à la longueur d'onde de la vibration considérée, des phénomènes d'«auto-interférences» vont apparaître pendant la réflexion, qui ne seront pas pris en compte. Un autre problème peut être lié au fait qu'on ne considère que des échanges «ponctuels» dans le temps alors que les surfaces sont de taille non nulle. Le fait de choisir un temps de propagation variable durant la simulation comme dans [Lew93] parait assez éloigné du phénomène physique sous-jacent. Enfin, lorsque l'on augmente le nombre d'éléments de surface utilisés, le coût en mémoire et en temps de calcul augmente très rapidement.

3.5 Modèles statistiques

Devant l'explosion rapide en complexité des approches géométriques, peu adaptées à simuler de grands ordres de réflexions dans des environnements complexes, des techniques statistiques ont été développées pour traiter la réverbération tardive [Hei93, MOD96]. Elles ont également été appliquées à la modélisation de réflexions diffuses, généralement en complément de méthodes géométriques traitant les premiers ordres de réflexions spéculaires [MvMV93].

3.5.1 Modélisation de la réverbération tardive

Les théories statistiques développées pour la modélisation de la réverbération tardive sont fondées sur la notion de champ diffus (*cf.* paragraphe 1.6.2). On peut alors montrer, dans le domaine fréquentiel comme dans le domaine temporel, que l'on peut assimiler la réponse impulsionnelle à un processus aléatoire gaussien. Cette modélisation n'est toutefois valable qu'à partir d'une certaine fréquence : la fréquence de Schroeder définie par $f_{Schroeder} = 2000\sqrt{T_r/V}$ (Hz) où T_r (s) est le temps de réverbération et V (m^3) est le volume de l'espace (supposé clos)[5]. On peut montrer que la représentation statistique n'est

5. comme dans la formule du temps de réverbération de Sabine, cette formule est empirique. On ne s'étonnera donc pas de l'inhomogénéité apparente des unités.

valable qu'au delà d'un certain temps : le temps de mélange qu'on peut approcher par $T_m = \sqrt{V}$ (ms). Il peut être considéré comme le temps nécessaire à l'établissement d'un champ «perceptiblement diffus». En général, il est inférieur à celui donné par l'approximation précédente. Différentes approches proposent alors de modéliser la partie de la réponse impulsionnelle correspondant à la réverbération tardive comme un bruit blanc gaussien (un processus aléatoire suivant une loi normale [PTVF92]) filtré et fenêtré temporellement [Moo79, MvMV93, MOD96, Hei93]. Ainsi Moorer [Moo79] propose de pondérer son bruit blanc avec une fonction exponentielle plus ou moins décroissante suivant la fréquence. Une autre approche plus récente [MvMV93], propose de filtrer plus finement un bruit blanc en fonction d'un filtre «moyen» dépendant de la source, du récepteur et de l'absorption des matériaux ; et de le fenêtrer avec une fonction dépendant de l'ordre de réflexion considéré.

3.5.2 Extension aux réflexions diffuses

Les approches statistiques ont également été utilisées pour modéliser les réflexions diffuses [MvMV93]. Là encore, un bruit blanc est filtré en fonction d'un filtre représentant les propriétés moyennes (sur l'espace des directions) de la source, du récepteur et de la diffusivité des parois. Toutefois, l'approche étant destinée à être utilisée pour les premières réflexions, l'expression de la fenêtre temporelle est modifiée pour pouvoir commencer au début de la réponse (c'est-à-dire dès le temps d'arrivée du son direct).

3.5.3 Validité et limitations des modèles statistiques

Les méthodes développées en acoustique statistique permettent un calcul très rapide de la partie tardive de la réponse impulsionnelle par filtrage d'un bruit blanc. Elles sont, toutefois, fondées sur l'hypothèse de champ diffus et ne sont donc valables *a priori* que sous ces conditions. Nous avons vu qu'un critère constitué d'une limite inférieure dans l'espace temps/fréquence définit leur validité. Néanmoins, le champ diffus n'est que rarement atteint en réalité, puisqu'il nécessite des conditions particulières concernant la géométrie et la diffusivité des surfaces du milieu (*cf.* paragraphe 1.6.2). Le même problème existe lorsque de telles méthodes sont utilisées pour calculer les réflexions diffuses, puisqu'aux faibles ordres de réflexions, le champ n'est certainement pas encore diffus. Hors, comme nous l'avons déjà vu, «réflexions diffuses» n'est pas équivalent à «champ diffus». De manière plus générale, malgré leur efficacité, les approches statistiques, négligent complètement l'influence de la géométrie et restent de ce fait limitées à des situations particulières comme les espaces clos. Dans des environnements plus généraux (espaces ouverts par exemple) les hypothèses sur lesquelles elles sont fondées ne sont certainement plus valides.

Enfin, lorsqu'elles sont utilisées en complément d'approches géométriques, se pose le problème du «recollement» des deux parties de réponse impulsionnelle obtenues, qui peut donner lieu à des artefacts audibles [Hei93, MvMV93].

Toutefois, dans leur cadre de validité, elles permettent d'obtenir des réverbérations réalistes lors de l'écoute. Elles sont de ce fait à la base de tous les «réverbérateurs artificiels» utilisés pour la production audio.

Chapitre 4

Acoustique et réalité virtuelle

APRÈS avoir exploré les principaux modèles utilisés pour les simulations acoustiques, nous présentons ici comment ils ont été utilisés dans le contexte de la réalité virtuelle auditive. Nous discuterons les avantages et inconvénients de ces méthodes tant du point de vue de leur réalisme que de leur complexité algorithmique. Nous commençons par présenter le principe à la base de la synthèse d'un champ sonore virtuel à partir d'une simulation acoustique : l'*auralisation*.

4.1 Auralisation

Rendre audible un champ sonore virtuel implique une phase de traitement du signal, ou auralisation, dans laquelle un son «brut» (ou *anéchoïque*) va être filtré par la réponse impulsionnelle de l'environnement de synthèse avant d'être présenté à l'auditeur (*cf.* Figure 4.1). La phase de génération de la réponse impulsionnelle et la phase d'auralisation sont bien évidemment étroitement liées et peuvent ne pas être distinctement séparées. Cependant la phase d'auralisation est assez souvent effectuée en post-traitement du calcul de réponse impulsionnelle [LB92, KDS93].

On distinguera deux types d'approches dans le cadre d'applications interactives :

– les approches plutôt quantitatives, héritées des simulations géométriques en acoustique prévisionnelle et plus ou moins simplifiées pour permettre des temps de traitement interactifs ;

– des approches plus qualitatives, basées sur des «effets» perceptifs et destinées pour la plupart à des applications «temps-réel» de type production audio (mixages, effets, etc.) pour des studios d'enregistrements, mais également à présent pour les jeux vidéo.

4.1.1 Réponse impulsionnelle et filtrage numérique

L'un des points importants dans la réalisation d'un système d'auralisation est le choix du type de filtre utilisé pour modéliser la réponse impulsionnelle. Deux choix existent : les filtres à *Réponse Impulsionnelle Finie* (RIF) et les filtres à *Réponse Impulsionnelle Infinie* (RII) [SH94, Jot92a] dont nous allons expliquer les différences. Dans les systèmes numériques discrets, le filtrage d'un signal se traduit par une convolution discrète (*cf.* paragraphe 1.3) :

$$y(n) = \sum_m h(m)x(n-m), (n,m) \in \mathbb{N}^2,$$

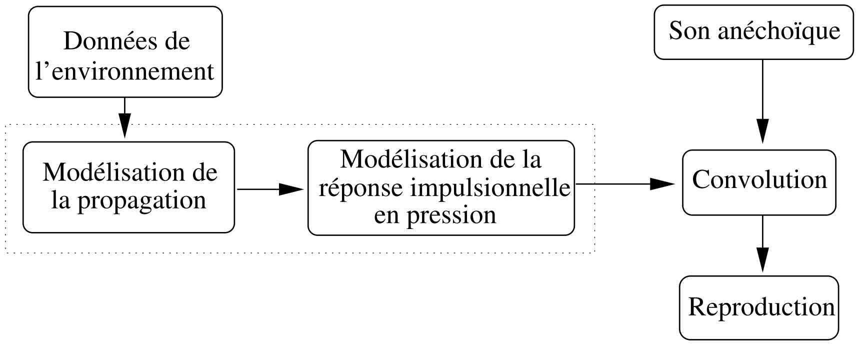


Figure 4.1 - *Processus d'auralisation. La réponse impulsionnelle en pression de l'environnement est simulée et utilisée pour filtrer un son «brut» avant écoute.*

où $y(n)$ est le signal filtré et $x(n)$ le signal d'entrée. Deux cas se présentent alors : soit la somme est une somme finie, les $h(m)$ étant non nuls uniquement pour un nombre fini de termes ; soit la somme est infinie car il existe une infinité de valeurs $h(m)$ non nulles. Dans le premier cas, le filtre est un filtre RIF et les $h(m)$ sont appelés les coefficients du filtre ; ils le définissent complètement. Ils peuvent se calculer de manière directe, par exemple à partir de la fonction de transfert (par transformée de Fourier inverse discrète). Dans le deuxième cas, le filtre est un filtre RII et son fonctionnement est régi par une équation récursive (liant un certain nombre d'éléments du signal d'entrée à un certain nombre d'éléments du signal de sortie) du type :

$$y(n) = \sum_{l=0}^{L} a_l x(n-l) - \sum_{k=1}^{K} b_k y(n-k).$$

Dans ce cas, les coefficients sont les a_l et les b_k et leur détermination n'est pas directe. En contrepartie, le coût (en nombre de multiplications par exemple) d'une convolution par un filtre RII est généralement beaucoup plus faible que celui d'une convolution par un filtre RIF. Pour plus d'information sur le sujet, on pourra consulter les références suivantes [OS75, Cot97, Bel95, dEV92, PTVF92].

4.1.2 Son «3D». Binauralisation et transauralisation

La plupart des recherches effectuées ces dernières années en acoustique virtuelle s'intéressent essentiellement à la reproduction d'un champ sonore réellement tridimensionnel (au sens où l'auditeur va pouvoir percevoir la position relative d'une source dans l'espace, y compris au dessus, au dessous et derrière lui). Ces approches permettent d'augmenter de manière extrêmement significative la sensation d'immersion par rapport à une reproduction monophonique ou stéréophonique classique. On songera au parallèle avec l'image où les systèmes de stéréovision permettent une immersion décuplée par rapport à la projection 2D sur écran. Ces techniques sont basées sur la modélisation des fonctions de transfert de la tête de l'auditeur (*cf.* paragraphe 2.2.2). Nous décrivons ici rapidement les principales approches pour la simulation d'un champ sonore tridimensionnel ou *spatialisation*.

Dès-lors qu'il s'agit de reproduire un champ sonore «3D» aux oreilles d'un auditeur, l'auralisation devient dépendante du dispositif de reproduction sonore : casque (on parle alors de *binauralisation*) [LB92, Bur92a, Bur92b] ou enceintes (*transauralisation*) [CB89, Gar95, Møl92a, JLW95, ETA98].

La binauralisation , conçue pour l'écoute au casque, consiste à générer, en fonction de la direction d'incidence du son, deux signaux (un pour chaque oreille) à l'aide du filtrage approprié. Elle permet d'obtenir des résultats très convaincants, exception faite d'une confusion avant-arrière assez fréquente, (les sons frontaux étant perçus à l'arrière) et de problèmes pour extérioriser les sons qui sont souvent perçus «à l'intérieur de la tête» [Rub91]. En outre, l'auditeur est contraint de porter un casque.

La transauralisation , basée sur la reproduction avec enceintes, permet de s'affranchir de ce problème mais au prix d'un traitement plus complexe. En effet, supposons que l'on veuille générer des signaux précis aux deux oreilles (situées en un point précis), il faut utiliser au minimum deux enceintes. Or, une partie du son émis par l'enceinte gauche va être perçue par l'oreille droite et vice-versa. Il faut donc supprimer cette «contribution croisée» pour reproduire l'effet voulu. En théorie, il est également nécessaire de compenser l'effet de l'environnement où se trouvent les enceintes, ce qui est une opération délicate. Une conséquence des techniques de transauralisation est que le champ sonore est valide dans une zone très réduite autour du point où le système a été calibré, mais permet une sensation d'extériorisation du son améliorée et la réduction des confusions avant-arrière (*cf.* paragraphe 2.2.2), en particulier quand les enceintes sont devant l'auditeur. Des recherches ont également été effectuées pour permettre un meilleur contrôle et un agrandissement de la zone de validité de l'effet à l'aide de multiples haut parleurs (on parle alors de *transauralisation généralisée*) [Der97, ETA98, Lee98]. Des systèmes de suivi de position permettent également de corriger en temps réel le signal émis pour maintenir sa cohérence en fonction des mouvements de l'auditeur par rapport aux enceintes. On se reportera à [Beg94] pour des informations complémentaires sur le sujet.

Ces techniques de «son 3D» permettent de réaliser des effets perceptiblement très «efficaces». Il est remarquable de constater qu'un nombre très important d'approches en acoustique virtuelle, si ce n'est la majorité, sont focalisées sur la spatialisation du son et que bien peu en comparaison s'intéressent à la simulation des phénomènes de propagation, pourtant essentiels.

4.1.3 Simulations en bandes fines et en bandes larges

Comme nous l'avons déjà remarqué, la phase d'auralisation et la phase de simulation sont très liées, même si l'auralisation peut être vue comme un post-traitement de la phase de simulation. En effet, la méthode utilisée pour gérer l'aspect fréquentiel de la simulation va influer sur le traitement du signal à effectuer durant l'auralisation. On peut ainsi catégoriser les méthodes de simulation en deux classes [LB92] : les approches dites «*larges bandes*» et celles dites «*en bandes fines*».

Les approches larges bandes effectuent le travail de simulation pour différentes bandes de fréquence, en général des bandes d'octaves ou de tiers d'octaves (on les appelle aussi approches «en bandes d'octaves»). Cela ne signifie pas forcément que toute la simulation est répétée (on ne va pas recalculer plusieurs fois le lancer de rayons par exemple). On va plutôt construire plusieurs échogrammes, chacun valable (du point de vue des coefficients de réflexion choisis, *etc.*) pour la fréquence centrale d'une bande d'octave. En filtrant ensuite chacun de ces échogrammes par le filtre passe-bande approprié puis en les additionnant, on obtient une approximation de la réponse impulsionnelle recherchée (*cf.* Figure 4.2). Il est donc clair que des interférences ne pourront se produire que sur les fréquences centrales considérées. Ces approches sont motivées par le fait que toutes les données acoustiques sont généralement données en bandes d'octaves ou de tiers d'octaves et la perception en bandes critiques de l'oreille (proches de tiers d'octaves) (*cf.* paragraphe 2.1.2) .

Dans les approches en bandes fines, on dispose de toutes les réponses impulsionnelles élémentaires correspondant aux réflexions, à la source, au récepteur, *etc.* [1]. On effectue alors une simulation indépendante de la fréquence. On obtient comme résultat un échogramme unique dont l'amplitude des «pics» est uniquement une fonction du temps de propagation (l'atténuation de l'onde en fonction de la distance parcourue). Pour chacun de ces pics on somme (en tenant compte de son retard et de son amplitude)

1. on peut aussi construire des réponses impulsionnelles à partir des seules données de leur spectre en bandes d'octaves en fixant certaines contraintes sur la phase (*cf.* Annexe C.1).

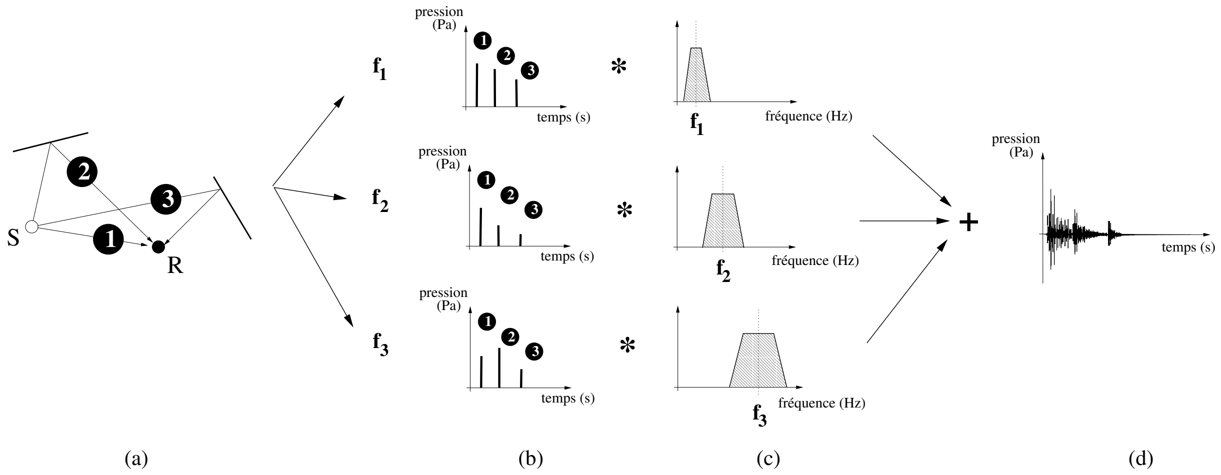


Figure 4.2 - *Simulation en bandes larges. (a) La simulation est conduite pour différentes bandes de fréquences. On obtient comme résultat plusieurs échogrammes chacun valable pour une bande de fréquence. (b,c) Par convolution avec les filtres passe-bande correspondants, on obtient un ensemble de réponses impulsionnelles, chacune également valable pour une bande de fréquence spécifique. (d) La réponse impulsionnelle recherchée est finalement obtenue par sommation de ces différentes réponses.*

la réponse impulsionnelle du chemin correspondant. Celle-ci est définie comme la convolution de tous les filtres élémentaires intervenant le long de ce chemin (*cf.* Figure 4.3). Ces approches sont donc plus proches des phénomènes physiques, puisque prenant en compte automatiquement les phénomènes d'interférences pour toutes les fréquences (lors de la superposition des réponses impulsionnelles de chaque chemin).

Ces aspects fréquentiels sont *a priori* indépendants du type de simulation effectué. Toutefois, il faut noter qu'une simulation comportant des phénomènes de réflexion diffuse (lancer de rayons stochastique ou échanges radiatifs) est peu appropriée pour un traitement en bandes fines. En effet, le nombre de contributions atteignant le récepteur est alors extrêmement important, ce qui implique un traitement du signal au coût prohibitif. Dans ce cas, comme nous l'avons vu au chapitre précédent, on utilise une approche en bandes larges et des échogrammes intégrés sur de petits intervalles de temps constants. La réponse impulsionnelle peut être alors obtenue par reconstruction du spectre en bande fine (on interpole entre les valeurs pour les différentes bandes) pour chaque pas de temps de ces histogrammes puis par transformée de Fourier inverse. L'information de phase étant manquante, on ajoute une phase aléatoire, ce qui s'éloigne des phénomènes physiques[2] [Kut93]. À l'inverse, les simulations dans lesquelles on ne considère que des réflexions spéculaires (sources-images en particulier, quelle que soit la manière de les calculer) se prêtent mieux à un traitement en bandes fines.

Pour résumer, on peut dire que ces deux approches ne gèrent pas la complexité fréquentielle de la simulation au même niveau ; les approches en bandes larges l'intègrent au calcul de la propagation, les approches en bandes fines la repoussent au niveau du traitement du signal (dans la mesure où on dispose d'une carte câblant les opérations de traitement du signal, les approches en bandes fines peuvent donc paraître plus appropriées). Il est donc difficile de privilégier l'une ou l'autre *a priori*.

2. néanmoins, l'importance de la phase par rapport au contenu fréquentiel de la réponse impulsionnelle est un aspect encore sujet à controverses [Kut93, LB92] (*cf.* paragraphe 2.3).

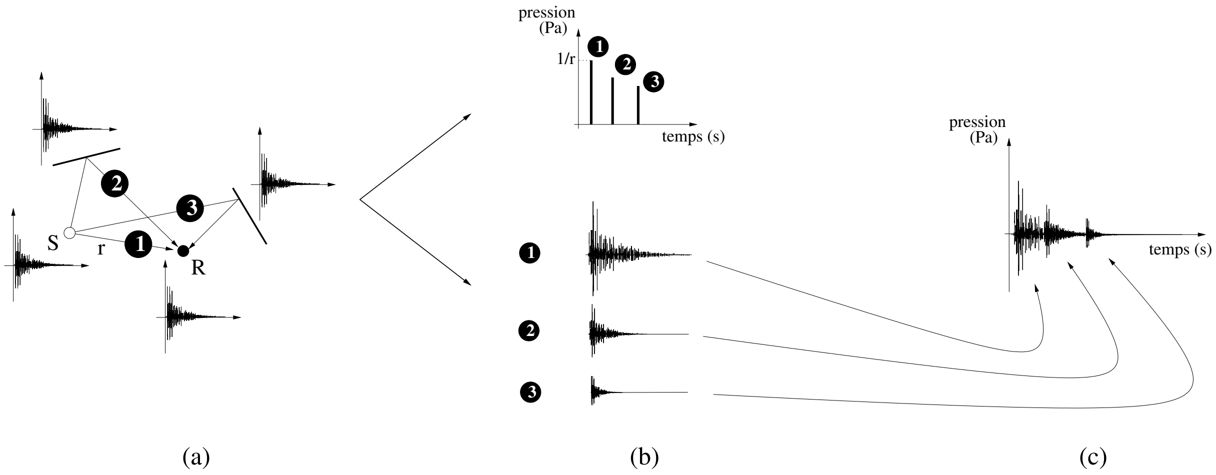


Figure 4.3 - *Simulation en bandes fines. (a) on dispose à l'origine de toutes les réponses impulsionnelles «élémentaires» des matériaux, de la source et du récepteur. (b) Après avoir simulé les différents chemins de propagation, on reconstruit pour chacun d'entre-eux la réponse impulsionnelle correspondante. (c) Enfin, toutes ces réponses sont sommées pour donner la réponse impulsionnelle recherchée.*

4.2 Simulation et auralisation interactive

Nous commençons par évoquer ici les techniques d'auralisation utilisées en acoustique prévisionnelle, puisque c'est dans ce cadre que la plupart des travaux ont d'abord été réalisés. L'auralisation est souvent dans ce cas une phase bien séparée de la simulation qui est généralement basée sur des techniques géométriques (*cf.* paragraphe 3.3). Les approches en bandes fines [MvMV93] , comme en bandes larges sont utilisées [Kut93]. Dans tout les cas, la réponse impulsionnelle finale est représentée par un filtre RIF, car ils permettent une modélisation plus fine que les filtres RII. Une fois la réponse impulsionnelle calculée (ou les réponses impulsionnelles dans le cas d'une simulation binaurale), un signal «brut» est filtré et présenté à l'auditeur. Il faut noter que ce processus peut ne pas être négligeable en temps de calcul si la durée de la réponse impulsionnelle est importante (une réponse impulsionnelle de plus de 1.5 secondes (66000 coefficients à 44kHz) est classique dans des salles de tailles moyennes ($\simeq 3000$ m^3)). Le traitement du signal est généralement effectué en utilisant des cartes spécialisées dans le traitement des convolutions en temps-réel [RM95]. Toutefois, dans le cadre d'environnements dynamiques, il faudrait pouvoir effectuer la simulation et mettre à jour les très nombreux coefficients du filtre en temps-réel, ce qui reste encore impossible. Ces approches sont donc généralement réservées aux environnements statiques.

Vers l'interactivité

La plupart des approches pour réaliser une auralisation interactive sont basées sur un calcul des sources-images, puisque, comme nous l'avons déjà vu (*cf.* paragraphe 3.3.1), elles permettent un traitement unifié et rapide. L'un des inconvénients de l'approche est sa limitation aux faibles ordres de réflexion. Or, comme nous l'avons vu au paragraphe 2.3, dans les environnement réverbérants, la réverbération tardive dépend peu de la position des sources et des récepteurs. L'approche communément adoptée est alors de calculer une unique réverbération tardive de référence. On l'utilisera alors quelles que soient les positions des sources et des récepteurs, alors que les premières réflexions seront, elles, recalculées [WF90, RM95].

Au niveau traitement du signal, ce sont les approches en bande fines qui sont le plus utilisées ; les convolutions élémentaires étant réalisées pour chaque source-image en temps réel à l'aide de cartes spécialisées comprenant plusieurs processeurs de traitement du signal (DSP) [WF90, RM95]. Toutefois ces approches restent limitées à de faibles ordres de réflexions (pour information, l'une de ces cartes, le «convolvotron», permet de recalculer dix chemins seulement en temps réel). Des systèmes assez complets ont toutefois été mis au point, basés sur l'utilisation de sources-images [SB95, Str98]. Pour les grands ordres de réflexion, on préférera généralement des techniques de type «*réverbérateur artificiel*» que nous présenterons dans le prochain paragraphe. Ces modèles hybrides au niveau du filtrage correspondent généralement à un modèle hybride de simulation géométrique et statistique.

Dans le cas particulier où seul l'auditeur se déplace, des approches basées sur un précalcul de l'arbre des réflexions, dans le cadre d'un lancer de faisceaux adaptatif, permettent une mise à jour rapide [MOD96]. Il en est de même lorsque les propriétés des matériaux sont modifiées. Récemment, une telle approche combinée avec le précalcul d'une structure codant les relations de visibilité dans l'environnement (basée sur la définition de «cellules» et de «hublots» [TS91, Tel92, FST92]) a permis de recalculer en temps réel la réponse impulsionnelle dans des environnements complexes. Jusqu'à dix ordres de réflexions spéculaires peuvent être mis à jour de manière interactive (voire temps-réel) dans des environnements complexes de type villes ou bureaux comptant de l'ordre de 10 000 polygones, ce qui est une première dans le domaine de l'acoustique virtuelle [FCE$^+$98]. Néanmoins, on peut s'interroger sur l'adaptation d'une telle approche à des sources mobiles, qui impliquerait de recalculer une partie ou l'intégralité de la structure d'optimisation. Son efficacité s'en trouverait sans doute réduite. Il en est de même dans le cadre de la simulation de phénomènes de propagation plus complexes.

4.3 Acoustique virtuelle temps-réel pour la production audio

Un deuxième type d'approches est celui utilisé dans les applications de production audio de type «boîte à effets», pour le mixage par exemple. Dans ce genre d'applications, l'ingénieur du son va régler «à l'oreille» les paramètres voulus. Deux points essentiels caractérisent donc ces approches : peu de contrôles, le plus pertinence possible d'un point de vue perceptif et surtout une réponse temps-réel du système, ce qui implique l'utilisation de techniques de filtrage peu coûteuses (on préférera ici les RII aux RIF). Ces systèmes de *réverbération artificielle* présents dans de nombreux processeurs d'ambiance et tables de mixage, sont basés sur les méthodes d'acoustique statistique (*cf.* paragraphe 3.5). La réponse impulsionnelle est générée par des bancs de filtres réverbérants (à base de filtres «*en peigne*» [Jot92a]). Ces filtres sont mis en boucle, un filtre réverbérant générant plusieurs copies du signal d'entrée filtrées et décalées en temps. Dans le cadre des applications à la production audio, on peut néanmoins distinguer l'approche suivie par Jot pour son «spatialisateur» [Jot92a]. Cette approche permet le contrôle des premières réflexions à l'aide de paramètres issus des critères subjectifs d'évaluation utilisés en acoustique prévisionnelle. Les paramètres du filtre de réverbération tardive peuvent en outre être calculés par optimisation pour simuler une réverbération mesurée. Toutefois, il n'existe pas vraiment de géométrie sous-jacente, ce qui limite l'utilisation de l'approche.

D'autres «boîtes à effets» ont été développées (comme le Roland RSS10 [Sal95]) pour reproduire un son spatialisé (binauralisation ou transauralisation) mais sont également très limitées puisqu'elles ne permettent pas de prendre en compte d'autres effets (environnement, propagation, *etc.*). Leur architecture modulaire leur permet néanmoins de s'intégrer facilement dans la chaîne des traitements effectués en studio.

4.4 Rendu «intégré» image-son

Très peu d'approches, en acoustique virtuelle, se sont intéressées à la notion d'un *rendu intégré image-son*. Nous entendons par là, des techniques ou algorithmes permettant une modélisation et un rendu *cohérent* du son et de l'image pour des applications de réalité virtuelle ou d'animation 3D. Cela implique en particulier d'être capable de synchroniser le son et l'image et bien sur de faire intervenir l'environnement tridimensionnel dans le rendu sonore. Takala et Hanh ont proposé une approche permettant de réaliser du rendu sonore pour des applications d'animation 3D [TH92]. Ils utilisent pour cela des «arbres de rendu» dont les feuilles sont les sons bruts (ou «textures sonores») qu'il combinent entre eux à l'aide de différents opérateurs ou filtres (les nœuds de l'arbre). Durant une animation, il génèrent, par pas de temps fixes (25 fois par seconde en général), des paramètres de contrôle pour ces opérateurs et filtres. Dans une deuxième passe, le rendu de l'arbre est effectué pour chaque pas de temps et une bande son est générée. Des extensions interactives ont également été envisagées [HFGL95]. Plus récemment, un rendu sonore virtuel à été utilisé pour simuler l'audition d'acteurs virtuels durant une animation [NT95] mais la simulation de la propagation reste très simple. Enfin, ces dernières années, sont apparues des primitives de gestion de sources sonores et récepteurs dans de nombreuses bibliothèques de graphique 3D ou «multi-média», en particulier pour des applications sur réseau (VRML, Java, Direct Sound, Intel RSX). Là encore, les possibilités offertes restent assez limitées au niveau de la simulation de la propagation (la réverbération reste très «artificielle», et l'environnement n'influence pas le rendu du son). De plus, ces bibliothèques, même si elles permettent de rajouter des sources sonores dans des environnements virtuels dynamiques, ne permettent pas vraiment une modélisation unifiée de la scène sonore et visuelle (définition d'obstacles et de réflecteurs de manière géométrique par exemple, propriétés des matériaux, des sources, des récepteur, *etc.*).

Chapitre 5

Discussion générale et bilan

A LA LUMIÈRE des deux chapitres précédents, nous donnons ici une discussion générale des approches précédentes dans le cadre de la mise au point d'une méthode de simulation de haute qualité destinée à des environnements graphiques interactifs. Quatre thèmes nous paraissent pertinents pour définir notre cadre de recherche :

– *la généralité de l'environnement* à simuler, étendue aux environnements *dynamiques*, très importants dans le cas d'applications interactives ;

– *la généralité des phénomènes de propagation* simulés, nécessaire au réalisme de la simulation ;

– *le contrôle de la complexité algorithmique et du réalisme* de la simulation, nécessaire pour s'adapter à des besoins divers ;

– *l'intégration avec le graphique 3D*, essentielle aux applications de réalité virtuelle de plus en plus présentes.

5.1 Traitement des environnements dynamiques généraux

L'un des points clés de l'approche que nous souhaitons développer est qu'elle soit applicable à des environnements généraux y compris des environnements dynamiques, où la source, l'environnement, les matériaux et le récepteur en particulier peuvent évoluer au cours du temps. Contrairement aux simulations en synthèse d'image où l'on peut rapidement considérer les sources lumineuses fixes, les sources sonores sont presque toujours en mouvement (hors cas particulier de sonorisation d'une salle de spectacle avec des enceintes fixes, on songera à des véhicules, des locuteurs, *etc.*) Seules les approches géométriques (les méthodes de sources-images en particulier) ont une complexité algorithmique et une souplesse leur permettant d'être appliquées à des environnements dynamiques, au moins pour les faibles ordres de réflexion. Nous avons vu que des approches par lancer de faisceaux existent permettant de recalculer en temps-réel jusqu'aux dix premiers ordres de réflexions spéculaires dans des environnements complexes lorsque l'auditeur se déplace. Pour la réverbération tardive, les approches statistiques paraissent attractives du fait de leur efficacité. Toutefois, elles sont fondées sur des hypothèses qui peuvent rapidement devenir limitatives dans des environnements généraux (par exemple non clos). En particulier elles négligent tout lien avec la géométrie sous-jacente du problème. Les approches par échanges radiatifs semblent mieux appropriées puisqu'elles restent géométriques. Elle peuvent permettre, en outre d'obtenir une solution qui est indépendante du point de réception. Des recherches récentes en synthèse d'image ont montré qu'il était possible de mettre à jour de manière interactive une solution de radiosité dans le cas d'environnements dynamiques [DS97].

5.2 Traitement des phénomènes de propagation complexes

Le principal phénomène auquel on s'intéresse pour le calcul de réponse impulsionnelle est la réflexion du son puisqu'il est responsable de la réverbération. Néanmoins, d'autres phénomènes, comme la diffraction qui est souvent négligée, restent importants. Si l'on peut argumenter sur la perceptibilité du phénomène de diffraction par des obstacles, il est des cas où un traitement binaire de la visibilité ne peut pas donner une solution satisfaisante au problème. On songera, pour s'en convaincre, au simple cas d'une source invisible (au sens géométrique) d'un auditeur (dans une fosse d'orchestre par exemple) et pourtant clairement audible. La prise en compte de phénomènes de transmission et diffraction est donc nécessaire mais implique un surcoût non négligeable dans un algorithme géométrique (de type lancer de rayons). De même, la simulation de phénomènes de réflexion généraux (autres que le modèle spéculaire idéal) est également un point intéressant. En outre, la réflexion diffuse reste encore un problème coûteux à traiter avec des approches de type lancer de rayons alors que les méthodes d'échanges radiatifs permettent de le traiter plus efficacement. Toutefois ces approches restant énergétiques, on peut s'interroger sur leur validité. D'autres phénomènes de propagation liés aux environnements dynamiques et significatifs du point de vue perceptif devront également être simulés. C'est en particulier le cas du «décalage Doppler».

5.3 Maîtrise de la complexité et de la précision des algorithmes

La plupart des approches géométriques existantes offrent très peu de compromis entre la qualité de la simulation et son coût. De plus, l'influence des paramètres de simulation sur le résultat est difficilement quantifiable *a priori* (nombre de rayons tirés dans un algorithme de lancer de rayons, par exemple). A l'opposé, les approches de type «réverbérateur artificiel» offrent des contrôles simples et significatifs du point de vue perceptif, mais au détriment de toute possibilité d'ajustement automatique à un environnement donné. Une méthode géométrique adaptative en complexité, comme celles développées ces dernières années pour la simulation de l'éclairage [HSA91] serait plus appropriée puisqu'elle permettrait de maintenir un lien avec la géométrie tout en contrôlant la finesse de la solution et son temps de calcul. Une telle approche est essentielle si l'on songe à des applications de réalité virtuelle distribuées où l'on pourrait adapter la complexité à la puissance des différentes machines *a priori* variable. Un algorithme géométrique adaptatif pourrait en outre permettre de passer continûment d'applications qualitatives du type mixage audio à des applications quantitatives en acoustique prévisionnelle.

Enfin, un point important est le traitement du signal nécessaire pour l'auralisation. Il faut en effet garder à l'esprit que le temps nécessaire au traitement du signal est loin d'être négligeable. Il faut donc essayer de fournir les résultats de la simulation sous une forme propice à un traitement le plus direct et rapide possible. Dans le cadre de cette étude, nous n'aborderons pas les problèmes d'accélération du traitement du signal lui-même, même s'il est déjà clair que les cartes spécialisées existantes sont loin de répondre à l'attente imposée par des simulations d'acoustique virtuelle. En effet, leur architecture reste très spécifique et offre peu de points de contrôle (généralement on se contente de spécifier les filtres). Elles sont exclusivement prévues pour effectuer des convolutions et n'offrent aucune accélération de la simulation elle-même. Des calculs coûteux doivent donc encore être effectués par le processeur central de la machine. Leur efficacité reste donc limitée et difficile à exploiter.

5.4 Intégration avec le graphique 3D

Le dernier point à considérer est l'intégration du son dans des applications graphiques interactives. En effet, la réalité virtuelle acoustique n'est pas destinée à exister en soi mais dans le cadre d'applications multi-modales combinant graphiques 3D immersifs et son spatialisé. Contrairement au domaine de l'acoustique virtuelle[1], dont le développement est très récent, le domaine du graphique 3D a déjà atteint une certaine «maturité» d'utilisation, en particulier au travers de cartes graphiques très puissantes permettant d'effectuer en temps-réel les opérations complexes nécessaires au rendu d'une scène polygonale. Dès lors qu'un calcul d'acoustique géométrique doit être intégré dans un environnement graphique 3D, on peut envisager l'opportunité de mettre à profit la puissance de calcul de telles cartes de rendu graphique pour effectuer des opérations relatives au rendu du son de manière souple et efficace. On pourrait dès lors utiliser des outils et des structures de données communes pour la modélisation graphique et celle de l'environnement acoustique. Cela permettrait d'obtenir une plus grande interaction entre modèle géométrique et système de rendu sonore et permettrait d'améliorer du même coup la cohérence et donc le réalisme de la simulation.

Enfin, il semble clair qu'il n'est pas envisageable d'utiliser la même complexité géométrique pour les simulations visuelles et auditives. Il parait donc judicieux dans le cas d'une approche géométrique de pouvoir définir une géométrie pour la simulation visuelle et une géométrie (simplifiée *a priori*) pour la simulation acoustique.

5.5 Bilan

Les différents aspects que nous venons de décrire ouvrent bien évidemment des perspectives de recherche extrêmement vastes, que l'on ne prétend pas résoudre entièrement ici. Dans le cadre de cette thèse nous présentons toutefois quelques solutions à ces différents problèmes sous deux aspects qui constituerons les deux prochaines parties de ce document :

– La réalisation d'un environnement graphique 3D permettant un rendu du son intégré et bien adapté à des simulations dynamiques. Nous montrerons également qu'il est possible dans ce contexte d'utiliser les capacités de cartes graphiques pour effectuer la simulation de phénomènes complexes comme la diffraction par des obstacles.

– La mise au point d'un algorithme de simulation adaptatif en complexité, fondé sur les approches de *radiosité hiérarchique* développées ces dernières années en synthèse d'image. En particulier, nous nous intéresserons dans ce cadre à la simulation combinée de réflexions globales spéculaires et diffuses à différents niveaux de détail au cours du temps et à l'extension de telles méthodes, à l'origine énergétiques, à des phénomènes ondulatoires (interférences par exemple).

1. nous entendons ici, rendu sonore complexe intégré dans un environnement de réalité virtuelle. Il est clair que les premières expériences d'auralisation, binauralisation et l'utilisation de réverbérateurs artificiels se sont développées depuis bien plus longtemps.

❒ ❒ ❒

Dans cette partie, nous avons examiné les différentes méthodes utilisées en acoustique pour la modélisation de la réponse impulsionnelle d'un environnement virtuel. Deux types d'approches se distinguent dans le cadre d'applications d'acoustique virtuelle. D'une part, des approches inspirées par les simulations effectuées en acoustique prévisionnelle précises mais coûteuses en temps de calcul et ne prenant pas toujours en compte certains phénomènes de propagation pourtant non négligeables. D'autre part, des approches plus qualitatives, permettant d'obtenir de faibles temps de calcul au détriment du réalisme, de la flexibilité ou des possibilités de contrôle offertes. En outre, peu de méthodes existantes permettent une intégration dans un environnement où son et image sont directement synchronisés. Dans les deux prochains chapitres, nous présentons la première contribution de cette thèse : un système de rendu sonore pour l'animation de synthèse et la réalité virtuelle permettant un rendu interactif du son directement synchronisé avec l'image. Ce système permet en outre de prendre en compte de manière semi-quantitative des phénomènes de propagation complexes comme l'occultation et la diffraction des ondes sonores, leur transmission, réflexion et atténuation dans des environnements virtuels dynamiques généraux.

Troisième partie

Un système de rendu sonore pour la réalité virtuelle traitant les occultations

Chapitre 6

Traitement des occultations sonores

Il y a encore à considérer dans l'émanation de ces ondes, que chaque particule de la matière, dans laquelle l'onde s'étend, ne doit pas communiquer son mouvement seulement à la particule prochaine qui est dans la ligne droite tirée du point lumineux ; mais qu'elle en donne aussi nécessairement à toutes les autres qui la touchent, et qui s'opposent à son mouvement. De sorte qu'il faut qu'autour de chaque particule il se fasse une onde dont cette particule soit le centre. [...] Mais chacune de ces ondes ne peut être qu'infiniment faible comparée à l'onde [principale], à la composition de laquelle toutes les autres contribuent par la partie de leur surface qui est la plus éloignée du point [lumineux].

Christiaan Huygens, *Traité de la lumière,* 1690[1].

COMME nous l'avons vu précédemment, peu d'approches, y compris dans le cadre d'applications de simulation quantitative sont à même de prendre en compte de manière satisfaisante les effets, pourtant non négligeables, de la diffraction du son par d'éventuels obstacles. De plus, le coût de calcul des algorithmes existants les rend inutilisables dans le cadre d'applications interactives. Ce chapitre présente la première contribution de ce travail de thèse et concerne le calcul interactif des effets d'obstacles sur la propagation du son dans des environnements virtuels dynamiques. Nous présentons deux méthodes, l'une qualitative mais pouvant être utilisée en temps réel, l'autre plus précise mais nécessitant un coût de calcul supérieur tout en restant interactive. Les deux approches tirent avantage d'une description de la scène sous forme polygonale et du rendu 3D câblé rapide de cartes graphiques spécialisées pour effectuer un calcul de diffraction approché entre une source et un récepteur ponctuels.

6.1 Diffraction du son et interactivité

La plupart des travaux concernant la diffraction du son sont consacrés au calcul d'atténuation par des écrans ; historiquement parce que les obstacles plans permettent d'obtenir des solutions analytiques (à l'aide de la théorie de Fresnel-Kirchhoff par exemple [Fil94, Hec87]). Ils sont à présent étendus à des écrans de formes plus complexes (en particulier «épais») dans le cadre de l'atténuation de bruit de trafic (routier, ferroviaire) qui en est l'une des applications principales. Les approches analytiques

1. édition française actualisée, Dunod éditeur, 1992.

aux problèmes de diffraction peuvent être utilisées dans le cas d'obstacles de formes simples (écran droit infini) [Mae68]. Néanmoins, dès que la forme des obstacles devient plus complexe, deux approches sont utilisées : les éléments finis de frontière (*cf.* paragraphe 3.2) [Jea98, Dau78, HCWH91] et la théorie géométrique de la diffraction (*cf.* paragraphe 3.3.2) [Kaw81]. Les approches par éléments finis restent les approches de référence, mais demandent une puissance de calcul les rendant peu adaptées à des calculs interactifs. Dans le cas de fréquences élevées (longueurs d'onde centimétriques) en particulier, il faut utiliser un maillage des surfaces très fin. On obtient alors rapidement des temps de calculs de l'ordre de l'heure. La théorie géométrique de la diffraction nécessite de construire des rayons diffractés, processus qui peut être coûteux dans des environnements complexes. Même si des structures d'optimisation peuvent être envisagées [FCE$^+$98], il est nécessaire de tout recalculer dans le cas d'un environnement totalement dynamique (sources, récepteurs et environnement en mouvement). Enfin, à notre connaissance, aucune approche ne permet de traiter de manière plus qualitative et rapide les problèmes de diffraction.

6.2 Une approche temps-réel utilisant les ellipsoïdes de Fresnel

Pour combler le manque d'approches permettant un traitement temps-réel des effets de l'occultation du son par des obstacles, nous avons développé une technique qualitative basée sur les *ellipsoïdes de Fresnel*. Le but recherché ici est double :

– développer une approche souple et applicable de manière satisfaisante dans des environnements virtuels généraux et dynamiques ;

– développer une approche suffisamment rapide pour être utilisée en temps réel ou au moins d'une manière interactive.

L'approche que nous proposons calcule l'atténuation de l'intensité acoustique pour différentes bandes de fréquences le long d'un chemin source-récepteur en fonction des obstacles obstruant les premiers ellipsoïdes de Fresnel correspondants.

6.2.1 Ellipsoïdes et zones de Fresnel

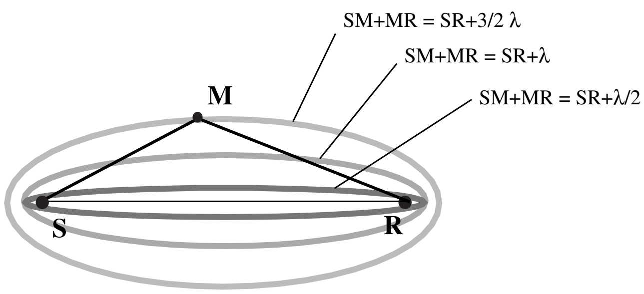
Les ellipsoïdes de Fresnel, définis pour une longueur d'onde λ donnée, sont une décomposition de l'espace en zones d'interférences alternativement constructives et destructives deux à deux[2] (*cf.* Figure 6.1 (a)). Ils ont pour foyers la source S et le récepteur R et sont définis comme l'ensemble des points $M \in \mathbb{R}^3$ vérifiant l'équation :

$$\|\vec{SM}\| + \|\vec{MR}\| = \|\vec{SR}\| + \frac{k\lambda}{2}, k \in \mathbb{N}^+. \tag{6.1}$$

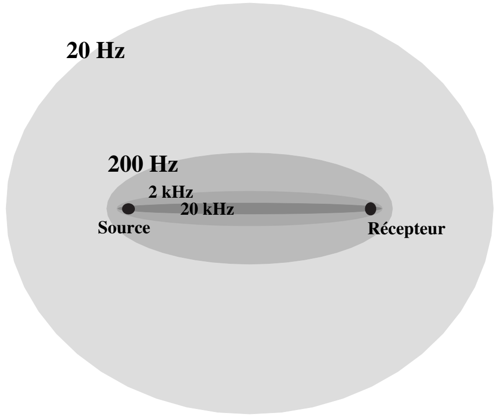
En général, on s'intéressera principalement aux *premiers ellipsoïdes de Fresnel*, obtenus pour $k = 1$ dans l'équation 6.1. On peut montrer en effet, qu'en *champ libre*, l'intensité reçue par le récepteur correspond à la moitié de l'intensité émise par la seule partie du front d'onde contenue dans les premiers ellipsoïdes de Fresnel [Hec87, Def96]. Ce résultat confère à ces zones un intérêt particulier, puisqu'elles délimitent l'espace où se propage la majeure partie de l'énergie contribuant au champ au point de réception. Les premiers ellipsoïdes de Fresnel sont fréquemment utilisés en électromagnétisme, dans

2. cette décomposition à été utilisée à l'origine par Fresnel pour exprimer la contribution d'un élément différentiel de surface d'un front d'onde sphérique au champ (libre) lumineux en un point de réception [Hec87].

le cadre de la propagation d'ondes radio, pour l'étude de la diffraction le long d'une liaison antenne-récepteur [Ber88, LS97, Vau94, Lam92]. Ils ont également été utilisés en acoustique pour la réflexion ou la diffraction du son mais ils constituent dans ce dernier cadre une approximation valide uniquement pour les hautes fréquences et des obstacles simples (un demi plan infini par exemple) [Def96]. Ils ne peuvent donc pas être vraiment utilisés dans le cadre de simulations quantitatives mais conviennent parfaitement, comme nous le verrons, à des applications plus qualitatives.



(a)



(b)

Figure 6.1 - *(a) Les ellipsoïdes de Fresnel : L'espace est divisé en régions d'interférences constructives et destructives bordées par des ellipsoïdes, dont les foyers sont la source S et le récepteur R. (b) Premiers ellipsoïdes de Fresnel (i.e. pour $k = 1$) en 2D pour une source et un récepteur distants de 10 m (la figure est représentée à l'échelle).*

Zones de Fresnel

On appelle *zones de Fresnel*, la surface correspondant à l'intersection d'un front d'onde sphérique et du volume délimité par deux ellipsoïdes de Fresnel consécutifs. Dans le cas de la *première zone de Fresnel*, correspondant à l'intersection du front d'onde avec le premier ellipsoïde, on obtient une calotte sphérique. Pour les zones suivantes, ce sont des secteurs circulaires sphériques (*cf.* Figure 6.2). Dans le cas d'ondes planes, la première zone est un disque et les zones suivantes sont des couronnes circulaires.

On trouvera dans l'Annexe A.3 des détails supplémentaires sur les ellipsoïdes et les zones de Fresnel.

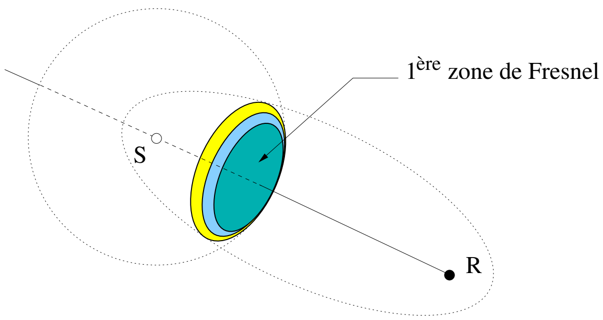


Figure 6.2 - *Les zones de Fresnel : intersections entre un front d'onde (ici sphérique) et les ellipsoïdes de Fresnel.*

En supposant que l'onde ne subisse pas de réflexion lors de son trajet depuis la source jusqu'au récepteur, on va donc supposer qu'il est suffisant de prendre en compte l'obstruction des premiers ellipsoïdes de Fresnel, pour déterminer l'atténuation du signal sonore. On supposera également que les obstacles situés hors de ce volume n'influencent pas la propagation de l'onde.

6.2.2 Utilisation du rendu 3D câblé pour les calculs de diffraction

Calculer le champ sonore diffracté implique de résoudre deux problèmes :

– identifier de manière rapide les objets contenus dans le premier ellipsoïde de Fresnel qui vont diffracter l'onde sonore ;

– évaluer les effets de la diffraction dus à chacun de ces objets.

Nous proposons de résoudre le premier problème en utilisant le rendu câblé de cartes graphiques spécialisées pour accéder de manière souple et rapide à l'information de visibilité entre la source et le récepteur. Pour cela, on utilise un rendu des polygones des objets la scène 3D à l'aide d'une caméra orthographique (*cf.* Figure 6.3). Afin d'obtenir une information de visibilité dépendante de la fréquence, la taille a du «cône» (ici, c'est donc un parallélépipède) de projection de la camera est choisie de manière à ce que celui-ci s'appuie sur le premier ellipsoïde de Fresnel correspondant :

$$a = \sqrt{\lambda\left(r + \frac{\lambda}{4}\right)}, \tag{6.2}$$

où λ est la longueur d'onde et r est la distance entre la source et le récepteur.

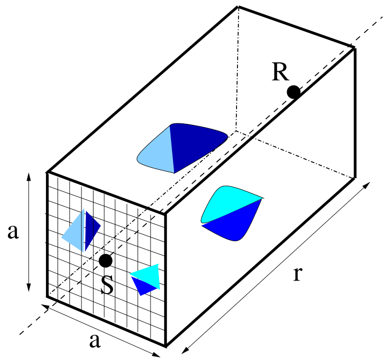


Figure 6.3 - *Cône de projection utilisé pour le calcul de la carte d'occlusion. Une projection parallèle est utilisée pour obtenir une image des obstacles entre la source S et le récepteur R.*

Le résultat du rendu est donc une image projetée de tous les obstacles entre la source et le récepteur qui intersectent le premier ellipsoïde de Fresnel (*cf.* Figure 6.4). Si la fréquence de l'onde augmente, le rayon de l'ellipsoïde diminue et on se rapproche d'une information de visibilité géométrique que l'on pourrait obtenir par lancer d'un rayon depuis la source jusqu'au récepteur. Au contraire, quand la fréquence de l'onde diminue, le rayon de l'ellipsoïde augmente et la «visibilité» est prise en compte à l'aide de ce «rayon» d'épaisseur non nulle.

(a) (b) (c)

Figure 6.4 - *Utilisation du rendu 3D câblé pour le calcul de la "visibilité sonore". (a) Vue 3D montrant un microphone, une source, les premiers ellipsoïdes de Fresnel associés (pour des fréquences de 400 et 4000 Hz) ainsi que des obstacles. (b) Visibilité depuis la source à 400 Hz. (c) Visibilité depuis la source à 4000 Hz. La zone blanche circulaire correspond à la première zone de Fresnel à mi-distance de la source et du récepteur.*

6.2.3 Calcul des coefficients d'atténuation liés à la diffraction

Nous choisissons de modéliser les effets de la diffraction due aux obstacles sous la forme d'une atténuation dépendante de la fréquence de l'onde sonore. On définit en conséquence le coefficient $V^\lambda \in [0,1]$ comme le rapport entre la surface projetée des obstacles et l'aire $A_{Fresnel}$ de la première zone de Fresnel, prise à mi-distance du récepteur. En effet, c'est l'aire apparente résultant de la projection du premier ellipsoïde de Fresnel. En pratique, on évalue très simplement ce coefficient en comptant les pixels correspondant aux obstacles dans la carte de visibilité $\mathrm{Pix}_{\mathrm{obstacles}}$, on obtient alors :

$$V^\lambda = \min\left(\left(\frac{\mathrm{Pix}_{\mathrm{obstacles}}}{A_{Fresnel}}\right)\cdot\left(\frac{a^2}{\mathrm{Pix}^2}\right),1\right), \tag{6.3}$$

où Pix est la taille de la fenêtre de rendu en pixels (la résolution étant donc Pix×Pix pixels) et a est la taille du cône de projection comme défini par l'équation 6.2. La valeur de $A_{Fresnel}$ est détaillé en Annexe A.3.

L'intensité acoustique au point de réception, après atténuation par diffraction, est alors donnée par :

$$I^\lambda_{diff} = I^\lambda_0(1-V^\lambda),$$

où I^λ_0 est l'intensité émise par la source.

Dans la formule 6.3, on peut remarquer que l'on prend en compte tous les pixels de la fenêtre de rendu. Or, une partie de ces pixels n'est pas contenue dans le premier ellipsoïde de Fresnel. Ce sont les pixels dans les «coins» de la fenêtre (par exemple les pixels noirs dans la Figure 6.4 (b) et (c)) qui ne correspondent pas à la projection de points dans l'ellipsoïde. Cela peut donc conduire à surestimer l'atténuation de l'onde sonore. Toutefois, il est très simple d'éviter le problème en construisant une table permettant de savoir quels pixels doivent être pris en compte (cette table indique pour chaque ligne le pixel de départ et de fin entre lesquels on doit effectivement comptabiliser les obstructions et peut être obtenue à partir d'un algorithme de tracé de cercle de Bresenham par exemple). Dans ce cas la valeur de V^λ est simplement donnée par :

$$V^\lambda = \frac{\mathrm{Pix}_{\mathrm{obstacles}}}{\mathrm{Pix}_{Fresnel}},$$

où $\text{Pix}_{\text{obstacles}}$ est le nombre de pixels correspondant à des obstacles *dans l'ellipsoïde* et $\text{Pix}_{Fresnel}$ est le nombre de pixels correspondant à la projection de l'ellipsoïde.

Traitement qualitatif de la transmission

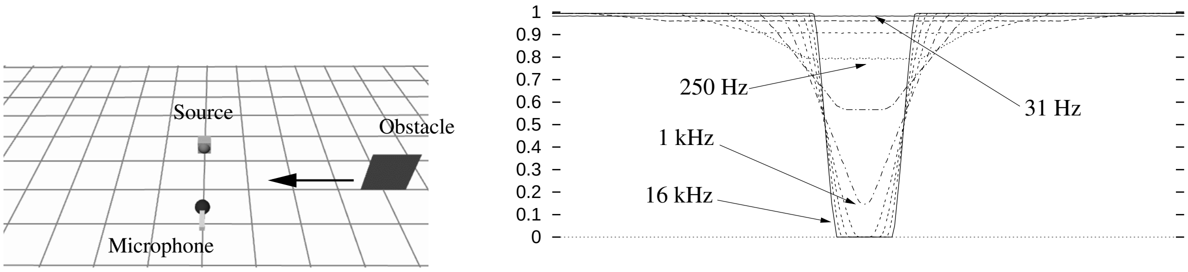
On peut également utiliser ce même principe et ajouter un coefficient permettant de simuler une «transmission» qualitative du son au travers des obstacles. En effet, lors du rendu de la scène, on peut utiliser la couleur de chaque objet comme identificateur de son matériau (on lira donc une valeur différente pour la couleur du pixel en fonction du matériau) et effectuer des calculs séparés pour chacun d'entre eux. On pourra alors attribuer à chaque matériau un coefficient de transmission $T^\lambda \in [0,1]$ dépendant de la fréquence. L'intensité acoustique au point de réception, pourra alors être exprimée par :

$$I^\lambda_{diff} = I^\lambda_0 \sum_{i=1}^{\text{nbMat}} \Bigg(\underbrace{(1-V_i^\lambda)}_{\text{diffraction}} + \underbrace{(V_i^\lambda T_i^\lambda)}_{\text{transmission}} \Bigg),$$

où nbMat est le nombre de matériaux différents dans la scène. On remarquera que, dans cette approche, le coefficient de diffraction dépend uniquement de la géométrie alors que le coefficient de transmission dépend de la géométrie et des matériaux. Le coefficient de diffraction agit comme un terme de visibilité géométrique étendue.

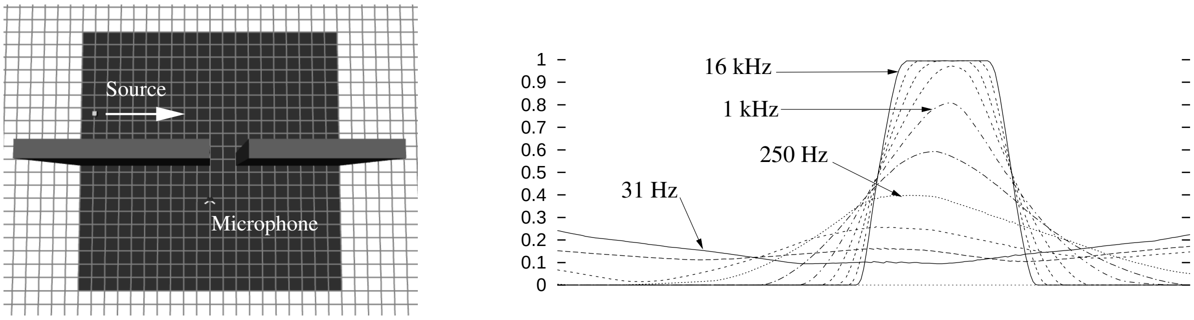
6.2.4 Résultats

Cette méthode a été implémentée sur station Silicon Graphics avec les librairies graphiques *OpenGL* et *OpenInventor* [NDM93, Wer94a, Wer94b], ce qui la rend portable sur une vaste variété de plateformes. Les Figures 6.5 et 6.6 présentent des exemples de calcul des coefficients d'atténuation $(1-V^\lambda)$ pour 10 bandes d'octaves en fonction du temps dans des environnements dynamiques. Les coefficients sont recalculés tout les 1/25ème de seconde pendant le mouvement des objets. Dans le premier exemple, on remarque que le passage de l'obstacle n'affecte pas les basses fréquences mais que le coefficient de visibilité tend vers un coefficient de visibilité géométrique (valant soit 0, soit 1) pour les hautes fréquences Diffraction!«visibilité sonore».



Environnement de test — Coefficient de «visibilité» en fonction du temps

Figure 6.5 - *Évolution de la «visibilité» en fonction du temps lorsqu'un obstacle en mouvement passe entre la source et le récepteur. L'obstacle est une plaque carrée de surface égale à 1 m^2. La source et le récepteur sont distants de 2 m. Les valeurs sont données en bandes d'octaves de 31 Hz à 16 kHz.*



Environnement de test — Coefficient de «visibilité» en fonction du temps

Figure 6.6 - *Évolution de la «visibilité» en fonction du temps lorsqu'une source mobile se déplace devant une «porte» taillée dans un mur. Le mur est haut de 10 m, large de 20 m et l'ouverture est large de 2 m. Les valeurs sont données en bandes d'octaves de 31 Hz à 16 kHz.*

Une implémentation directe utilisant un rendu par bandes d'octaves, permet de calculer les coefficients en temps-réel pour trois bandes d'octaves et des fenêtres de 50×50 pixels. Néanmoins on peut imaginer optimiser le processus en effectuant un rendu unique pour toutes les bandes d'octaves considérées, en choisissant la taille du cône de projection de manière à contenir la plus grande première zone de Fresnel (c'est-à-dire pour la fréquence la plus basse). Pour calculer le coefficient dans une bande d'octave donnée, il faut alors prendre garde à ne compter que les pixels correspondant aux objets contenus dans la zone de Fresnel correspondante. On pourra utiliser pour cela les pixels contenus dans la zone carrée contenant la zone de Fresnel ou bien utiliser là encore une table permettant d'accéder directement les bons pixels. Il faudra également prendre garde à choisir une résolution suffisante au départ pour prendre en compte suffisamment de pixels à haute fréquence.

Auralisation des résultats

À partir des coefficients d'atténuation de l'intensité par bandes d'octaves, on peut reconstruire le spectre en bandes fines de la fonction de transfert correspondante par interpolation (linéaire par exemple). On lui associe alors une *phase minimale* (*cf.* Annexe C.1) pour obtenir la fonction de transfert définitive. On peut remarquer qu'en général ce filtre est un filtre passe-bas puisque les basses fréquences seront moins atténuées que les hautes fréquences, ce qui est conforme à ce qui se passe dans des environnements réels.

6.2.5 Avantages et limitations

Cette méthode offre un moyen simple et rapide pour calculer des effets qualitatifs de la diffraction et de la transmission entre une source et un récepteur ponctuels. Elle est générique dans son principe et permet de traiter des environnements arbitraires (les objets devant être néanmoins polygonalisés) et complètement dynamiques. En outre, la modification subie par le signal sonore reste cohérente avec l'environnement de synthèse puisque qu'elle utilise directement la géométrie des objets. De plus, elle peut être appliquée aux chemins de propagation indirects dus à des réflexions spéculaires et modélisées à l'aide de sources-images, comme nous le verrons par la suite.

Une caractéristique du processus est sa non-symétrie, puisque nous avons choisi arbitrairement de prendre en compte une information de visibilité calculée depuis le point de vue de la source. Cette information n'est évidemment pas la même prise du point de vue du récepteur (dans le cas général). Une possibilité pourrait être d'effectuer deux rendus et de calculer une moyenne des deux coefficients de visibilité obtenus. Néanmoins cela nécessiterait deux fois plus d'opérations graphiques, pour un gain difficile à évaluer puisque la méthode reste essentiellement qualitative. De même, nous avons choisi d'utiliser une projection orthographique alors qu'un projection perspective aurait pu être envisagée. Là encore, un processus de calibration serait sans doute nécessaire pour obtenir les meilleurs résultats d'un point de vue perceptif mais il était difficile d'effectuer une telle étude dans le cadre de ce travail.

Enfin, l'efficacité de la méthode pourrait sans doute être encore accrue en utilisant d'autres modes de rendu graphique comme le mode «*feedback*» d'*Open GL* (qui supprimerai l'étape de calcul de l'image mais nécessiterai d'effectuer le calcul des aires des triangles affichés par le processeur graphique) ou sa nouvelle extension histogramme («*EXT_histogram*» qui permet de compter les pixels affichés directement au moment de la génération de l'image par le processeur graphique) et qui est câblée sur certains modèles de machines Silicon Graphics (cette extension est en outre prévue en standard pour la prochaine version d'*Open GL*).

6.3 Extension : diffraction de Fresnel-Kirchhoff

Le modèle que nous avons décrit dans les paragraphes précédents n'offre néanmoins qu'une solution qualitative au problème de la diffraction d'une onde sonore. Nous proposons une extension semi-quantitative à cette méthode, utilisant la théorie de Fresnel-Kirchhoff que nous avons introduite au paragraphe 3.1.2).

6.3.1 Évaluation de l'intégrale de Kirchhoff à l'aide d'une carte de profondeur

Comme précédemment nous supposerons la source et le récepteur ponctuels. Étant donné un modèle géométrique de l'environnement virtuel, on peut utiliser le rendu câblé pour calculer en temps réel une *carte de profondeur* («Z-buffer») des objets de l'environnement contenus dans un certain volume entre la source et le récepteur. Comme précédemment, on obtient une image des obstacles entre la source et le récepteur mais l'information portée par les pixels n'est plus ici une couleur dépendante de l'objet ou de son matériau, mais la «profondeur» de l'objet en ce point relativement à l'origine du cône de vue (la position de la source dans notre cas) (*cf.* Figure 6.7). Cette carte (ou «tampon») de profondeur est directement accessible lors d'un rendu 3D câblé puisqu'elle est utilisée par le processeur graphique pour ordonnancer l'affichage des objets et éliminer les parties cachées [FVFH90].

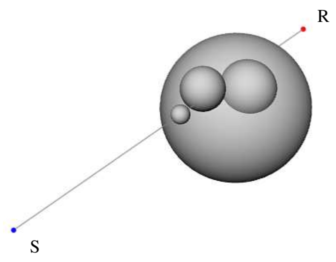
La carte de profondeur nous permet d'identifier immédiatement les objets qui vont bloquer l'onde sonore. On peut, en outre, grâce à l'information de profondeur, reconstruire une approximation 3D de ce domaine diffractant, que nous noterons Σ, en reprojetant chaque pixel de la carte dans l'espace (*cf.* Figure 6.8).

On se place dans le cas d'une source ponctuelle harmonique en régime établi. La pression en champ libre au point de réception situé à la distance ρ_0 de la source est donnée par :

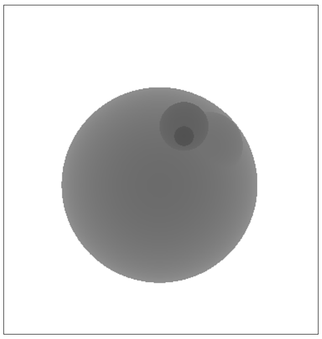
$$\hat{P}_u(t,\rho_0) = \frac{P_o}{\rho_0} e^{i(k\rho-\omega t)} = \hat{P}_u(\rho_0)e^{-i\omega t},$$

où P_o est l'amplitude de la pression émise par la source.

En régime établi, on rappelle (*cf.* paragraphe 3.1.2) que la théorie de la diffraction de Fresnel-Kirchhoff permet d'exprimer la contribution en pression acoustique d'un petit élément différentiel de surface dS du front d'onde à la pression totale (en champ libre) en un point de réception quelconque



(a)



(b)

Figure 6.7 - *Exemple de carte de profondeur calculée à l'aide du rendu 3D câblé. (a) Scène. (b) Carte de profondeur générée. Les nuances de gris plus sombres représentent les points les plus proches de la source.*

situé au delà sous la forme :

$$d\hat{P}(M) = -\frac{P_o e^{ik(\rho+r)}}{4\pi\rho r}\left(ik(1+\mathbf{v}\cdot\mathbf{r}) - \frac{1}{\rho} - \frac{\mathbf{v}\cdot\mathbf{r}}{r}\right), \tag{6.4}$$

où **r** et **v** sont des vecteurs unitaires (voir également les notations de la Figure 6.8).

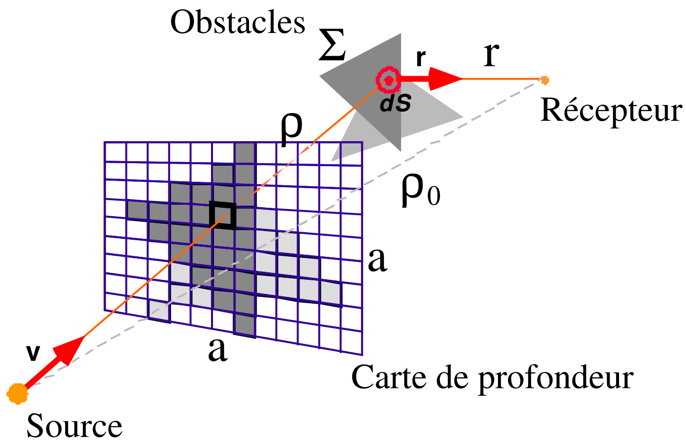


Figure 6.8 - *Évaluation de l'intégrale de Kirchhoff à l'aide d'une carte de profondeur*

On peut donc exprimer la contribution reçue par le récepteur R comme la différence entre la contribution en champ libre et la contribution de toute la surface de front d'onde bloquée par les obstacles. La pression au point de réception R est donc donnée par :

$$\hat{P}(R) = \hat{P}_u(R) - \iint_\Sigma d\hat{P}(R)dS. \tag{6.5}$$

À partir de la carte de profondeur, on peut évaluer l'intégrale de manière discrète pour chaque pixel et la valeur de la pression est alors égale à :

$$\hat{P}(R) = \hat{P}_u(R) - \sum_{i=1}^{\mathrm{Pix}} \sum_{j=1}^{\mathrm{Pix}} V_{ij} \frac{a^2}{\mathrm{Pix}^2} d\hat{P}_{K_{ij}}(R),$$

où a est comme dans la partie précédente la taille du cône de projection, Pix×Pix est la résolution de la carte de profondeur en pixels, $d\hat{P}_{K_{ij}}$ est le contribution différentielle (*cf.* équation 6.4) calculée en K_{ij}, projection inverse en 3D du pixel(i,j) et V_{ij} est la fonction qui vaut 1 si le pixel (i,j) correspond à un objet (*i.e.* est «allumé») et 0 sinon.

La Figure 6.9 récapitule l'algorithme permettant d'effectuer ce calcul.

```
complex FresnelKirchhoffDiffraction( Point S, M; Scene blockers; float λ; int a, Pix)
{
   setupOrthographicCamera(S,M,Pix,a)
   renderScene(blockers)
   Vector Z = S-M
   float ρ0 = norm(Z)
   float k = 2π/λ
   complex Pressure = e^(ikρ0)/ρ0
   for int i = 0 to Pix
       for int j = 0 to Pix
           if isPixelOccluded(i,j)
           {
               Point K = unprojectPixel(i,j)
               Vector R = M - K
               Vector V = K - S
               float r = norm(R)
               float ρ = norm(V)
               normalize(R)
               normalize(V)
               complex deltaOccluded = dP̂(k,r,ρ,R,V)*a²/Pix²
               Pressure = Pressure - deltaOccluded
           }
   return Pressure
}
```

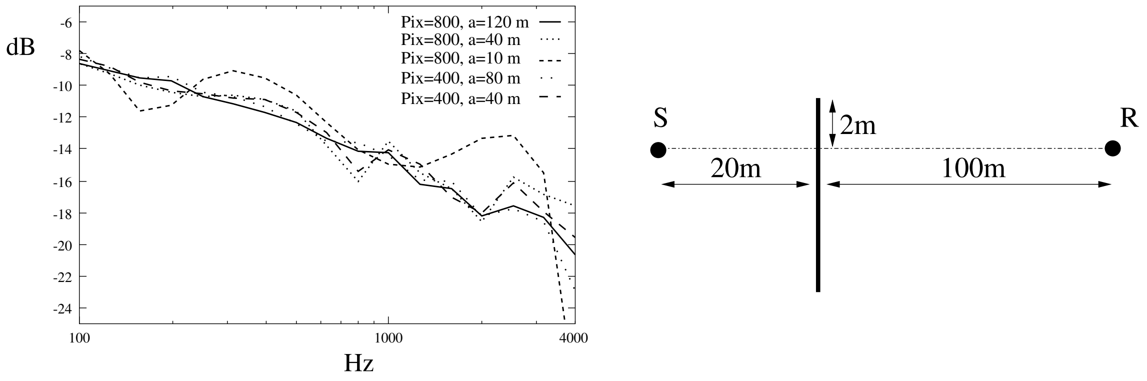
Figure 6.9 - *Pseudo-code pour le calcul du champ diffracté par la théorie de Fresnel-Kirchhoff, au point de réception M, d'une onde sphérique de longueur d'onde λ émise par la source ponctuelle S.*

6.3.2 Influence des paramètres de rendu

Les paramètres contrôlant la qualité de la simulation sont a priori les paramètres de rendus de la carte de profondeur : la taille a du cône de projection considéré et la résolution Pix × Pix en pixels de la carte de profondeur. Comme dans la première partie de ce chapitre, on choisira pour simplifier une projection orthographique, telle que la largeur du cône soit égale à sa hauteur (*cf.* Figure 6.3).

Ces deux paramètres ont une influence importante sur la qualité des résultats. En effet, ils sont en partie reliés à la longueur des chemins obtenus et conditionnent donc un traitement correct des interférences. En effet, si l'on se place dans le cas d'un obstacle plan entre la source et le récepteur, la taille du cône de projection conditionne la longueur maximale des chemins obstrués que l'on va obtenir sur la carte d'ombre. La résolution, elle, va conditionner les déphasages relatifs entre deux chemins voisins. Il faut donc que la résolution soit suffisante pour ne pas perdre des interférences. Il faut également que la taille de la fenêtre soit suffisante pour prendre en compte suffisamment de chemins significatifs. Dans le cas général, la valeur de la profondeur en un point de la carte peut varier arbitrairement et il est moins facile de quantifier l'influence des paramètres de rendu. Toutefois, une taille de fenêtre supérieure à 10λ et une résolution telle que $a/\mathrm{Pix} < \lambda/10$ permettent d'obtenir de bons résultats.

La Figure 6.10 montre un exemple de spectre d'atténuation[3] calculé pour différentes valeurs des paramètres de rendu. On remarque que la taille du cône de rendu est un paramètre important qu'il ne faut pas négliger (comparer la courbes pour $a = 10$ m et Pix = 800 pixels et celles pour $a = 40$ m et Pix = 800/400 pixels respectivement). La courbe obtenue pour $a = 120$ m et Pix = 800 pixels est très proche de celle que l'on pourrait obtenir avec des éléments finis de frontière [Def96].



(a) (b)

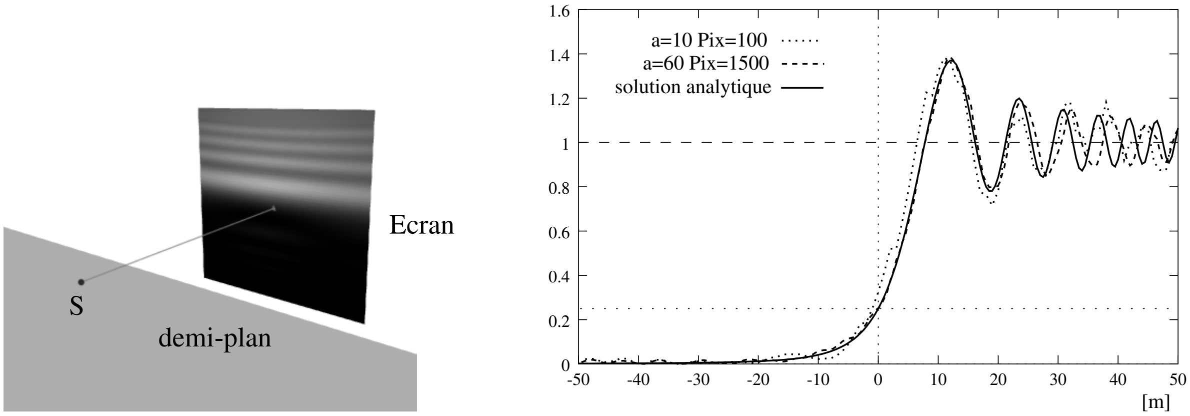
Figure 6.10 - *(a) Spectre d'atténuation (intensité acoustique) dans le cas où l'obstacle est un demi-plan infini, pour différents paramètres (taille a et résolution* Pix*) de la carte de profondeur. (b) Positions relatives de la source, du récepteur et du demi-plan utilisées pour le calcul.*

La Table 6.1 donne des temps d'exécution comparés pour différentes valeurs des paramètres de rendus (pour une station Silicon Graphics Indy R5000 150 MHz). Un temps d'exécution de quelques secondes permet une très bonne approximation des résultats.

Table 6.1 - Temps d'exécution comparés en fonction de la résolution de la carte de profondeur dans le cas d'un demi-plan infini. Calcul effectué à 1000 Hz pour les conditions de la Figure 6.10 dans l'axe de la source.

résolution ($pixels \times pixels$)	**a (m)**	**temps d'exécution (s)**	**erreur (%)**
100x100	10.0	0.1	7
200x200	20.0	0.6	5
400x400	30.0	2.5	1.5
1500x1500	60.0	36	0.2

3. de l'intensité acoustique : $20\log(|\hat{p}|/|\hat{P}_u|)$.



(a) (b)

Figure 6.11 - *(a) Franges de diffraction (carte d'intensité) pour un demi-plan infini calculées à 1000 Hz. La distance de la source au demi-plan est de 20 m. La distance du demi-plan au plan de réception est de 100 m. L'écran de réception est un carré de 100 m de côté. (b) Variations du rapport de l'intensité en présence de l'obstacle sur l'intensité en champ libre, pour différentes valeurs des paramètres de rendu (taille de la fenêtre et résolution), le long d'un axe vertical dans le plan de réception. Les abscisses correspondent à des hauteurs relatives à l'arête du demi-plan.*

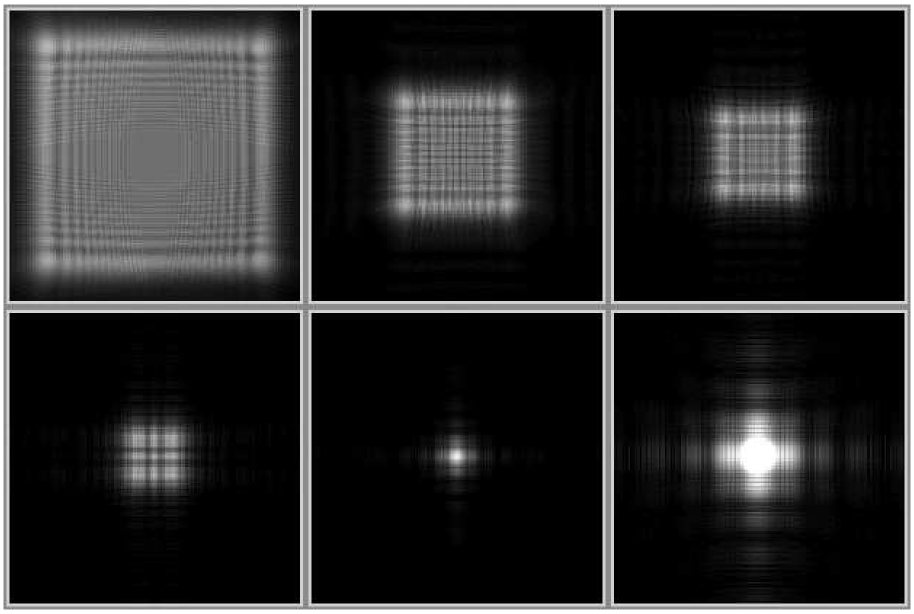
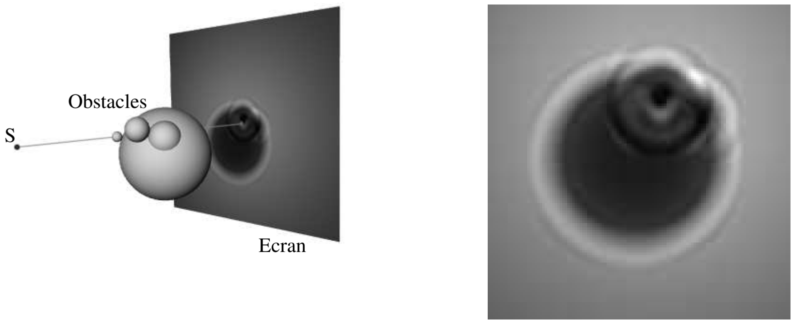


Figure 6.12 - *Figures de diffraction pour une ouverture carrée de taille décroissante. La taille de l'ouverture décroît de la figure supérieure gauche à la figure inférieure droite. On observera la différence entre des figures de diffraction* «de Fresnel» *et* «de Fraunhofer» *(avant dernière image en bas). Dans ce dernier cas, les hypothèses de champ lointain sont valides (le front d'onde sphérique peut être assimilé à une onde plane). La dernière image en bas à droite est un gros plan sur la figure de diffraction de Fraunhofer. Les conditions sont les mêmes que pour la Figure 6.11 (a) et la taille de l'ouverture varie de* 15×15 *m à* 2×2 *m.*

6.3.3 Visualisation de figures de diffraction

L'algorithme précédent peut être utilisé pour visualiser des *figures de diffractions*, qui sont des représentations de la valeur de l'intensité de l'onde sur un plan de réception. Pour cela, on échantillonne le plan de réception en différents point et pour chacun de ces points, on exécute le processus décrit dans la Figure 6.9. Le module de la pression obtenu pour chaque point est proportionnel à l'intensité acoustique incidente. Les Figures 6.11, 6.12 et 6.13 présentent des résultats de cette visualisation pour différentes configurations d'obstacles. On pourra comparer certains résultats avec [Dau96, Hec87]. Le temps de calcul dépend bien sur du nombre de points d'échantillonnage sur le plan de réception, de la taille et de la résolution de la carte de profondeur utilisée et du nombre de pixels bloqués. Il dépend donc fortement de la configuration des obstacles. A titre d'exemple, la figure de diffraction pour les obstacles sphériques de la Figure 6.13, a été calculée en échantillonnant l'écran de réception avec une résolution de 200×200 pixels. Pour chaque pixel, le calcul de l'intensité a été effectué avec une carte de profondeur de 4×4 m à une résolution de 100×100 pixels. Le temps de calcul est de 2 heures (sur station SGI O_2, processeur R5000 à 180 MHz), ce qui donne un temps moyen pour le calcul de diffraction point à point de 0.02 s. (rendu de la carte de profondeur compris).



(a) (b)

Figure 6.13 - *(a) Exemple de figure de diffraction pour un obstacle composé d'un assemblage de sphères. Calcul effectué à 1000 Hz avec des cartes de profondeur de* 4×4 *m à une résolution de* 100×100 *pixels. La distance de la source à l'écran est de 120 m. L'écran est large de 100 m. (b) Gros plan sur la figure de diffraction.*

6.3.4 Tests de validation : atténuation du son par des écrans acoustiques

Afin de vérifier la validité de la méthode, nous avons procédé à des tests dans le cadre de l'atténuation du son par un écran[4]. La méthode a été comparée à des simulations par éléments finis de frontière («Boundary Element Method»)[5]. La Figure 6.14 illustre les différentes configurations étudiées. L'écran est supposé infiniment long et parfaitement absorbant. De plus nous avons étendu notre technique à un sol parfaitement réfléchissant en utilisant un principe de source-image (*cf.* Figure 6.15), approche déjà utilisée par Maekawa [Mae68, Fil94].

4. ces tests ont été réalisés en collaboration avec Y. Gabillet et J. Defrance du Centre Scientifique et Technique du Bâtiment (CSTB) à Grenoble.
5. algorithme développé par P. Jean du CSTB Grenoble [Jea98].

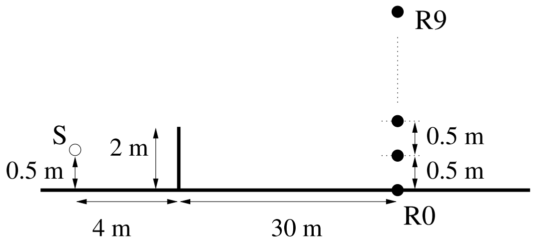
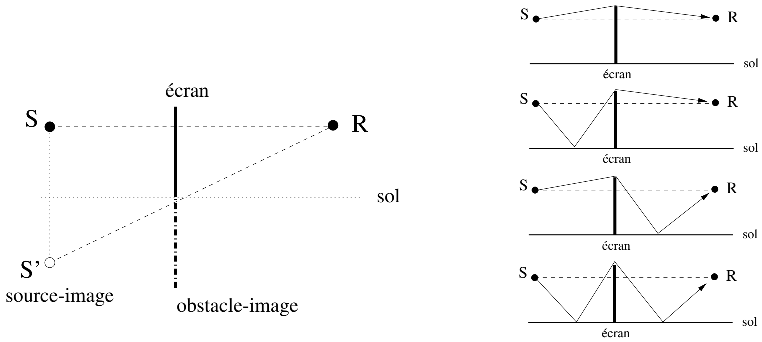
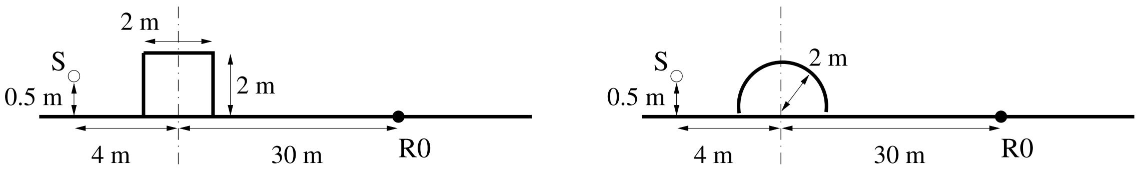


Figure 6.14 - *Configurations de sources et de récepteurs utilisées pour les tests de comparaison avec un calcul par éléments finis de frontière.*



(a) (b)

Figure 6.15 - *(a) Extension à des chemins de propagation indirects. Pour prendre en compte les effets de la réflexion spéculaire sur un sol réfléchissant, on ajoute la contribution d'une source-image. Un «obstacle-image» est ajouté à la scène et une carte de profondeur complémentaire est donc calculée depuis le point de vue de la source-image. (b) Les chemins de propagations équivalents simulés par cette méthode.*



(a) (b)

Figure 6.16 - *Autres configurations d'écrans utilisées pour les tests. (a) Écran droit épais. (b) Écran cylindrique. Les configurations de source et récepteurs restent identiques.*

Les Tables 6.2 à 6.5 regroupent les résultats obtenus pour différentes fréquences (125, 250, 500 et 1000 Hz) : atténuation (rapport de l'intensité sur l'intensité en champ libre), module et phase (entre 0^o et 360^o) du rapport de pressions complexe $\hat{p} = |\hat{p}|e^{i\varphi} = \hat{P}/\hat{P}_u$. Les valeurs des paramètres pour le calcul de la carte de profondeur sont : $a = 3.4$ m et Pix $= 1000$ pixels. On remarque un très bon accord avec les valeurs des BEM, qui s'améliore lorsque la fréquences augmente. En effet, puisque l'on ne considère que la partie de l'obstacle visible depuis la source, on travaille plutôt dans des hypothèses proches de l'optique, donc valable *a priori* en hautes fréquences.

6.3.5 Avantages et limitations

Cette approche permet de calculer de manière semi-quantitative les effets de la diffraction par des obstacles de manière très efficace. Utilisant le rendu câblé, elle n'est pas limitée à des géométries simples. Les tests effectués dans des cas simples ont montré un bon accord avec des techniques de référence comme les éléments finis de frontière, y compris en présence de réflexions secondaires. Toutefois, notre méthode utilise une projection des obstacles et reste donc limitée à la partie «apparente» du domaine diffractant, ce qui constitue une limitation dans le cas ou des obstacles se superposent. D'autres tests, effectués sur des écrans plus complexes (demi-cylindre, écran large) donnent d'ailleurs des valeurs moins satisfaisantes (*cf.* Figure 6.16 et Tables 6.6 et 6.7). Enfin, les obstacles sont supposés ici parfaitement absorbants. Elle permet néanmoins d'obtenir très rapidement et quelle-que-soit la fréquence une approximation intéressante de la solution dans des cas de géométrie complexe où d'autres approches plus rigoureuses peuvent être extrêmement coûteuses à appliquer.

Table 6.2 - Diffraction par un écran plat à 125 Hz

Hauteur (m)	**Att (dB)**		Δ_{Att} (dB)	$(\lvert\hat{\mathbf{p}}\rvert, \varphi)$		$\Delta_{(\lvert\hat{p}\rvert,\varphi)}$
	BEM	Z-buffer		BEM	Z-buffer	
0.0 (R0)	-0.10	-2.88	**2.78**	(0.987, 107.58)	(0.717, 99.57)	**(0.270, 8.01)**
0.5 (R1)	-0.14	-3.21	**3.07**	(0.984, 108.55)	(0.691, 94.14)	**(0.293, 14.41)**
1.0 (R2)	-0.24	-2.66	**2.42**	(0.973, 109.42)	(0.735, 90.52)	**(0.238, 18.90)**
1.5 (R3)	-0.39	-3.57	**3.18**	(0.955, 110.23)	(0.662, 92.11)	**(0.293, 18.12)**
2.0 (R4)	-0.62	-3.61	**2.99**	(0.931, 110.96)	(0.659, 93.12)	**(0.272, 17.84)**
2.5 (R5)	-0.91	-4.12	**3.21**	(0.900, 111.55)	(0.621, 100.43)	**(0.279, 11.12)**
3.0 (R6)	-1.28	-4.81	**3.53**	(0.863, 112.03)	(0.574, 94.48)	**(0.289, 17.55)**
3.5 (R7)	-1.72	-4.36	**2.64**	(0.820, 112.34)	(0.605, 88.98)	**(0.215, 23.36)**
4.0 (R8)	-2.24	-5.35	**3.11**	(0.772, 112.43)	(0.540, 86.76)	**(0.232, 25.67)**
4.5 (R9)	-2.85	-5.45	**2.60**	(0.720, 112.22)	(0.533, 87.66)	**(0.187, 24.56)**

Table 6.3 - Diffraction par un écran plat à 250 Hz

Hauteur (m)	**Att (dB)**		Δ_{Att} (dB)	$(\lvert\hat{\mathbf{p}}\rvert, \varphi)$		$\Delta_{(\lvert\hat{p}\rvert,\varphi)}$
	BEM	Z-buffer		BEM	Z-buffer	
0.0 (R0)	-8.02	-10.65	**2.63**	(0.397, 168.90)	(0.293, 151.51)	**(0.104, 17.39)**
0.5 (R1)	-8.14	-9.06	**0.92**	(0.391, 170.74)	(0.352, 161.83)	**(0.039, 8.91)**
1.0 (R2)	-8.51	-10.38	**1.87**	(0.375, 172.34)	(0.303, 160.13)	**(0.072, 12.21)**
1.5 (R3)	-9.15	-10.81	**1.66**	(0.348, 173.63)	(0.288, 160.73)	**(0.060, 12.90)**
2.0 (R4)	-10.10	-11.12	**1.02**	(0.312, 174.44)	(0.278, 163.76)	**(0.034, 10.68)**
2.5 (R5)	-11.42	-14.26	**2.84**	(0.268, 174.43)	(0.194, 148.17)	**(0.074, 26.26)**
3.0 (R6)	-13.23	-14.77	**1.54**	(0.218, 172.89)	(0.182, 153.52)	**(0.036, 19.37)**
3.5 (R7)	-15.69	-17.47	**1.78**	(0.164, 167.91)	(0.134, 124.95)	**(0.030, 42.96)**
4.0 (R8)	-18.95	-18.35	**-0.60**	(0.113, 155.06)	(0.120, 110.92)	**(-0.007, 44.14)**
4.5 (R9)	-21.85	-20.61	**-1.24**	(0.081, 122.61)	(0.093, 74.83)	**(-0.012, 47.78)**

Table 6.4 - Diffraction par un écran plat à 500 Hz

Hauteur (m)	**Att (dB)**		Δ_{Att} (dB)	$(\lvert\hat{\mathbf{p}}\rvert, \varphi)$		$\Delta_{(\lvert\hat{p}\rvert,\varphi)}$
	BEM	Z-buffer		BEM	Z-buffer	
0.0 (R0)	-10.30	-11.59	**1.29**	(0.305, 172.25)	(0.263, 175.53)	**(0.042, -3.28)**
0.5 (R1)	-10.72	-13.62	**2.90**	(0.291, 175.75)	(0.208, 172.76)	**(0.083, 2.99)**
1.0 (R2)	-12.05	-13.77	**1.72**	(0.250, 178.17)	(0.205, 172.27)	**(0.045, 5.90)**
1.5 (R3)	-14.63	-17.04	**2.41**	(0.186, 178.13)	(0.141, 169.47)	**(0.045, 8.66)**
2.0 (R4)	-19.30	-22.95	**3.65**	(0.108, 169.57)	(0.071, 146.14)	**(0.037, 23.43)**
2.5 (R5)	-25.22	-22.58	**-2.64**	(0.055, 114.21)	(0.074, 72.62)	**(-0.019, 41.59)**
3.0 (R6)	-19.12	-19.34	**0.22**	(0.111, 59.37)	(0.108, 50.73)	**(0.003, 8.64)**
3.5 (R7)	-14.31	-16.69	**2.38**	(0.192, 51.07)	(0.146, 47.55)	**(0.046, 3.52)**
4.0 (R8)	-11.57	-13.54	**1.97**	(0.264, 51.28)	(0.210, 48.98)	**(0.054, 2.30)**
4.5 (R9)	-10.02	-10.74	**0.72**	(0.315, 53.60)	(0.290, 55.47)	**(0.025, -1.87)**

Table 6.5 - Diffraction par un écran plat à 1000 Hz

Hauteur (m)	**Att (dB)**		Δ_{Att} (dB)	$(\lvert\hat{\mathbf{p}}\rvert, \varphi)$		$\Delta_{(\lvert\hat{p}\rvert,\varphi)}$
	BEM	Z-buffer		BEM	Z-buffer	
0.0 (R0)	-12.30	-12.95	**0.65**	(0.243, 78.91)	(0.225, 61.57)	**(0.018, 17.34)**
0.5 (R1)	-14.14	-14.63	**0.49**	(0.196, 85.86)	(0.186, 67.61)	**(0.010, 18.25)**
1.0 (R2)	-22.40	-24.26	**1.86**	(0.079, 85.45)	(0.062, 77.78)	**(0.017, 7.67)**
1.5 (R3)	-22.03	-22.55	**0.52**	(0.079, 302.33)	(0.075, 302.08)	**(0.004, 0.25)**
2.0 (R4)	-14.02	-14.67	**0.65**	(0.199, 302.67)	(0.185, 303.88)	**(0.014, -1.21)**
2.5 (R5)	-12.21	-12.19	**-0.02**	(0.245, 309.49)	(0.245, 303.34)	**(0.000, 6.15)**
3.0 (R6)	-13.95	-15.34	**1.39**	(0.201, 314.58)	(0.171, 304.24)	**(0.030, 10.34)**
3.5 (R7)	-20.77	-22.43	**1.66**	(0.091, 301.66)	(0.076, 282.11)	**(0.015, 19.55)**
4.0 (R8)	-20.02	-20.40	**0.38**	(0.099, 207.98)	(0.096, 204.59)	**(0.003, 3.39)**
4.5 (R9)	-13.41	-14.42	**1.01**	(0.213, 198.06)	(0.190, 194.97)	**(0.023, 3.09)**

Table 6.6 - Diffraction par un écran cylindrique à 1000 Hz

Hauteur (m)	**Att (dB)**		Δ_{Att} (dB)	$(\lvert\hat{\mathbf{p}}\rvert, \varphi)$		$\Delta_{(\lvert\hat{p}\rvert,\varphi)}$
	BEM	Z-buffer		BEM	Z-buffer	
0.0 (R0)	-17.00	-18.29	**1.29**	(0.141, 58.49)	(0.122, 153.33)	**(0.019, -94. 84)**
0.5 (R1)	-18.80	-20.93	**2.13**	(0.115, 65.01)	(0.090, 150.55)	**(0.025, -85. 54)**
1.0 (R2)	-26.62	-29.08	**2.46**	(0.047, 62.44)	(0.035, 200.69)	**(0.012, -138. 25)**
1.5 (R3)	-27.12	-20.39	**-6.73**	(0.044, 287.60)	(0.095, 326.72)	**(-0.051, -39. 12)**
2.0 (R4)	-18.98	-14.71	**-4.27**	(0.112, 283.55)	(0.183, 0.85)	**(-0.071, 282. 70)**
2.5 (R5)	-17.06	-14.47	**-2.59**	(0.140, 288.08)	(0.188, 19.64)	**(-0.048, 268. 44)**
3.0 (R6)	-18.57	-16.88	**-1.69**	(0.117, 289.52)	(0.143, 33.68)	**(-0.026, 255. 84)**
3.5 (R7)	-23.76	-30.53	**6.77**	(0.065, 269.63)	(0.030, 53.91)	**(0.035, 215. 72)**
4.0 (R8)	-22.98	-19.87	**-3.11**	(0.071, 201.26)	(0.101, 217.65)	**(-0.030, -16. 39)**
4.5 (R9)	-17.85	-16.39	**-1.46**	(0.128, 184.88)	(0.151, 231.68)	**(-0.023, -46. 80)**

Table 6.7 - Diffraction par un écran droit épais à 1000 Hz

Hauteur (m)	**Att (dB)**		Δ_{Att} (dB)	$(\lvert\hat{\mathbf{p}}\rvert, \varphi)$		$\Delta_{(\lvert\hat{p}\rvert, \varphi)}$
	BEM	Z-buffer		BEM	Z-buffer	
0.0 (R0)	-17.43	-14.71	**-2.72**	(0.134, 91.25)	(0.184, 83.44)	**(-0.050, 7. 81)**
0.5 (R1)	-19.50	-18.85	**-0.65**	(0.106, 96.91)	(0.114, 111.25)	**(-0.008, -14. 34)**
1.0 (R2)	-29.05	-32.58	**3.53**	(0.035, 83.05)	(0.023, 310.15)	**(0.012, -227. 10)**
1.5 (R3)	-24.33	-18.16	**-6.17**	(0.061, 320.47)	(0.123, 295.39)	**(-0.062, 25. 08)**
2.0 (R4)	-18.19	-14.49	**-3.70**	(0.123, 317.90)	(0.188, 321.00)	**(-0.065, -3. 10)**
2.5 (R5)	-17.30	-13.57	**-3.73**	(0.136, 320.48)	(0.209, 337.86)	**(-0.073, -17. 38)**
3.0 (R6)	-20.17	-16.82	**-3.35**	(0.098, 315.34)	(0.144, 18.47)	**(-0.046, 296. 87)**
3.5 (R7)	-24.84	-21.05	**-3.79**	(0.057, 265.27)	(0.088, 95.93)	**(-0.031, 169. 34)**
4.0 (R8)	-19.42	-15.41	**-4.01**	(0.106, 218.31)	(0.169, 176.03)	**(-0.063, 42. 28)**
4.5 (R9)	-16.19	-11.20	**-4.99**	(0.155, 212.47)	(0.275, 210.61)	**(-0.120, 1. 86)**

Chapitre 7

Intégration du son dans une plateforme d'animation 3D interactive : le système *NARSYS*

Ce chapitre présente la deuxième contribution de ce travail de thèse sous la forme d'un système, que nous avons baptisé *NARSYS* (Natural Audio Rendering SYStem) permettant le rendu interactif de l'image et du son dans un système d'animation 3D. Basé sur un modèle de sources-images, l'originalité de ce système est d'offrir une intégration complète du rendu simultané son-image, permettant de garantir la cohérence des pistes audio et vidéo et d'utiliser les capacités de cartes graphiques spécialisées pour des calculs qualitatifs de diffraction et de transmission sonore.

7.1 Modélisation de la scène sonore

Nous commençons par décrire ici, les différents éléments nécessaire à la définition d'une «scène sonore», c'est à dire l'environnement virtuel acoustique qui sera simulé. Nous proposons une modélisation de la scène sonore à l'aide de quatre types de «primitives» :

– les sources sonores,

– les microphones,

– les objets de l'environnement (donnés sous forme de polygones) qui vont être utilisés pour le calcul de «visibilité» du son que nous appellerons des «obstacles».

– les objets de l'environnement (également donnés sous forme de polygones) qui vont être utilisés pour le calcul de la réflexion du son : les «réflecteurs».

On remarque que l'on retrouve transposés les éléments classiques d'une scène 3D «visuelle» : sources lumineuses, caméras, modèles polygonaux des objets.

7.1.1 Sources et microphones

Les sources et les récepteurs, que nous appellerons ici microphones dans un contexte de prise de son, seront supposés ponctuels. Il sont caractérisés par leur position au cours du temps et un ensemble de réponses impulsionnelles normalisées (valeurs à 1m par exemple) dépendant de la direction d'incidence, ce qui permet d'utiliser des réponses mesurées sur de vrais transducteurs (enceintes, microphones, oreilles). La source et le microphone sont, en outre, caractérisés par un gain permettant d'en adapter la puissance ou la sensibilité. La réponse impulsionnelle dans une direction donnée est alors le produit entre la réponse impulsionnelle normalisée et ce gain. La position des sources et des microphones peut être modifiée en cours de simulation. En outre, plusieurs sources peuvent être combinées en imposant leurs positions relatives les unes par rapport aux autres. Ceci permet d'approcher à l'aide de plusieurs sources le comportement de sources d'émission complexes (groupes d'enceintes, sources linéiques, etc.). Il en va de même pour les microphones, ce qui permet de simuler des dispositifs de prise de son comme les «*paires stéréophoniques*[1]» par exemple [MR92, HW94].

On peut également définir des sources «d'ambiance» qui ne seront pas simulées mais dont les contributions seront directement additionnées a la bande son généré à la fin du processus de rendu sonore. Dans le cadre de génération de bandes son pour l'animation, cela permet d'intégrer des effets ou une musique d'ambiance de manière simple.

Textures sonores

A chaque source est également associé un signal émis. Ce signal est constitué d'une *texture sonore*. Une texture sonore est un signal échantillonné représentant le son «brut» (en champ libre) émis par la source. On ne s'intéressera pas ici à l'aspect synthèse de la texture sonore, comme ont pu le faire Takala et Hahn avec leurs «timbre-trees» [TH92]. Il est pourtant clair qu'obtenir un son «brut» n'est pas un objectif simple à réaliser du point de vue du dispositif d'acquisition à mettre en œuvre et que l'aspect synthèse constitue un point intéressant. Néanmoins, on préfère en général construire un modèle à partir d'une base d'enregistrements pour des questions de réalisme. La combinaison de textures échantillonnées parait donc être une option plus attractive comme base pour la génération de sons plus complexes. Dans notre modèle plusieurs textures peuvent être combinées à partir d'un même source en utilisant différents paramètres : retards relatifs, gains relatifs et hauteurs («pitch») relatives. La combinaison de plusieurs textures sur plusieurs sources permet de créer des sources complexes tant du point de vue du signal émis que de son rayonnement.

7.1.2 Réflecteurs et obstacles

Nous avons choisi de séparer les objets de la scène sonore en deux groupes : les réflecteurs et les obstacles. Dans les deux cas, les objets sont décrits sous forme polygonale et leur position peut évoluer au cours du temps. Les polygones des réflecteurs seront utilisés pour générer des sources-images afin de traiter les réflexions spéculaires du son. Les polygones des obstacles seront utilisés pour le calcul de diffraction et transmission du son. En toute rigueur, il n'y a aucune raison de différencier les deux groupes d'objets. Toutefois, cela offre des possibilités de contrôle supplémentaire (on pourra utiliser uniquement les murs d'une pièce comme réflecteurs et les autres objets, le mobilier par exemple, uniquement comme obstacles). Il faut noter ici que les polygones utilisés pour le rendu du son peuvent être différents de ceux utilisés pour le rendu visuel. En général, on pourra utiliser une version simplifiée du modèle visuel (en particulier pour les réflecteurs ce qui permet de limiter le nombre de sources-images), avec une approche de type niveaux de détail géométriques (on conserve alors la cohérence avec le modèle visuel) [RPV93].

1. assemblage de deux microphones monophoniques pour la prise de son stéréophonique très couramment utilisé en radiotélédiffusion.

Matériaux

Les réflecteurs et obstacles portent également des informations concernant leurs matériaux : respectivement filtre de réflexion et de transmission. Nous avons choisi ici de travailler avec des filtres en bandes fines pour plus de généralité. Dans le cas de valeurs données en bandes d'octaves, il faudra les convertir en bandes fines. Cela pourra être fait, suivant le cas à différents stades du processus de rendu ; nous donnerons quelques détails sur ce point dans la section 7.5.4.

7.2 La boucle d'animation et la synchronisation audio-vidéo

Nous décrivons à présent l'architecture globale de notre système de rendu audio-vidéo (*cf.* Figure 7.1). Il peut être séparé en trois parties :

– *la boucle de simulation* qui simule le mouvement des objets et met à jour les paramètres nécessaires au rendu vidéo et audio.

– *le moteur de rendu vidéo* qui permet de produire l'image. Dans notre cas cela sera effectué par la carte graphique spécialisée des stations de travail (Silicon Graphics dans notre cas). Nous ne détaillerons pas ce point.

– *le moteur de rendu audio* synchronisé avec la boucle de simulation qui va effectuer la synthèse des échantillons sonores.

La boucle d'animation simule le mouvement des objets en utilisant un pas de temps de simulation dt_{simu} (ce pas de temps peut être adaptatif dans le cas où le mouvement doit être simulé par des modèles physiques). A chaque pas de temps dt_{image} (généralement à 25 Hz), les positions courantes des objets sont utilisées pour modifier la géométrie de la scène 3D à afficher et pour calculer les paramètres de rendu sonore. La bande son est ainsi synchronisée à l'image près, ce qui est suffisant pour les applications qui nous intéressent. C'est d'ailleurs également le cas pour la plupart des supports vidéo et cinématographiques. La géométrie est alors envoyée au processeur graphique pour le rendu d'une image alors que les paramètres de rendu sonore sont envoyés au moteur de rendu sonore pour la synthèse des échantillons audio correspondant à ce pas de temps. Le rendu audio est effectué aux fréquences d'échantillonnage classiques $1/dt_{audio}$ (à 8, 11.25, 16, 32, 44.1 ou 48 kHz). On générera donc le signal audio par blocs de dt_{image}/dt_{audio} échantillons.

Séquencement et gestion d'événements sonores

Le système permet à chaque pas de temps dt_{image} de traiter un certain nombre d'événements sonores, comme le déclenchement ou l'arrêt d'une source ou la réponse à une collision entre objets par exemple. Les paramètres des textures sonores peuvent alors être modifiés à partir des données du mouvement, vitesse, position voire des données physiques comme le module des forces de contact, l'énergie de la collision (dans le cas d'une simulation par modèles physiques par exemple).

7.3 Traitement de la réverbération sonore

Nous avons choisi de traiter les réflexions du son par un modèle de sources-images , bien adapté aux environnements dynamiques. Pour chaque polygone réflecteur spécifié, chaque source et chaque microphone (*cf.* Figure 7.3 et 7.4), on commence par construire les sources images associées jusqu'à l'ordre

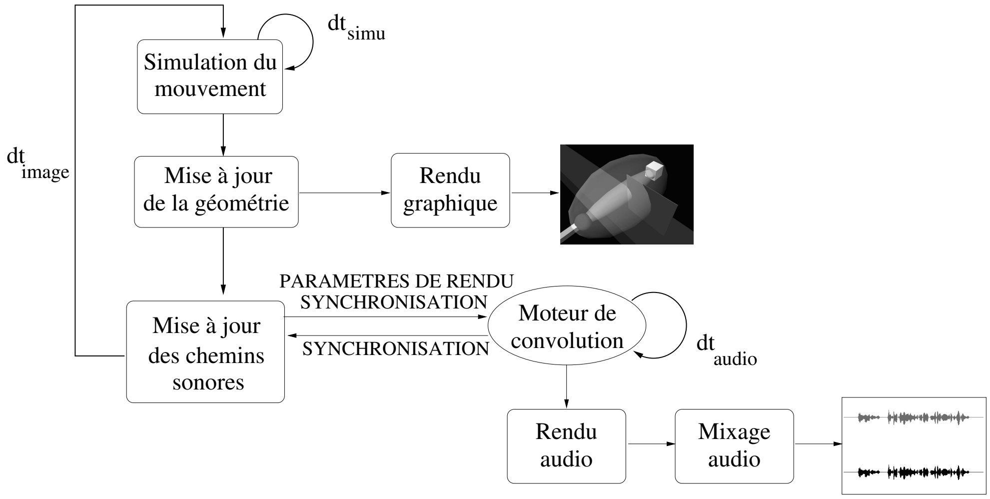


Figure 7.1 - *Boucle d'animation pour le rendu intégré audio-vidéo. Le mouvement des éléments dans la scène est simulé par pas de temps adaptatifs* dt_{simu}*. Tout les pas de temps* dt_{image}*, la scène 3D est rendue et une image est générée. Les paramètres de rendu des chemins sonores sont également mis à jour et transmis au moteur de rendu sonore synchrone qui génère les échantillons sonores à la fréquence d'échantillonnage audio* $(1/dt_{audio})$ *par blocs de taille* (dt_{image}/dt_{audio})*.*

voulu (suivant la qualité désirée de la simulation ou la puissance de calcul disponible). Pour chaque couple source-microphone, on va alors définir un «*chemin-sonore*», brique de base de notre système de rendu, qui représente le canal acoustique associé.

7.3.1 Chemins sonores

Les chemins sonores contiennent toute l'information nécessaire au rendu du son pour le canal acoustique source-microphone considéré. Dans le cas d'environnements dynamiques cette information est composée de deux éléments (*cf.* Figure 7.2)

– le *délai de propagation* du son de la source jusqu'au microphone.[2] Ce délai est nécessaire pour rendre l'effet Doppler, dont nous avons vu qu'il est perceptiblement important. Cela sera effectué en *ré-échantillonnant* la texture sonore de la source, puisque le délai de propagation n'est pas en général un multiple de la fréquence d'échantillonnage de ce signal (*cf.* paragraphe 1.4.2). On peut faire un parallèle avec l'image de synthèse en rapprochant cette opération d'une projection de la texture dans l'espace du microphone.

– la *fonction de transfert* du canal acoustique associé prenant en compte les effets de filtrage de la source, du microphone, du réflecteur, des obstacles et du milieu de propagation. Cette fonction de transfert permettra le rendu proprement dit de la texture (un équivalent du calcul de l'illumination locale dans le cadre de la synthèse d'image).

2. en toute rigueur il faudrait prendre en compte deux délais de propagation, l'un avant la réflexion entre la source et le réflecteur, l'autre après la réflexion entre le réflecteur et le microphone. En effet, la fréquence du son étant modifiée par effet Doppler avant la réflexion, le filtrage du matériau varie également. Dans notre cas, nous n'en tiendrons pas compte. On pourra se référer à [Str98] pour plus de détails à ce propos.

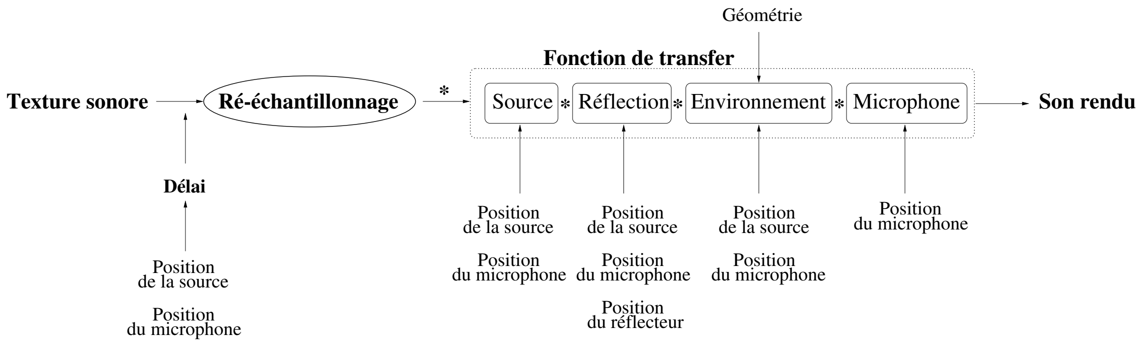


Figure 7.2 - *Un chemin sonore modélise le canal acoustique entre une source, réelle ou virtuelle et le microphone. Il est composé de deux informations principales : le délai de propagation de la source au micro (permettant de simuler l'effet Doppler) et la fonction de transfert du canal acoustique. Le rendu d'un chemin consistera donc en deux étapes : une étape de ré-échantillonnage de la texture en fonction du délai de propagation et un étape de convolution par la fonction de transfert du canal acoustique associé au chemin.*

7.3.2 Construction des sources image et des chemins sonores

Les chemins sonores sont construit récursivement et stockés sous forme d'arbre en fonction de l'ordre de réflexion de la source correspondante, ce qui permettra ensuite d'éliminer rapidement les sources invalides pour les ordres de réflexion supérieurs (*cf.* Figure 7.4).

```
struct SoundPath
{
    SoundPathList*sons
    SoundPath     *father
    ImageSource   *source
    Microphone    *mic
    double        delay
    Filter        *environnementFilter
    Filter        *reflectionFilter

    double        oldDelay
    Filter        *oldFilter
    Filter        *newFilter
    Sample        *oldSamples
    Sample        *newSamples
}
```

```
struct ImageSource
{
    Reflector*reflectionPolygon
    Filter    *sourceFilter
    Position  position
    Texture   *sourceSignal
}
```

Figure 7.3 - *Structures de données pour les chemins sonores et les sources-images. Les types Filter et Sample désignent des tableaux d'échantillons représentant respectivement les coefficient du filtre et les échantillons du signal sonore. Le type Texture désigne une texture sonore, c'est à dire un échantillon «brut» muni de ses paramètres de contrôle. Le type Reflector représente un polygone réfléchissant muni de ses propriétés de réflectivité. Enfin le type Position désigne une position au sens d'un couple (translation,orientation).*

```
void buildPaths(SoundPath path,Reflector r,int order)
{
    if (order == 0) return
    ImageSource s = new ImageSource
        s→signal = path→source→signal
        s→reflectionPolygon = r
        s→sourceFilter = path→source→sourceFilter
    SoundPath p = new  SoundPath
        p→father = path
        p→source = s
        p→microphone = path→mic
        p→reflectionFilter = r→material→transferFunction
    path→sons.add(p)
    forall reflect in Reflectors
    {
        if (reflect==r) continue
        buildPaths(p,reflect,order-1)
    }
}

forall src in Sources
    forall mic in Microphones
    {
        SoundPath p = new  SoundPath
        p→father = NULL
        p→source = src
        p→microphone = mic
        forall reflect in Reflectors
            buildPaths(p,reflect,MAXORDER)
    }
```

Figure 7.4 - *Pseudo-code pour la construction des chemins sonores.*

7.4 Mise à jour des chemins sonores

Au cours de la simulation, les différents chemins sonores vont être modifiés. En particulier, certains ne devrons pas être rendus puisque la source-image correspondante sera devenue invalide. A chaque pas de simulation dt_{image}, il faut donc retrouver le sous-ensemble des chemins valides à transmettre au moteur de rendu et mettre à jour leurs paramètres : délai de propagation et différents filtres (*cf.* Figure 7.5). Nous allons à présent détailler ces différentes opérations.

7.4.1 Détection des chemins invalides et visibilité

La détection des chemins invalides se fait de manière simple en testant si le rayon source-microphone intersecte le polygone réflecteur (*cf.* Figure 7.5). Comme les chemins sont parcourus récursivement par ordre de réflexion croissants, ce test est suffisant pour garantir la validité de la source-image associée.

```
void updatePath(SoundPath p,int order)
{
    if (order == 0) return
    if (p→father!=NULL)
        p→source→position = mirror(father→source→position, p→source→reflectionPolygon)
    Ray r = new Ray (p→source→position.translation,p→microphone→position.translation)
    if r→intersect(p→source→reflect)
    {
        updateDelay(p)
        updateMicrophoneFilter(p)
        updateSourceFilter(p)
        updateReflectionFilter(p)
        updateEnvironmentFilter(p,ObstaclesScene scene)
        renderEngine.addPath(p)
        forall path in p→sons
            updatePath(path,order-1)
    }
}


renderEngine.clearAllPaths()
forall path in SoundPath
    updatePath(path,MAXORDER)
```

Figure 7.5 - *Pseudo-code pour la mise à jour des chemins sonores. Les chemins sonores valides sont identifiés et leurs paramètres mis à jour avant d'être transmis au moteur de rendu. On pourrait intégrer ici un test sur l'«importance» du chemin d'un point de vue énergétique avant de l'envoyer au moteur de rendu, ce qui permettrait d'alléger les calculs en élaguant l'arbre des réflexions [MOD96].*

Pour évaluer la visibilité nous proposons d'utiliser la méthode décrite au chapitre 6.2. On peut également utiliser cette même technique à partir d'une source image mais il faut alors veiller à considérer également comme obstacles potentiels, les réflexions des obstacles réels par rapport au plan du polygone réflecteur (*cf.* Figure 7.6). Ceci peut être effectué très facilement durant le rendu des obstacles par le processeur graphique. Il suffit de rendre une scène constituée :

– de tout les obstacles situés devant le plan de réflexion. Ceci peut être très facilement réalisé en rendant tous les obstacles mais en insérant un plan de coupe («*clipping plane*»).

– de leurs symétriques par rapport au plan de réflexion. Il suffit donc d'ajouter une matrice de transformation à la scène précédente et de la rendre à nouveau[3].

La méthode présentée au chapitre 6.2 peut donc être étendue pour une source-image sans coût de calcul supplémentaire, les transformations géométriques étant effectuées par le processeur graphique.

En pratique, nous avons implémenté cette technique en utilisant la librairie graphique *OpenInventor*. Cette librairie permet d'effectuer le rendu d'une scène structurée sous forme d'arbre [Wer94a]. Les

3. Il faut également veiller à inverser l'orientation des polygones si on utilise le mode d'élimination des faces arrières («*backface culling*»).

feuilles de cet arbre sont les primitives géométriques et ses nœuds des opérations de transformation géométriques, de groupement, de modification des attributs. Une telle représentation est très appropriée ici, puisque l'on peut partager les mêmes nœuds de transformation entre plusieurs arbres, en particulier celui du modèle visuel et celui du modèle acoustique. La cohérence du mouvement des primitives dans les deux modèles peut ainsi être maintenue à peu de frais. On notera que le calcul du filtre de visibilité pour un pas de temps fixé utilise les positions des objets calculées pour ce même pas de temps. On suppose donc la position des obstacles invariante durant le temps de propagation de la source au récepteur. Cette hypothèse est raisonnable dans la mesure où les objets se déplace à une vitesse largement infrasonique, ce qui est le cas dans les applications qui nous intéressent.

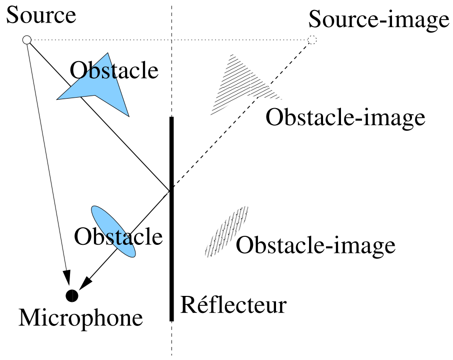


Figure 7.6 - *La visibilité pour une source image peut être évaluée en considérant comme obstacles potentiels les objets de la scène réelle ainsi que leurs images-miroir par rapport au plan de la réflexion.*

7.4.2 Calcul du délai de propagation

En toute rigueur, le délai de propagation devrait être recalculé pour chaque échantillon de son généré. Toutefois, un tel calcul serait très coûteux et il semble suffisant de le calculer par interpolation linéaire entre des délais «clé», évalués pour chaque dt_{image} [NT95, TH92, Str98]. La détermination de ces délais clé reste néanmoins non triviale. Nous avons comparé ici deux méthodes : la première utilise un processus itératif convergent similaire à celui utilisé dans [NT95], nécessitant de stocker les positions passées de la source[4]. Le second par évaluation directe du temps de propagation entre la source et le microphone au temps considéré et très peu coûteux et ne nécessite aucun stockage particulier autre que la position courante des sources et microphones. On aurait pu également utiliser ici une approche basée sur un développement limité de l'expression du délai de propagation en fonction du temps comme proposé dans [Str98]. Nous avons toutefois préféré évaluer la méthode la plus simple possible.

Calcul itératif et calcul direct

On rappelle que l'on peut exprimer le délai de propagation $\tau(t)$ entre une source S et un microphone M mobiles par l'équation suivante :

$$\tau(t) = \frac{1}{c}\|\mathbf{M}(t) - \mathbf{S}(t - \tau(t))\|.$$

4. on peut également utiliser de manière satisfaisante une extrapolation linéaire à partir du dernier vecteur vitesse connu de la source.

On peut appliquer cette formule à chaque pas de temps dt_{image} en considérant, par exemple, que la position $\mathbf{M_{dt}}$ du microphone est fixe durant cette durée. On obtient alors :

$$\tau_{dt} = \frac{1}{c}\|\mathbf{M_{dt}} - \mathbf{S}(t - \tau_{dt}))\|.$$

Dans les hypothèses où la vitesse de la source est inférieure à la vitesse du son c, on peut calculer la valeur du délai de propagation en utilisant le processus itératif décrit dans la Figure 7.7. La preuve de la convergence de cette méthode peut être trouvée en Annexe A.5. Entre une et cinq itérations sont nécessaires en pratique pour déterminer de manière satisfaisante le délai (à 1.10^{-6} près) à une fréquence de re-calcul de 25 Hz, dans le cas de trajectoires où la vitesse relative source/récepteur varie continûment, ce qui reste acceptable. En cas de discontinuité, due à une collision par exemple, plus d'itérations pourront être nécessaires.

```
void updateDelay(SoundPath p)
{
    p→oldDelay = p→newDelay
    double d = p→newDelay
    double old = -1
    int iter = 0
    Vector mic = p→microphone→getPosition(t)
    while (|old−d| >EPSILON) and (iter<MAXITER)
    {
        Vector src = p→source→getPosition(t-d)
        old = d
        d = 1/C * || src-mic ||
        iter = iter + 1
    }
    p→newDelay = d
}
```

Figure 7.7 - *Calcul itératif du délai de propagation entre une source et un microphone mobiles. A $t = 0$, le processus est initialisé avec la valeur du délai courant entre la source et le microphone. A chaque pas de temps dt_{image}, on ré-utilise, comme valeur de départ, l'ancienne valeur du délai ce qui accélère la convergence du processus.*

Une autre manière plus directe est de définir le délai de propagation tout simplement comme le délai de propagation entre la position courante de la source et du microphone.

Nous avons comparé les deux approches dans le cadre d'une expérience simple : une source mobile (se déplaçant en ligne droite) passant devant un microphone fixe à différentes distances et à différentes vitesses. Les Figures 7.8 et 7.9 illustrent le résultat de ce test.

La Figure 7.8 illustre un comparatif des délais obtenus au cours du temps par les deux méthodes pour une source se déplaçant à 50 et 100 km/h à une distance de 50, 10 et 4 m. Comme on pouvait s'y attendre, l'écart entre les deux résultats est d'autant plus important que la vitesse relative source-microphone est importante. Toutefois, on constate que les deux méthodes offrent des résultats très similaires. Le très léger décalage temporel observé entre les deux courbes n'est pas perceptiblement audible.

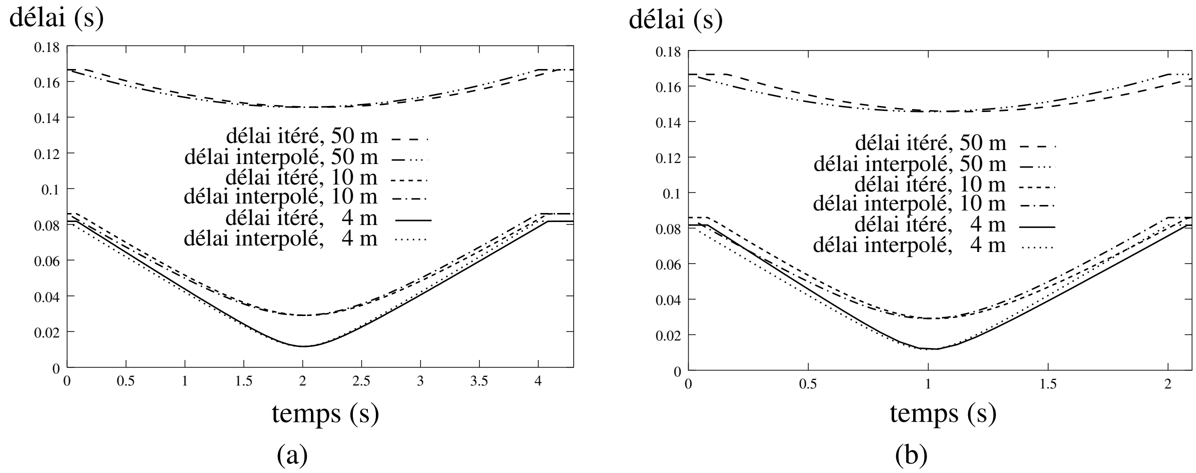


Figure 7.8 - *Comparaison entre délais de propagation calculés itérativement ou par interpolation directe, pour des distances minimales entre source et microphone de 4, 10 et 50 m. La fréquence de rafraîchissement des délais est 25 Hz. (a) La vitesse de la source est 50 km/h. (b) La vitesse de la source est 100 km/h.*

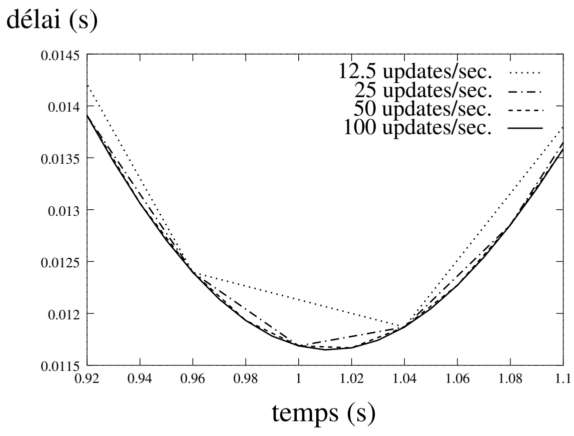
La Figure 7.9 compare les délais obtenus pour différentes fréquences de re-calcul (valeurs de dt_{image}) allant de 12,5 Hz à 100 Hz. La source passe ici en ligne droite à 4 m devant le récepteur (toujours fixe) et à une vitesse de 100 km/h. On obtient un variation très importante de la fréquence et de l'intensité du signal reçu. On observe mieux sur cette figure le décalage entre les deux courbes (le délai obtenu par interpolation directe étant «en avance» d'un centième de seconde environ). L'échantillonnage correct du délai de propagation reste le critère important. En effet, comme on peut le voir sur la figure, 12,5 mises à jour par seconde ne sont pas suffisantes ici pour échantillonner le «creux» de la courbe. Cela résulte en des artefacts clairement audibles sur des signaux harmoniques. Dès que le taux de rafraîchissement devient supérieur à 25 Hz (inclus), on observe un rendu sans artefact perceptif notable. Il faut noter ici que, dans le cas de signaux au contenu fréquentiel riche, les artefacts sont également moins audibles.

Néanmoins, pour une qualité de simulation optimale, on pourrait imaginer utiliser un pas de temps variable, comme dans le cas de simulations de mouvement par modèles physiques, afin de recalculer plus fréquemment le délai de propagation lorsqu'il varie beaucoup ; par exemple, en divisant par deux le pas de temps de simulation lorsque la variation du délai dépasse un certain seuil. Plusieurs choix sont alors possibles. En particulier on peut estimer la quantité $|d\tau/dt| \approx |\frac{\tau_t - \tau_{t-dt}}{dt}|$ qui mesure la variation du délai en fonction du temps. Comme on le remarque sur la Figure 7.10 (a), cette quantité ne permet toutefois pas une subdivision correcte puisque elle ne tend pas à diminuer lorsque l'on divise le pas de temps.

Une autre quantité parait bien plus appropriée : $|d\tau/\tau| \approx \frac{1}{2}|\frac{\tau_t - \tau_{t-dt}}{\tau_t + \tau_{t-dt}}|$, représentant l'erreur relative commise sur l'évaluation du délai. La Figure 7.10 (b) montre les variations de cette quantité. On remarque qu'elle diminue lorsque l'on divise le pas de temps. De plus, elle admet des *maxima* au moment où les variations du délai sont importantes, quand la source va passer devant le microphone (et juste après). On va donc subdiviser le pas de temps judicieusement à ces instants là. L'autre avantage de ce critère est qu'il va rester «indépendant» de la simulation.

7.4.3 Mise à jour des filtres

Nous abordons à présent le problème de la mise à jour des filtres. Dans le cas des filtres de la source et du microphone ils sont directement mis à jour en fonction de leurs positions et orientations relatives et



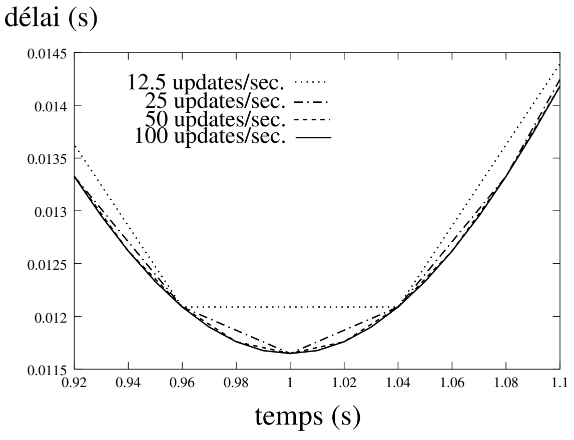


Figure 7.9 - *Influence du taux de rafraîchissement des positions sur le délai de propagation. La source se déplace à 100 km/h et passe à 4 m du récepteur. (a) Délais obtenus par méthode itérative. (b) Délais obtenus par interpolation directe.*

ne nécessitent pas de traitement particulier.

Filtre de l'environnement

Le filtre de l'environnement modélise la divergence géométrique de l'onde sonore (en «$1/r^2$» où r est la distance de la source au microphone), l'absorption atmosphérique et la visibilité le long du chemin. Le module de son spectre est donc donné par le produit de trois termes :

$$|S(f)| = \frac{V(f)\mu(f,r)}{r^2},$$

où V et μ sont respectivement les coefficients de diffraction/transmission (*cf.* paragraphe 6.2) et d'absorption atmosphérique (*cf.* Annexe A.2.2).

En pratique, on évalue ces coefficients pour les fréquences centrales de bandes d'octaves et on interpole linéairement les valeurs obtenues, pour obtenir le spectre du filtre recherché. On peut alors en déduire le filtre à phase minimale associé.

Filtre de réflexion

Le filtre de réflexion pour un chemin donné correspond à la convolution de tout les filtres de réflexion des parois rencontrées le long du chemin (*cf.* Figure 7.11).

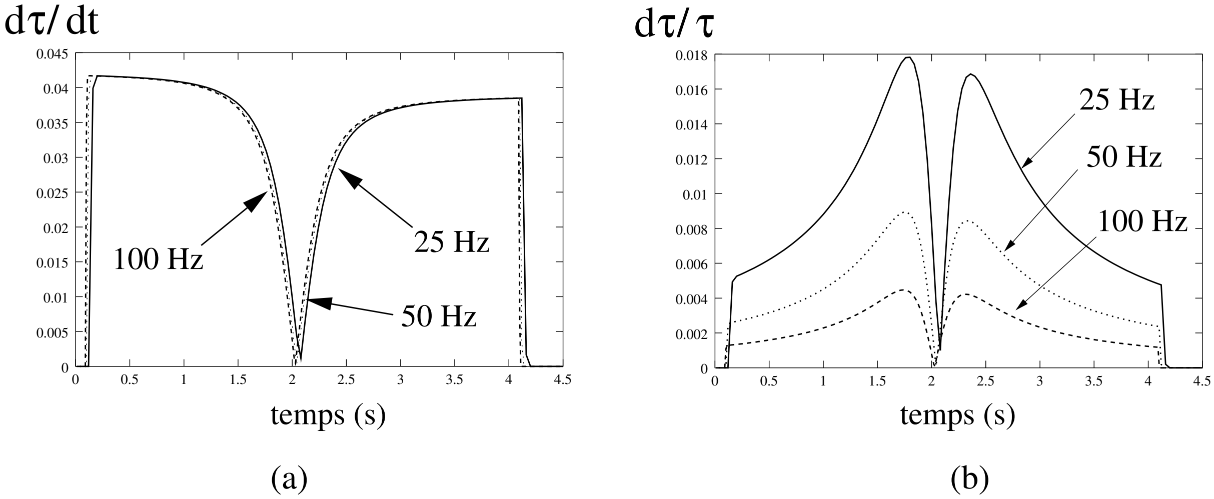


Figure 7.10 - *Critères de subdivision temporelle pour le calcul du délai de propagation. Variations comparée des critères pour différents pas de temps de simulation. (a) Le critère utilisé est $d\tau/dt$. On remarque qu'il demeure* quasi *constant lorsque l'on raffine le pas de temps de simulation. (b) Le critère utilisé est $d\tau/\tau$. Dans ce cas, le critère diminue lorsque le pas de temps est subdivisé. De plus, il augmente de manière appropriée lorsque la source arrive à proximité du microphone (et de manière symétrique lorsqu'elle commence à s'en éloigner). Dans cet exemple la source se déplace en ligne droite devant un microphone fixe. La distance de la droite au microphone est de 4 m et la vitesse de la source de 50 km/h. Dans ce cas les calculs ont été effectués avec la méthode itérative.*

```
void updateReflectionFilter(SoundPath p)
{
    if (p→father==NULL) return
    Vector exitanceDir = new Vector (p→source,p→mic)
    Vector incidenceDir = mirror(exitanceDir, p→source→reflectionPolygon→normal)
    setValue(p→reflectionFilter,incidenceDir)
    p→reflectionFilter = convolve( p→reflectionFilter, father→reflectionFilter)
}
```

Figure 7.11 - *Mise à jour du filtre de réflexion.*

7.5 Le pipeline de rendu sonore

Lors de leur mise à jour, les chemins sonores sont transmis au moteur de rendu sonore. Celui-ci va se charger des opérations de traitement du signal nécessaire au rendu d'un chemin (*cf.* Figure 7.12) :

– Création des dt_{image}/dt_{audio} nouveaux échantillons de la texture sonore approprié pour chaque chemin en fonction du délai de propagation.

– Convolution de cette texture par les filtres de source, microphone, environnement, et réflexion.

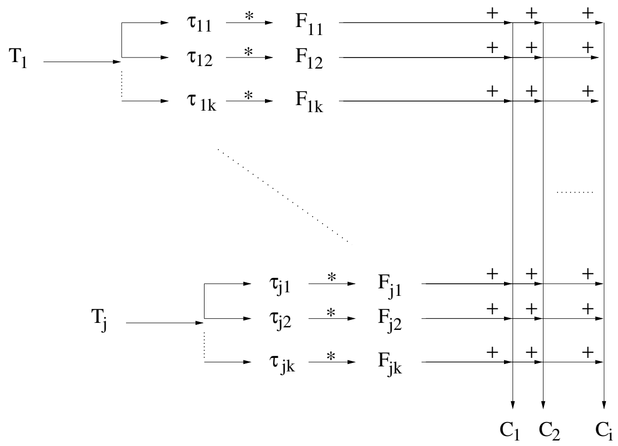


Figure 7.12 - *Pipe-line de rendu sonore. Pour chaque source S_j (y compris les sources-images) et chaque microphone M_k, on génère un bloc de dt_{image}/dt_{audio} échantillons à partir de la texture T_j en fonction du délai de propagation source-microphone τ_{jk}. Puis ce bloc est filtré par le filtre F_{jk} du chemin sonore correspondant. Les signaux résultant sont enfin mixés de manière appropriée vers les différents canaux de sortie C_i.*

– Mixage de tous les signaux ainsi crées en fonction du nombre et type de canaux disponibles en sortie.

La phase d'auralisation n'est donc pas vraiment séparée de la phase de calcul de la réponse impulsionnelle.

7.5.1 Rendu d'un chemin sonore

Comme nous l'avons déjà vu, le rendu d'un chemin sonore nécessite deux phase : le ré-échantillonnage de la texture sonore associée, puis son rendu proprement dit par convolution avec la réponse impulsionnelle du chemin. Nous détaillons à présent ces deux phases.

Ré-échantillonnage des textures sonores

Pour faire un équivalent avec la synthèse d'image, on peut voir cette opération comme une projection («mapping») de la texture dans l'espace du microphone. Comme nous l'avons vu dans la partie 1.4.2, le signal reçu par le microphone à un instant t est la texture sonore émise par la source au temps $t - \tau$, où τ est le délai de propagation :

$$m(t) = s(t - \tau(t))$$

Or, $t - \tau$ n'est pas, en général, un multiple de la fréquence d'échantillonnage de la texture sonore. Il faut donc ré-échantillonner la texture en interpolant le signal entre les deux échantillons les plus proches du temps désiré (choisir simplement l'échantillon le plus proche induirait des artefacts très perceptibles !). Une interpolation linéaire entre échantillons est suffisante [TH92, NT95]. Ce ré-échantillonnage de la texture s'accompagne néanmoins d'autres artefacts qu'il faut éliminer à l'aide de filtres appropriés. En effet, lorsque le délai de propagation diminue, on sous-échantillonne le signal et sa fréquence augmente. On peut alors dépasser la fréquence de Nyquist et il faut utiliser un filtre passe-bas pour éviter que des artefacts apparaissent. De même lorsque l'on sur-échantillonne le signal, il faut utiliser un filtre «anti-image» qui est, là encore, un filtre passe-bas. On se reportera à [Str98] pour une discussion plus approfondie sur les techniques de ré-échantillonage de signaux dans le contexte de simulation de l'effet Doppler pour la réalité virtuelle auditive. Des détails sur les algorithmes de ré-échantillonage peuvent être également trouvés dans [Dig79].

Rendu de la texture

Une fois la texture projetée dans l'espace du microphone, elle doit être filtrée par les quatre filtres définissant le chemin sonore correspondant à la source, au microphone, à l'environnement et aux réflexions. Ces quatre filtres ainsi que la texture sont donc convolués ce qui donne le signal sonore rendu pour le pas de temps courant. On peut voir cette opération comme analogue au calcul de l'éclairage local en synthèse d'image.

7.5.2 Implémentation parallèle du moteur de rendu

Une implémentation parallèle utilisant un processus par chemin sonore parait très naturelle à mettre en œuvre. La Figure 7.13 décrit le traitement effectué par chaque processus.

```
void convolutionThread(SoundPath p, double currentTime)
{
    swap(p→newFilter,p→oldFilter)
    swap(p→newSamples,p→oldSamples)
    updatePathImpulseResponse(p→newFilter)
    long length = dt_frame/dt_audio
    for i= 0 to len
    {
        double t = currentTime - i*p→newDelay - (len-i)*p→oldDelay
        p→newSamples[i] = interpolate(p→source→sourceSignal,t)
    }
    convolve(p→newSamples,p→newFilter,buffer1)
    convolve(p→oldSamples,p→newFilter,buffer2)
    crossfade(buffer1,buffer2,outputBuffer)

}


forall p in renderEngine.pathList
    launchThread(p,simulationTime)
```

Figure 7.13 - *Pseudo-code pour le rendu des chemins sonores. Le moteur de rendu est organisé de manière parallèle, le rendu de chaque chemin étant effectué par un processus indépendant.*

Nous avons implémenté ce moteur de rendu sur une architecture parallèle bi-processeurs (station SGI Onyx2 Infinite Reality R10000 180 MHz $\times 2$). Il permet d'obtenir le rendu stéréophonique temps réel d'un chemin à 44,1 kHz et avec un traitement des occultations est effectué pour 3 bandes d'octaves. Des exemples plus complexes comportant une cinquantaine de chemins et un traitement des occlusions sur 10 bandes d'octaves demandent un temps de traitement d'une quarantaine de seconde pour une seconde de son simulé. On est alors loin de l'interactivité mais ce temps reste raisonnable dans le cadre de sonorisation d'animations de synthèse. Il faut être conscient que, particulièrement dans le cadre de simulations utilisant des modèles physique, la majeure partie du temps va être consacrée au calcul des trajectoires des objets à l'aide des équations de la dynamique. Pour un surcoût de calcul moindre, on peut intégrer directement dans la simulation la synthèse d'une bande-son synchronisée avec l'image.

7.5.3 Prise de son virtuelle et restitution

Jusquá présent, nous n'avons pas évoqué le problème de la restitution du signal sonore généré. C'est pourtant elle qui va, dès le départ conditionner différents paramètres du système comme le nombre de microphones utilisés, leur type (deux «oreilles» pour de la binauralisation, une «paire stéréophonique » pour de la stéréophonie, etc.) et le nombre de canaux de sortie (un pour de la monophonie, deux pour de la stéréophonie, quatre pour de la quadriphonie, etc.).

Notre système permet de travailler dans une approche de preneur de son en plaçant différents microphones de types différents à différents endroits de l'environnement. Le son généré sera ensuite mixé par le système en fonction d'une «*matrice de mixage*» indiquant quelle proportion du signal arrivant à un microphone doit être dirigée vers chaque canal[5]. La matrice est donc de taille $M \times C$ où M est le nombre de microphones et C le nombre de canaux de sortie. On peut donc ajouter comme paramètres d'un chemin la liste des canaux (et des gains associés) auxquels sa contribution doit être additionnée. Ce *matriçage* des signaux peut également faire appel à des opérations plus complexes si on veut fournir en sortie un signal encodé suivant un protocole particulier du type «*Dolby Surround*» par exemple [Dre88] (*cf.* Annexe C.2).

Enfin, si l'on utilise des sources sonores d'«ambiance» (*cf.* paragraphe 7.1.1), leurs contributions sont ajoutées directement au niveau de ce mixage final, sans passer par le moteur de rendu et la simulation. On pourra ainsi spécifier de la même manière une matrice de mixage pour ce type particulier de sources.

7.5.4 Notes sur le traitement du signal

Nous avons implémenté nos fonctions de transfert sous forme de filtres RIF (128 coefficients à 32 kHz par exemple), permettant d'obtenir un traitement rapide du signal.

Tous les filtres de base sont stockés sous forme de leur transformée de Fourier (fonction de transfert) ce qui évite l'étape d'une transformée de Fourier lors de la convolution. Pour les filtres donnés par leur spectre en bandes d'octaves, ils sont convertis immédiatement en bandes fines au début du processus (on les choisira alors classiquement à phase minimale). Dans la plupart des cas, on dispose uniquement de données en bande d'octave. Pendant la mise à jour des chemins sonores, on ne calcule alors que les spectres des différents filtres en bandes d'octave ce qui permet d'accélérer le processus (le produit terme à terme de deux spectres est le spectre de la convolution). Au moment du rendu du chemin, on fabrique alors un filtre choisi à phase minimale, dont le spectre en bandes fines est interpolé à partir des valeurs en bandes d'octave. Les convolutions sont effectuées à l'aide d'une méthode de type «*overlap-save*» [PTVF92], nécessitant de doubler la taille N des filtres et du signal à filtrer (on conserve les N échantillons correspondant à la texture au pas de temps précédent, auxquels on ajoute les N nouveaux échantillons). De plus, afin d'éviter les artefacts («clics») lors de la mise à jour des filtres, on calcule en fait une paire de signaux pour chaque pas de temps : l'un en utilisant le nouveau filtre et l'autre en utilisant le filtre au pas de temps précédant ; on réalise alors un fondu-enchaîné entre les deux signaux

5. on notera que ce mixage pourrait également être effectué *a posteriori*.

(*cf.* Figure 7.14). Cette méthode, pourtant simple, permet d'éliminer efficacement les artefacts (aux pas de temps que nous utilisons du moins).

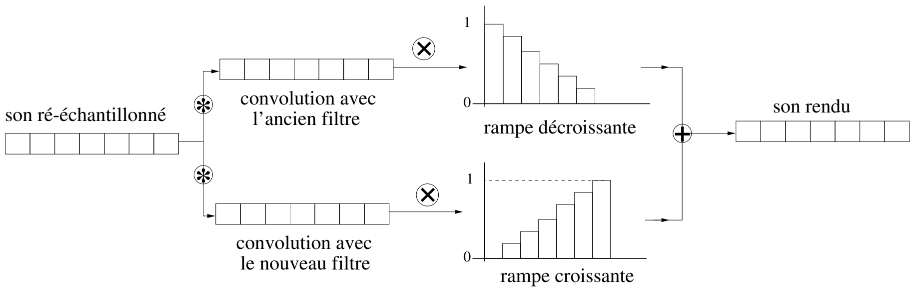


Figure 7.14 - *Fondu enchaîné lors de la convolution pour éliminer les artefacts du à la modification du filtre*

7.6 Applications

Le système de rendu sonore que nous venons de décrire à été implémenté dans la plateforme «*Fabule*» d'animation et de simulation de l'équipe *i*MAGIS [Gas94]. Nous l'avons utilisé principalement pour la sonorisation d'animations. Il permet d'obtenir très simplement une bande-son parfaitement synchronisé avec l'image. Nous l'avons testé dans différentes conditions de restitution : quadriphonie (quatre microphones restitués sur quatre enceintes dans une approche similaire dans le principe à [Cho71]), stéréo (paire ORTF), stéréo binaurale (deux microphones munis de fonctions de transfert mesurées sur une tête artificielle *KEMAR* au *M.I.T. Medialab* [GM94, Gar95]).

Il va être également utilisé au sein d'un simulateur de systèmes de téléconférence dans le cadre d'un contrat avec le Centre National d'Études des Télécommunications baptisé *Télémédia*. Il permettra d'évaluer de manière semi-quantitative un dispositif de prise de son/restitution pour un «puit» permettant un communication informelle entre deux groupes d'utilisateurs distants.

7.7 Bilan

Dans ce chapitre, nous avons présenté un système de rendu sonore permettant de générer des bandes son synchronisées dans un système d'animation de synthèse interactif.

Ce modèle utilise un concept de chemins sonores des sources vers les microphones virtuels. Ces chemins sont étendus aux chemins correspondant aux réflexions spéculaires des ondes sonores dans un formalisme de sources-images. On ne va pas construire la réponse impulsionnelle globale de l'environnement mais superposer les signaux obtenus pour les différents chemins. Ces signaux résultent de la convolution de la texture sonore d'origine et de la réponse impulsionnelle élémentaire de chaque chemin. Cette réponse est modélisée comme le produit de trois filtres correspondants aux propriétés de la source, du microphone et de l'environnement. Ce dernier permet de prendre en compte les effets d'atténuation du son par les obstacles en utilisant la méthode présentée au Chapitre 6. Nous avons également montré que cette méthode peut être étendue sans charge supplémentaire aux chemins réfléchis. Enfin, l'effet Doppler peut également être traité.

Cette approche permet d'obtenir des résultats très réalistes pour des temps de calculs rapides. Dans le cas d'environnements simples, le processus fonctionne en temps-réel. Le système est néanmoins très limité par les opérations de convolution qui demeurent coûteuses et sont effectués ici de manière logicielle.

Il est clair que l'utilisation de cartes de traitement du signal câblé permettrait d'augmenter sensiblement les capacités.

Le problème principal de l'approche provient sans nul doute du formalisme de source-images, peu adapté à simuler de grands ordres de réflexion. Nous n'avons pas proposé ici d'approche originale permettant de traiter efficacement la réverbération tardive. Dans le cadre d'applications de type production audio, une technique de réverbération artificielle pourrait parfaitement convenir. Néanmoins, comme nous l'avons déjà évoqué ces approches sont fondées sur des hypothèses fortes concernant l'environnement (lieux clos, champ diffus). De plus, on aimerait également pouvoir prendre en compte des phénomènes de réflexion plus complexes, particulièrement les cas de surfaces lambertiennes.

❒ ❒ ❒

Dans les chapitres précédents, nous nous sommes intéressés au problème du traitement efficace de phénomènes de propagation complexes, en particulier la diffraction. Nous y avons apporté une solution basée sur l'utilisation d'un modèle géométrique et du rendu graphique câblé de cartes spécialisées aujourd'hui disponibles sur la plupart des machines. Nous avons également présenté une application directe de notre solution sous la forme d'un système de rendu sonore pour des applications graphiques 3D interactives, permettant une intégration cohérente du son et de l'image. Toutefois, ce système est basé sur un modèle de sources-images et n'apporte pas une solution suffisamment satisfaisante pour traiter de grands ordres de réflexion. Elle ne permet pas non plus de traiter les réflexions diffuses dues à des surfaces rugueuses, ou utilisées fréquemment pour compenser une simplification géométrique. Dans la dernière partie de ce document, nous allons nous intéresser à la simulation efficace d'échanges diffus et spéculaires, y compris pour de grands ordres de réflexion. Nous présentons une nouvelle approche adaptative basée sur une technique d'échanges radiatifs hiérarchiques dont la résolution peut évoluer au cours du temps.

Quatrième partie

Échanges énergétiques temporels hiérarchiques : application aux ondes sonores

Chapitre 8

Radiosité hiérarchique et extensions

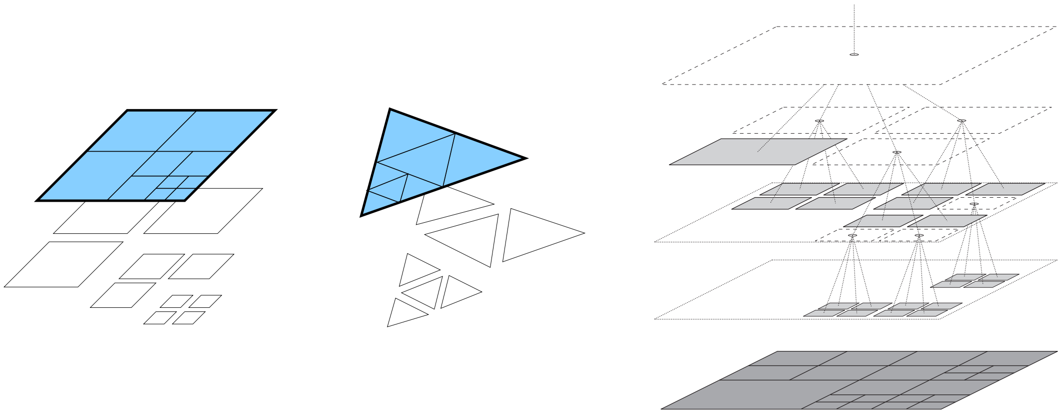
Nous revenons ici sur les techniques de simulation d'échanges énergétiques que nous avions déjà évoquées au paragraphe 3.4. Depuis le début des années 80 [GTGB84], la simulation de l'éclairage global par échanges radiatifs entre surfaces fait l'objet d'intenses recherches dans la communauté graphique. Ces techniques, dites aussi «de radiosité» (car on y calcule la puissance lumineuse par unité de surface), offrent une alternative intéressante aux coûteuses simulation d'éclairage global par lancer de rayons stochastique, en particulier dans le cas où les surfaces sont des réflecteurs diffus. Nous présentons ici un rapide tour d'horizon de quelques développements de cette méthode dans le cadre de la synthèse d'image et en particulier une de ses plus importantes extensions, la *hiérarchisation* des calculs, qui permet de traiter de manière efficace des scènes complexes. Nous présenterons également des extensions permettant de traiter des surfaces non lambertiennes, aux propriétés de réflectance complexes. Nous en discuterons l'applicabilité aux simulations de phénomènes acoustiques.

8.1 Représentations hiérarchiques et structures de données

Pour répondre à la complexité des techniques de radiosité nécessitant de calculer et stocker un nombre de facteurs de forme croissant avec le carré du nombre d'éléments dans la scène, se sont dévelop-pées des approches hiérarchiques permettant de réduire le nombre d'interactions entre objets. Ces approches s'appuient sur le maillage adaptatif et hiérarchique des surfaces de la scène en fonction d'un critère d'erreur. Une interaction lumineuse est, dans ce cas, simulée entre deux éléments à un niveau plus ou moins fin de la hiérarchie. Ainsi les zones où plus de précision est nécessaire sont maillées de manière plus dense. A l'inverse, on effectue moins de calculs dans les zones où ce n'est pas nécessaire. La précision accordée au calcul d'une interaction énergétique dépend de son importance dans la solution globale.

8.1.1 Hiérarchies de surfaces

Ces approches utilisent une représentation hiérarchique des surfaces de la scène permettant de les subdiviser de manière adaptative en facettes plus petites. Une surface est représentée par une arborescence dont la racine est la surface elle même et chaque nœud est une sous-partie de la surface de son père (les facettes à un niveau forment un pavage de leur face père). La structure la plus répandue est une structure dans laquelle quatre fils sont générés à chaque subdivision (*cf.* Figure 8.1) .



(a) (b) (c)

Figure 8.1 - *Subdivision hiérarchique de surfaces. Dans le cas de quadrilatères (a) et de triangles (b), on peut subdiviser le polygone en utilisant les milieux des arêtes. On peut ré-appliquer ce processus récursivement. (c) La structure de donnée correspondante est donc un arbre de facteur de branchement égal à 4 («quadtree»)*[1].

8.1.2 Représentation hiérarchique de la fonction de radiosité

Pour gérer les échanges énergétiques à différents niveaux de résolution, il ne suffit pas de pouvoir subdiviser les éléments de surface de manière adaptative, il faut également être capable de représenter la fonction de radiosité d'une manière cohérente à ces niveaux de raffinement. L'approche classiquement choisie est une représentation de la fonction sous forme de *pyramide* [SP94] , dans laquelle la fonction à un niveau de résolution donné est une *moyenne* des valeurs au niveau de résolution supérieur. Dans le cas de la fonction de radiosité, qui est on le rappelle une quantité définie *par unité de surface*, cette moyenne est donc effectuée au prorata des aires des surfaces au niveau de subdivision supérieur. La même structure arborescente est utilisée pour représenter de manière hiérarchique les éléments de surface et leur valeur de radiosité associée. Comme dans le cas des méthodes de radiosité classique, on choisira généralement de représenter la radiosité comme un fonction constante sur chaque élément (*cf.* paragraphe 3.4). Cette représentation correspond à une décomposition de la radiosité en ondelettes de Haar. On peut utiliser des bases d'ondelettes d'ordre supérieur [GSCH93].

8.2 Échanges hiérarchiques. Raffinement

Les interactions énergétiques entre facettes sont représentées par des «liens» , le principe étant que deux éléments vont être «liés» à un niveau de hiérarchie donné si l'on considère que l'échange énergétique est «suffisamment bien» représenté à ce niveau. Le principal problème est bien entendu, comme nous le verrons, de quantifier ce «suffisamment bien». En pratique, un lien est une structure de données comprenant une facette de départ, une facette d'arrivée et une estimation du facteur de forme associé au transfert[2]. À chaque facette, on associe donc en plus de sa radiosité, une liste de liens vers d'autres facettes de la scène.

1. figure tirée de l'ouvrage «*Radiosity and Global Illumination*» par F.X. Sillion et C. Puech [SP94].
2. on rappelle que le facteur de forme est la proportion d'énergie quittant la facette source qui arrive sur la facette destination (*cf.* paragraphe 3.4.

L'algorithme fonctionne alors de la manière suivante :

– Dans une *phase de précalcul*, on évalue la visibilité entre les surfaces d'entrée (le plus haut niveau de la hiérarchie) On établit un lien (on calcule donc une estimation du facteur de forme) pour chaque couple d'objets qui ne sont pas mutuellement invisibles.

– Ensuite à chaque étape du calcul on va :

   1. propager l'énergie le long de tous les liens.
   2. examiner tous les liens et raffiner ceux qui ne satisfont pas un certain critère de précision

La résolution est donc, comme dans le cas de la radiosité classique un processus itératif.

Complexité de l'algorithme

L'un des aspects essentiels de l'algorithme de radiosité hiérarchique est sa complexité. L'étape de prétraitement pour l'établissement des liens a une complexité en $O(m^2 \log m)$, voire $O(m^3)$ en temps où m est le nombre de surfaces initiales. Chaque étape de raffinement est liée au nombre total d'éléments N de la hiérarchie et peut être effectuée en temps en $O(N)$. Chaque étape de raffinement peut nécessiter un calcul de visibilité en $O(\log m)$, voire $O(m)$. La complexité totale de l'algorithme est donc en $O(m^3 + Nm)$ alors que celle d'un algorithme classique est en $O(N^2 m)$.

8.2.1 Échanges hiérarchiques

A chaque étape du calcul, l'énergie est échangée le long des liens au niveau de précision où ils ont été établis. Il est clair que pour connaître l'énergie reçue par un élément, il faut prendre en compte l'énergie reçue par ses «pères» et ses «fils». Après avoir collecté l'énergie sur chaque élément, il faut mettre à jour de manière cohérente l'information hiérarchique en vue de l'itération suivante. Cela nécessite une opération de traversée de la hiérarchie de chaque surface ou «*push-pull*» (*cf.* Figure 8.2)

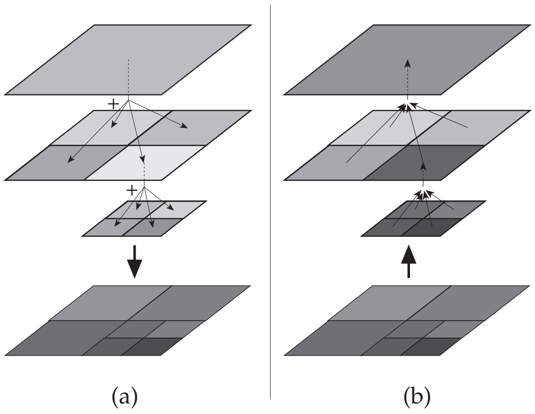


Figure 8.2 - *Combinaison des échanges énergétiques à différents niveaux de la hiérarchie ou* «push-pull»[3]. *Ici, les carreaux les plus foncés indiquent la valeur la plus élevée de la fonction de radiosité. (a) Les valeurs de radiosité sont propagées jusqu'aux feuilles et ajoutées à chaque niveau de la hiérarchie (*«push»*). (b) Les valeurs de radiosité sont ensuite remontée jusqu'à la racine de la hiérarchie tout en étant moyennées à chaque niveau (*«pull»*).*

3. figure tirée de l'ouvrage «*Radiosity and Global Illumination*» par F.X. Sillion et C. Puech [SP94].

8.2.2 Raffinement des échanges

Dans un second temps, on raffine les échanges dont on considère qu'ils n'ont pas été suffisamment bien représentés. On examine chaque lien et, si il ne satisfait pas le critère imposé, on le subdivise en plusieurs nouveaux liens. Ce raffinement s'accompagne de la subdivision de l'un ou l'autre des éléments source ou récepteur reliés par le lien. Les nouveaux liens sont créés entre les fils de l'élément subdivisé et l'élément non-subdivisé.

En général, c'est le récepteur qui est raffiné puisque la radiosité reçue va varier de manière non négligeable. La source doit être subdivisée, elle, dans le cas où elle est non-uniforme ou partiellement visible. Ce choix est également du ressort de la fonction d'évaluation de l'erreur commise sur le transfert. Les nouveaux liens ainsi créés sont alors examinés à leur tour récursivement jusqu'à ce que la précision requise soit atteinte ou que les surfaces aient atteint une taille limite fixée par l'utilisateur.

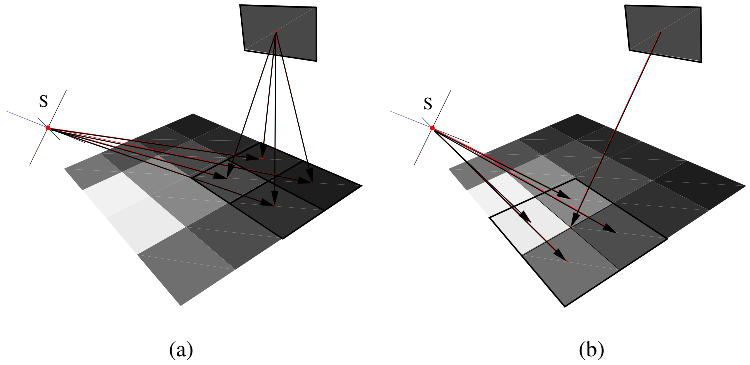


Figure 8.3 - *Établissement des échanges radiatifs à différents niveaux de précision. (a) (b)*

Oracles

Toute l'efficacité d'un algorithme de radiosité hiérarchique réside dans le critère ou *oracle* qui détermine à quel niveau un échange doit être établi. C'est également, bien entendu, l'un des points les plus délicats de l'approche. Plusieurs types d'oracles ont été développés, basés sur la valeur du facteur de forme, la quantité d'énergie transmise (produit du facteur de forme et de la de la radiosité, par exemple) [SP94] ou plus récemment sur une estimation de bornes d'erreurs sur le facteur de forme [LSG94].

8.3 Extension aux surfaces non diffuses

La technique de radiosité comme celles de radiosité hiérarchique sont basées sur l'hypothèse que les surfaces des objets sont lambertiennes. Pourtant, en optique comme en acoustique, ce modèle est loin de reproduire la réalité des matériaux qui nous entourent. Dans le cadre de la synthèse d'image, les techniques de radiosité ont été étendues pour pouvoir prendre en compte des phénomènes de réflexion plus complexes.

8.3.1 Approches à plusieurs passes

Ces approches sont basées sur une classification des chemins optiques en fonction du type de réflexion (purement diffus, purement spéculaire, hybride). Elles divisent le problème en deux étapes : une étape *indépendante du point de vue* et une étape *dépendante du point de vue* [SP94]. Dans la première étape on ne représente sur les surfaces que la composante diffuse de l'éclairage global. Une approche de type radiosité classique peut alors être utilisée pour simuler les échanges diffus. Des extensions utilisant des facteurs de forme étendus ont également été proposées. Ces facteurs de forme permettent de prendre en compte un nombre quelconque de réflexions spéculaires entre deux réflexions diffuses [SP94]. Dans la deuxième passe, un lancer de rayon simplifié depuis le point de vue est utilisé pour collecter l'éclairage diffus sur les surface visibles. Il permet également de traiter les chemins de réflexion complètement spéculaires. On notera toutefois qu'aucun rayon n'est alors nécessaire pour le calcul des ombres déjà traitées par la première passe.

Ces méthodes hybrides sont proches de celles que l'on rencontre en acoustique et que nous avons évoquées au paragraphe 3.3. Néanmoins, dans le cadre de l'acoustique, la solution indépendante est rarement stockée sur les faces mais «reprojetée» au point d'écoute au fur et à mesure du calcul. Ceci permet de réduire de manière significative l'espace mémoire nécessaire au détriment de la possibilité de conserver une information réellement indépendante du point de vue.

8.3.2 Radiosité directionnelle

Un problème des approches évoquées précédemment est qu'elles ne permettent de traiter simplement que les réflectances combinant un terme purement spéculaire et un terme purement lambertien. Or, comme nous l'avons vu au paragraphe 1.5.1, ces deux modèles ne permettent qu'une modélisation limitée des phénomènes de réflexion complexes que l'on rencontre dans la réalité.

Les approches de radiosité directionnelle sont basées sur une discrétisation des surfaces et de l'espace des directions en éléments, dans un formalisme d'éléments finis. Elle permettent de prendre en compte des distributions de radiance quelconques. A la différence des approches précédentes, elles sont complètement indépendantes du point de vue puisqu'une information de radiance dépendant de la direction est stockée sur chaque élément de surface.

On peut discrétiser l'espace des directions en utilisant une «grille» sur les faces d'un cube («global cube») centré sur chaque élément de surface et orienté de manière fixe (*cf.* Figure 8.4) [IC86]. Chaque «cellule» du cube porte l'information de visibilité et de facteur de forme directionnel qui est calculée une seule fois au début du processus. La résolution est ensuite similaire à celle d'un algorithme de radiosité standard, le système d'équations à résoudre étant donné par :

$$L_{kl} = L_e(x_k, \Theta_l) + \sum_j \rho_{bd}(x_k, \Theta_l, -\Theta_j) L_{v(j)j} \int_{\Theta \in -\Omega_j} \cos\theta d\omega,$$

où L_{kl} est la radiance correspondant à l'élément spatial et directionnel (k,l), émise au point x_k dans la direction Θ_l. $L_{v(j)j}$ représente la radiance reçue depuis la direction Θ_j en provenance de l'élément de surface visible $v(j)$ depuis cette direction (il sera généralement supposé unique) et ρ_{bd} est la fonction de distribution de la réflectance bidirectionnelle associée à la surface considérée.

Malgré le nombre relativement limité d'interactions pertinentes, ces simulations restent très coûteuses en temps de calcul et en espace mémoire (les temps de simulation annoncés, en 1986, se situent entre 25 et 192 heures pour des scènes de complexité très modérée).

Des approches par décomposition de la radiance directionnelle en ondelettes ont également été utilisées [CSSD96].

Pour le rendu final de toutes ces méthodes, les surfaces visibles depuis le point de vue sont identifiées, puis pour chaque pixel de l'image, on retrouve les éléments les plus proches du point correspondant sur la surface visible. Les valeurs de radiance dans la direction correspondante sont alors extraites et interpolées pour donner la radiance au point cherché. Cela peut être effectué en utilisant un lancer de rayon mais limité, ici, aux seuls rayons *directs*.

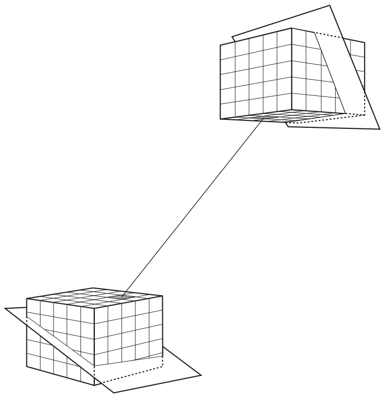


Figure 8.4 - *Radiosité directionnelle avec l'algorithme du «global cube»*[4]*. Des cellules à la surface de cubes centrés sur les éléments de surface sont utilisées pour stocker les informations directionnelles.*

Une autre approche, particulière, est basée sur la notion de «transport à trois points» [AH93]. Elle exprime comme variables du problème les interactions entre deux points avec une réflexion intermédiaire sur un troisième. Un algorithme hiérarchique a été développé sur cette base mais il reste extrêmement coûteux (sa complexité est $O(n+k^3)$ où n est le nombre de feuilles dans la hiérarchie et k est le nombre de surfaces de départ).

8.3.3 Radiosité avec des distributions directionnelles

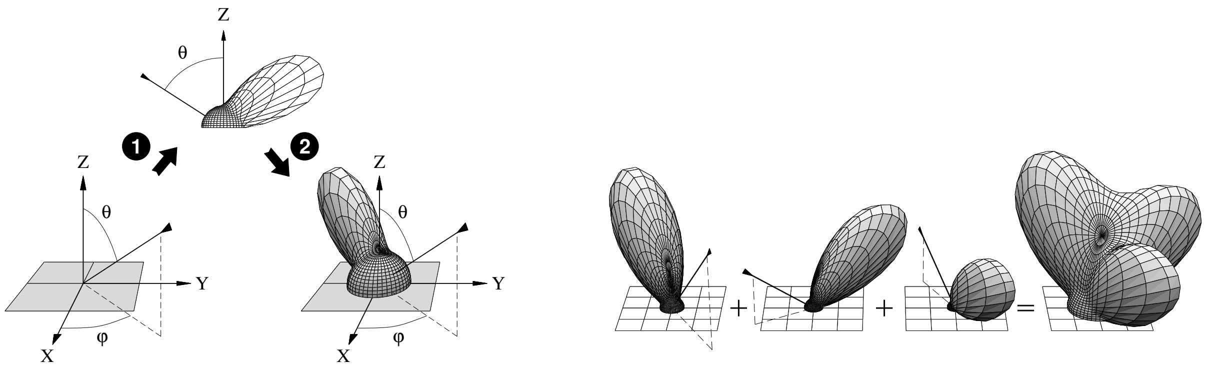
L'utilisation d'une information stockée sur les faces de cubes orientés de manière identique, comme dans les techniques de «global cube», produit sur les images des artefacts visuels. Il serait préférable de pouvoir représenter l'information de radiance directionnelle de manière continue. Les approches de radiosité avec distributions directionnelles proposent de simuler des échanges diffus directifs en utilisant directement une représentation fonctionnelle de la distribution de radiance directionnelle. Tous les calculs sont effectués directement sur ces distributions, comme ils sont effectués sur la valeur scalaire de radiosité (*cf.* Figure 8.5). On préférera une représentation compacte des fonctions de radiance, comme celle que l'on peut obtenir par décomposition en *harmoniques sphériques* [SP94]. En outre, les opérations de rotation et sommation des distributions peuvent alors être effectuées de manière plus efficace directement sur les coefficients de la décomposition. Toutefois, les calculs impliqués demeurent encore très importants.

Les effets des réflexions spéculaires idéales peuvent être ajoutés dans une seconde phase, d'une manière similaire à celle des approches à passes multiples [SAWG91].

8.4 Bilan et applicabilité à la simulation acoustique

Il est clair que les approches d'échanges énergétiques hiérarchiques apparaissent très attractives pour traiter les problèmes de simulation acoustique. Comme c'est le cas dans le cadre des simulations visuelles, elles peuvent permettre une amélioration notable de la complexité et de la qualité des calculs (en particulier, on pourra imaginer faire évoluer la précision des échanges au cours du temps). Néanmoins, la définition de nouveaux oracles de raffinement liés à l'aspect fréquentiel du phénomène sonore sera sans doute nécessaire.

4. figure tirée de l'ouvrage «*Radiosity and Global Illumination*» par F.X. Sillion et C. Puech [SP94].



(a) (b)

Figure 8.5 - *Radiosité avec distributions de radiance directionnelles*[5]. *(a) La fonction de distribution de la réflectance bidirectionnelle d'un matériau permet d'obtenir directement la distribution de la radiance émise pour une direction d'incidence* (θ, φ) *donnée (1). Celle-ci est ensuite replacée dans le repère de l'élément de surface considéré (2). (b) Ces distributions de radiance directionnelles sont sommées pour les différentes directions d'incidence de l'énergie sur l'élément de surface considéré avant d'être redistribuées.*

Par contre, il parait moins évident d'étendre les méthodes de radiosité directionnelles déjà très coûteuses en temps et en mémoire à une dimension supplémentaire ; le temps. Pour cela, les approches hybrides paraissent plus appropriées. Dans le cadre de simulations pouvant prétendre être interactives, il semble également peu réalisable d'introduire des modèles de réflexion complexes. En outre, aucun modèle n'existe actuellement (à notre connaissance du moins) pour simuler de tels phénomènes dans le cadre des ondes sonores et la mesure des distribution directionnelles demeure encore extrêmement complexe.

5. figure tirée de l'ouvrage «*Radiosity and Global Illumination*» par F.X. Sillion et C. Puech [SP94].

Chapitre 9

Radiosité hiérarchique temporelle

Ce chapitre présente la dernière contribution de ce travail : une extension des techniques de radiosité hiérarchique à la simulation d'échanges énergétiques temporels, en particulier dans le cadre de simulations acoustiques. La méthode que nous présentons permet de traiter des échanges globaux dans le cadre de réflexions diffuses. Pour réduire la complexité du processus, nous décrivons également une technique originale de «regroupement temporel» de l'énergie ; fidèle à la théorie ondulatoire, elle permet de prendre en compte des interférences. Puis nous proposerons une extension aux réflexions spéculaires à l'aide d'un modèle unifié de sources-images. Nous examinerons pour conclure les possibilités d'intégration d'une telle approche dans un environnement de simulation dynamique.

9.1 Caractérisation des échanges énergétiques temporels diffus

Nous commençons par reprendre, d'une manière un peu plus détaillée l'expression des échanges temporels entre surfaces lambertiennes, que nous avions présentés au paragraphe 3.4. Nous nous plaçons dans le cadre de la propagation d'ondes sonores, et considérons l'intensité acoustique, équivalent de la radiosité en synthèse d'image. Toutefois, ce n'est pas un point limitatif de l'approche.

On rappelle que l'équation décrivant les échanges d'*intensité acoustique* au cours du temps peut s'exprimer comme (*cf.* figure 9.1) :

$$I^\lambda(x,t) = Ie^\lambda(x,t) + \gamma^\lambda(x) \int_{y\in S} \kappa^\lambda(x,y) I^\lambda\left(y, t-\frac{r}{c}\right) dy, \tag{9.1}$$

avec $\kappa^\lambda(x,y) = \frac{\cos\theta\cos\theta'}{\pi r^2} V(x,y)\mu^\lambda(r)$, où $V(x,y)$ est un terme de visibilité binaire et μ^λ est l'atténuation atmosphérique de l'onde sonore.

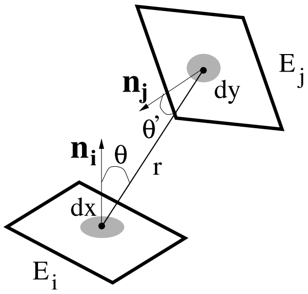


Figure 9.1 - *Notations pour les échanges radiatifs entre surfaces et le facteur de forme.*

9.1.1 Formulation du problème et hypothèses

On peut ré-écrire l'équation 9.1 en faisant intervenir plus clairement le délai de propagation de l'onde sonore entre deux points :

$$I^\lambda(x,t) = Ie^\lambda(x,t) + \gamma^\lambda(x) \int_{\tau=0}^{+\infty} \int_{y \in S} \kappa^\lambda(x,y)\delta(r - c\tau) I^\lambda(y, t - \tau)\, dy d\tau, \tag{9.2}$$

où $\delta(r - c\tau)$ est la distribution de Dirac non nulle lorsque $r = c\tau$. On remarque que cela consiste à exprimer l'intégrale sur des courbes «iso-délai de propagation».

Supposons à présent que l'on subdivise les surfaces de la scène en N éléments polygonaux E_i et que l'intensité est *constante en tout point* de ces éléments. Les éléments étant finis les délais de propagation point à point sont tous compris entre un délai minimum $\tau_{\min}$ et un délai maximum $\tau_{\max}$. L'équation 9.2 devient :

$$I_i^\lambda(x,t) = Ie^\lambda(x,t) + \gamma_i^\lambda \sum_{j=1}^{N} \int_{\tau_{\min}}^{\tau_{\max}} I_j^\lambda(t - \tau) \int_{y \in E_j} \kappa^\lambda(x,y)\delta(r - c\tau)\, dy d\tau,$$

où $[\tau_{\min}, \tau_{\max}]$ est la *durée* de l'échange entre les éléments.

On remarque un paradoxe apparent puisque l'intensité supposée constante sur les surfaces est susceptible de varier au point x. En fait, on suppose l'intensité constante sur un élément uniquement pour effectuer le calcul de propagation de l'énergie sur les autres surfaces [SP94]. Cette valeur est calculée comme une moyenne sur la surface des intensités aux différents points (il en va de même pour l'intensité «propre» Ie) :

$$I_i^\lambda(t) = \frac{1}{A_i} \int_{x \in E_i} I^\lambda(x,t)\, dx.$$

On a donc au final :

$$I_i^\lambda(t) = Ie_i^\lambda(t) + \gamma_i^\lambda \sum_{j=1}^{N} \int_{\tau_{\min}}^{\tau_{\max}} I_j^\lambda(t - \tau) \underbrace{\frac{1}{A_i} \iint_{(x,y) \in (E_i, E_j)} \kappa^\lambda(x,y)\delta(r - c\tau)\, dx dy}_{F_{ij}^\lambda(\tau)} d\tau. \tag{9.3}$$

Dans ce cas l'intensité qui atteint une surface peut s'exprimer sur la durée de l'échange $[\tau_{\min}, \tau_{\max}]$ comme une convolution entre un «*facteur de forme temporel*» $F_{ij}^\lambda(\tau)$ et l'intensité émise par la surface source au cours du temps $I_j^\lambda(t - \tau)$. Les Figures 9.2 et 9.3 montrent des exemples de ce facteur de forme temporel échantillonné dans des cas simples. Pour cela, nous avons approché l'intégrale de l'équation 9.3 en échantillonnant les surfaces puis en intégrant les termes par pas de temps successifs.

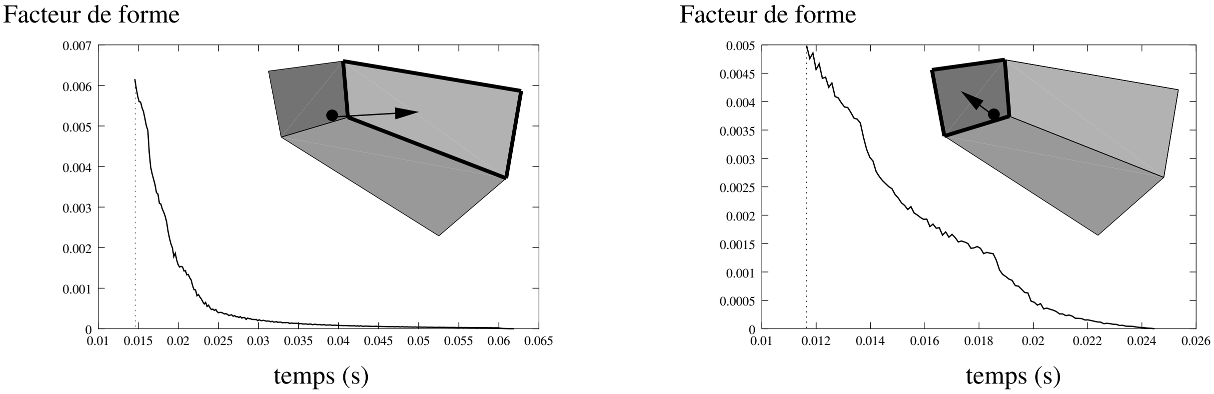
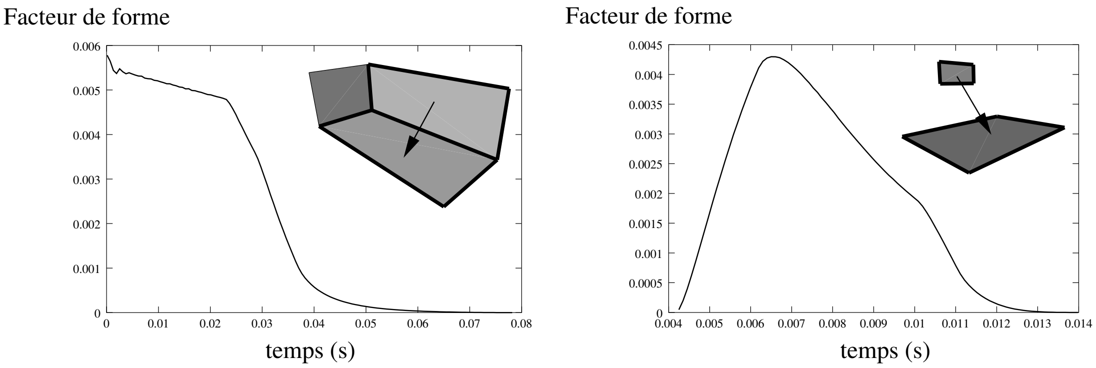


Figure 9.2 - *Exemples de «facteurs de forme temporels» pour une source ponctuelle et différentes configurations de récepteurs.*



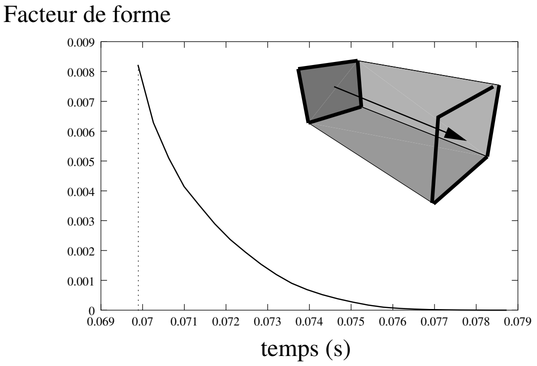


Figure 9.3 - *Exemples de «facteurs de forme temporels» pour une source surfacique et différentes configurations de récepteurs.*

Nous introduisons une simplification supplémentaire ; nous supposons également que l'intensité sur les éléments ne dépend pas du temps au cours de la durée de l'échange. Bien entendu l'intensité varie globalement au cours du temps. On la considère simplement constante sur de petits pas de temps qui correspondent à la durée des échanges. On obtient alors :

$$I_i^\lambda(t) = Ie_i^\lambda(t) + \gamma_i^\lambda \sum_{j=1}^{N} I_j^\lambda(t - T_{ij}) \underbrace{\frac{1}{A_i} \int_{\tau_{\min}}^{\tau_{\max}} \iint_{\{(x,y)\in(E_i,E_j)/r=c\tau\}} \kappa^\lambda(x,y)\,dx\,dy\,d\tau}_{F_{ij}^\lambda}, \tag{9.4}$$

où $T_{ij} = \tau_{\min}$ est le temps de propagation minimum entre les deux éléments en considérant les couples de points *mutuellement visibles*. F_{ij}^{λ} est dans ce cas le facteur de forme «classique», tel que nous l'avons introduit au paragraphe 3.4.

On obtient l'équation de radiosité temporelle :

$$I_i^{\lambda}[t] = Ie_i^{\lambda}[t] + \gamma_i^{\lambda} \sum_{j=1}^{N} F_{ij}^{\lambda} I_j^{\lambda}[t - T_{ij}] \tag{9.5}$$

Les échanges correspondent à différents instants discrets induits par les T_{ij}.

Nous allons voir à présent comment il est possible de représenter l'intensité acoustique et d'effectuer les échanges de manière hiérarchique au cours du temps.

9.2 Radiosité hiérarchique temporelle en régime impulsionnel

Dans ce paragraphe, nous présentons notre approche hiérarchique des échanges radiatifs temporels. En particulier, nous verrons comment il est possible de représenter la radiosité sur une surface au cours du temps de manière hiérarchique. Nous étendrons également les notions de facteur de forme et de lien «classiques» afin de prendre en compte l'aspect temporel et fréquentiel de la simulation.

9.2.1 Discrétisation du temps et de l'espace

Comme dans les techniques de radiosité hiérarchique que nous avons décrites précédemment, nous allons utiliser une structure de subdivision hiérarchique des facettes de l'environnement. Cette structure est un arbre dont chaque nœud aura quatre fils (chaque élément de surface étant récursivement découpable en quatre facettes «filles»). A chaque facette est attachée une représentation de la radiosité qu'elle a reçue. Bien évidemment, cette valeur varie avec le temps ce qui impose une structure de donnée permettant de gérer différentes valeurs pour différents instants.

Contrairement aux approches précédentes (*cf.* chapitre 3.4), nous n'assimilerons pas les éléments à des sources et récepteurs ponctuels. Cela implique de prendre en compte la durée de l'échange.

De plus, nous avons choisi de ne pas intégrer la fonction de radiosité par pas de temps constants, comme cela est fait dans [Lew93], par exemple. Nous conservons le temps «exact» de l'arrivée de l'énergie sur un récepteur. Cela nous permettra, comme nous le verrons, d'intégrer dans le processus le traitement de l'énergie échangée lors d'un nombre quelconque de réflexions spéculaires pures qui, elle, peut être associée à un temps de trajet unique.

Représentation de la radiosité temporelle : «échos»

Nous représenterons la fonction de radiosité au cours du temps comme une suite d'«impulsions» énergétiques de durée non nulle, que nous appellerons «*échos*». Un écho représente une certaine intensité sonore qui est émise ou reçue à un instant donné par un élément pendant une certaine durée. On le caractérise par trois paramètres (*cf.* Figure 9.4) :

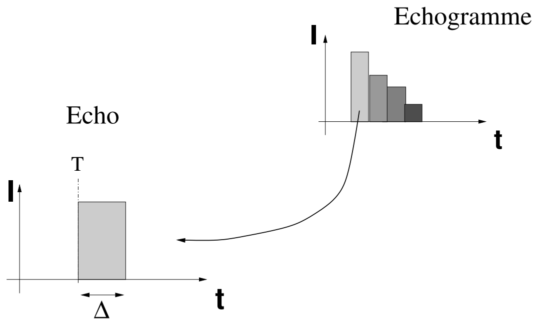
– un *temps de début d'émission* (*resp.* de réception) ;

– une *durée*, l'énergie étant échangée pendant un certain temps après le temps d'émission ;

– l'*intensité* de l'onde sonore échangée pendant cette durée.

Le temps d'émission est utilisé pour classer les échos qui sont reçus par un élément au cours du temps dans des structures semblables à des échogrammes.
Note : Au début du processus, seules les sources principales ponctuelles possèdent un écho à émettre. Cet écho est une «impulsion» de durée nulle.
Ce sont ces échogrammes attachés aux éléments surfaciques qui représentent les variations de l'intensité sonore au cours du temps. On stocke également sur chaque écho son élément de provenance (*cf.* Figure 9.4).

De plus, l'intensité dépend de la fréquence. Chaque écho est valide pour une certaine bande de fréquence faisant de notre approche une simulation en bandes larges.




```
struct Echo
{
    double   width
    double   time
    double   intensity
    Element  shooter
    int frequencyBand
}
```

Figure 9.4 - *L'intensité est représentée au cours du temps par une suite d'impulsions de durée non nulle ou «échos», stockées dans un «échogramme». A chaque écho, on associe le temps T de son début d'arrivée sur un récepteur. On lui associe également une durée d'échange Δ (les éléments ayant une taille non nulle) et l'intégrale de l'intensité acoustique échangée sur cette durée.*

Représentation hiérarchique

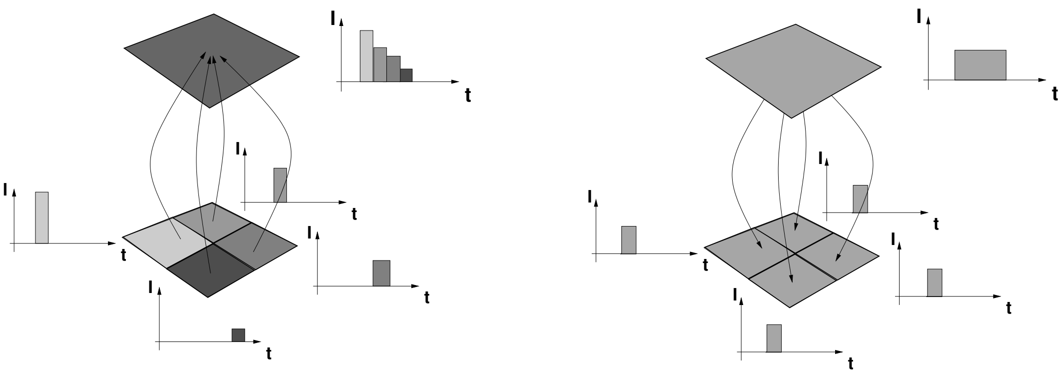
La radiosité ainsi définie peut également être stockée de manière hiérarchique, comme nous l'avons vu au chapitre précédent. Un écho à un niveau donné de la hiérarchie est représenté par quatre échos de même intensité et de même temps d'arrivée au niveau inférieur. Par contre, les durées sont différentes et calculées en considérant la durée initiale pondérée par les rapports d'aires «fils/père» (*cf.* Figure 9.5). Nous discuterons au paragraphe 9.3.2 les inconvénients que peut avoir une telle approche, et nous proposerons une solution susceptible de les limiter.

De manière symétrique, lorsqu'il s'agit de représenter plusieurs échos à un niveau de résolution moindre, on pourrait les combiner en un écho unique. Toutefois, le fait de subdiviser les surfaces permet non seulement d'augmenter la précision énergétique du processus, mais également sa précision temporelle. Nous avons choisi de conserver le même nombre d'échos. On ne modifie pas la structure temporelle dans ce cas mais on pondère l'intensité au *prorata* des rapports d'aires «fils/père» (*cf.* chapitre 8.1.2 et Figure 9.5).

Nous verrons que cette approche a bien évidemment l'inconvénient de multiplier rapidement le nombre d'échos. Nous proposerons une solution à ce problème au paragraphe 9.5.

9.2.2 Facteur de forme «étendus» et liens

Après avoir défini la représentation hiérarchique de la fonction de radiosité temporelle, nous allons maintenant nous intéresser à la modélisation des échanges et en particulier au facteur de forme.



(a) (b)

Figure 9.5 - *Représentation hiérarchique de la radiosité temporelle. (a) Un écho peut être représenté à un niveau supérieur de détail par quatre échos de même temps de propagation et intensité. (b) Lorsqu'ils sont utilisés à un niveau de détail moindre, la distribution temporelle des échos est maintenue alors que leurs intensités est pondérée au* prorata *des rapports d'aires «fils/père».*

Comme on peut le voir dans l'Équation 9.5, nous avons besoin de trois données pour caractériser les échanges entre une paire d'éléments. Ces données sont comparables à celles nécessaires pour représenter la radiosité (*cf.* Figure 9.6) :

– le *temps de propagation* le plus court entre les éléments ;

– la *durée de l'échange* ;

– le *facteur de forme* entre les éléments.

Le temps de propagation le plus court entre une paire de points *mutuellement visibles* (x,y) sur chaque élément est défini comme :

$$T_{ij} = \min_{x \in E_i, y \in E_j} \frac{r_{xy}}{c},$$

où r_{xy} est la distance entre les points.

La durée de l'échange est donnée par :

$$\tau_{ij} = \max_{x \in E_i, y \in E_j} \frac{r_{xy}}{c} - \min_{x \in E_i, y \in E_j} \frac{r_{xy}}{c},$$

où x et y sont là encore mutuellement visibles.

Enfin, le facteur de forme est calculé comme dans les approches de radiosité classiques par :

$$F_{ij}^{\lambda} = \frac{1}{A_i} \int_{x \in E_i} \int_{y \in E_j} \frac{\cos\theta \cos\theta'}{\pi r^2} V(x,y) \mu^{\lambda}(r) dxdy,$$

où $V(x,y)$ est le terme de visibilité entre les points x et y.

L'influence de la fréquence sur le facteur de forme reste limitée dans ce cas au coefficient d'atténuation atmosphérique. Néanmoins, on peut modifier ce facteur de forme pour prendre en compte un terme de visibilité fréquentielle V^{λ} comme présenté au Chapitre 6.

Ces trois termes constituent un «facteur de forme étendu[1]» que nous noterons $\mathcal{F}_{ij}^{\lambda}$. Le calcul de ces facteurs de forme étendus implique *a priori* un échantillonnage des surfaces puisque le facteur de forme ne peut être obtenue analytiquement dans le cas général. De plus, il est également difficile et coûteux de connaître la plus petite et la plus grande distance entre deux points respectifs sur deux éléments donnés. Nous donnons en Annexe B des détails supplémentaires sur le calcul des trois termes du facteur de forme étendu par échantillonnage des éléments.

```
struct ExtendedFormFactor
{
    double   delay
    double   duration
    double   formFactor
}
```

```
struct Link
{
    Element   source
    Element   receiver
    ExtendedFormFactor  ff
    int       frequencyBand
}
```

Figure 9.6 - *Structures de données pour les facteurs de forme étendus et les liens.*

Liens dépendant de la fréquence

Nous allons également étendre les liens entre les éléments afin de prendre en compte l'aspect fréquentiel de la simulation. Dans les approches de radiosité en synthèse d'image, on utilise généralement une radiosité dépendant de la fréquence. Souvent le calcul est effectué sur trois bandes de fréquences lumineuses : rouge, vert et bleu, correspondant à la base des couleurs des moniteurs informatiques. Toutefois, lorsqu'il s'agit de prendre en compte des sources dont le spectre d'émission est complexe, il est souhaitable d'utiliser plus de bandes (*i.e.* un spectre) afin d'obtenir une représentation plus exacte de la couleur [Zeg97]. Néanmoins, la plage des longueurs d'ondes visibles est très étroite. Le déroulement de la simulation et en particulier le raffinement des liens ne dépend donc pas en grande mesure de la longueur d'onde. En comparaison, les longueurs d'onde audibles sont réparties sur trois ordres de grandeurs (comme on peut le voir dans la Table 1.1), ce qui implique une plus grande influence de la longueur d'onde sur le processus de simulation. D'autre part, si l'on veut pouvoir prendre en compte des interférences dans notre modèle de simulation, il est nécessaire de pouvoir contrôler le niveau de raffinement en fonction de la fréquence (comme dans des approches par élément finis de frontière). En effet, les temps de propagation entre éléments (et donc les éventuelles interférences) dépendent bien entendu de leur taille.

Nous avons choisi d'établir des liens valides pour une bande de fréquence donnée (*cf.* Figure 9.6). Comme nous le verrons par la suite ces liens peuvent être raffinés en fonction de critères dépendant de la fréquence.

9.3 Transport hiérarchique de la radiosité temporelle

Après avoir défini les données nécessaires aux échanges, nous abordons maintenant le problème du transport de l'énergie au cours du temps entre éléments et à travers la hiérarchie.

1. on ne le confondra pas avec les facteurs de forme étendus utilisés en simulation de l'éclairage pour la simulation des échanges spéculaires par des méthodes de radiosité.

9.3.1 Opérateur de transport

Lorsqu'un écho est échangé d'une surface à l'autre le long d'un lien ses paramètres sont modifiés selon l'équation 9.5. Ainsi, comme on peut le voir sur la Figure 9.7, son énergie est multipliée par le facteur de forme associé à l'échange. Après le transfert, il représente de l'énergie sur la surface destination mais à un temps plus grand, puisque le temps de transfert T_{ij} n'est pas nul. Enfin, la durée de l'écho se trouve étendue par la durée du transfert associée au lien. On remarquera que cette formulation revient à exprimer l'intensité en avançant dans le temps, plutôt qu'en fonction des instants précédents comme dans l'équation 9.5.

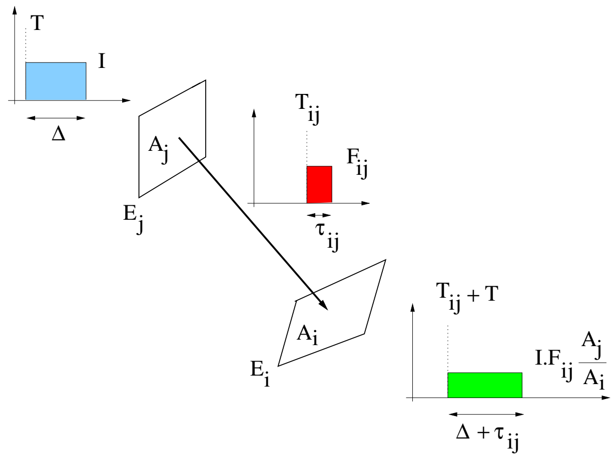


Figure 9.7 - *Opérateur de transport des échos d'un élément à un autre.*

On transporte de manière itérative tous les échos d'élément en élément le long des liens. Chaque itération du processus correspond à une réflexion de l'onde sur les parois. Comme dans les méthodes classiques de simulation en acoustique, on suppose que le processus a convergé lorsque l'intensité acoustique totale à ré-émettre a diminué d'un facteur 10^{-6}. Chaque élément dans la hiérarchie enregistre les différents échos qui l'ont atteint au cours du temps, constituant la solution du processus. Cette solution étant stockée sur les surfaces, on peut la reprojeter directement en n'importe quel point de réception dans l'environnement à peu de frais.

Afin de réaliser cette opération, il est nécessaire de maintenir trois listes d'échos ou échogrammes sur chaque élément :

- un *échogramme d'émission* qui contient tous les échos devant être ré-émis par l'élément à l'itération courante ;
- un *échogramme de réception* qui contient tous les échos reçus par l'élément (à son niveau dans la hiérarchie) à l'itération courante ;
- un *échogramme solution* qui contient tous les échos reçus par l'élément (à son niveau dans la hiérarchie) depuis le début du processus. C'est donc lui qui représente la *solution de radiosité temporelle*.

Afin de limiter le coût mémoire impliqué par le stockage des échogrammes, nous avons choisi de conserver dans l'échogramme solution uniquement la liste des échos reçus par un élément *à son niveau de hiérarchie*. Cela implique une traversée de la hiérarchie pour reconstruire la solution complète.

```
void gatherLinks()
{
    foreach l in links
    {
        Element shooter = l→shooter
        Element gatherer = l→gatherer
        double FF = l→ff.formFactor
        double t = l→ff.delay
        double τ = l→ff.duration
        double A_s = shooter.getArea()
        double A_g = gatherer.getArea()
        foreach echo in shooter→shootingEchogram
        {
            double I = echo.intensity
            double T = echo.time
            double Δ = echo.width
            gatherer→gatheringEchogram →insertEvent(I*FF*A_s/A_g, T+t, Δ+τ)
        }
    }
}
```

Figure 9.8 - *Pseudocode pour l'émission/réception des échos.*

9.3.2 Transport hiérarchique

Comme dans les approches de radiosité hiérarchique traditionnelles, chaque échange est effectué à un niveau précis de la hiérarchie. La radiosité reçue par un élément dépend de celle reçue par ses pères et ses fils. Après chaque itération une opération de «push-pull» est donc là aussi nécessaire pour maintenir une information cohérente à tous les niveaux de la hiérarchie.

Push-pull

Durant l'opération de «pull» tous les échos reçus par un élément sont transmis à ses fils et ce de manière récursive à travers toute la hiérarchie. Une fois les échos transmis jusqu'aux feuilles, ils sont alors «remontés» jusqu'à l'élément de départ suivant la méthode présentée au paragraphe 9.2.1.

L'algorithme est récapitulé dans la figure 9.9. Durant ces deux phases les échos reçus dans les échogrammes de réception sont combinés à différents niveaux de hiérarchie et transférés dans les échogrammes d'émission pour l'itération suivante. Comme on peut le constater l'opération de «push-pull» est coûteuse puisqu'il faut combiner entre eux tous les échogrammes à chaque niveau de la hiérarchie. De plus, puisque nous avons choisi de ne pas maintenir une représentation de la solution à tous les niveaux de la hiérarchie, cette opération n'est utile qu'au niveau ou des échanges sont réellement effectués (l'énergie émise par un éléments dépend effectivement de tout ce qui a été reçu dans sa hiérarchie).

Afin de limiter le coût de l'opération, nous proposons de maintenir des quantités scalaires correspondant à la somme des intensités des échos contenus dans un échogramme. Ces valeurs scalaires sont maintenues comme des valeurs de radiosité hiérarchique par un «push-pull» classique peu coûteux et sont utilisées pour les décisions concernant le raffinement. Lorsque le niveau de raffinement a été établi pour un lien, conformément au critère fixé, le véritable «push-pull» des échogrammes est effectué pour l'élément source et les échos sont échangés.

```
void pushPull(Element e)
{
    e→shootingEchogram→clear()
    e→shootingEchogram→merge(e→gatheringEchogram)

    if (e→father!=NULL)
        mergeFathersGatheringEchograms(e→shootingEchogram,e→father)

    foreach son in e→sons
        mergeSonsGatheringEchograms(e→shootingEchogram,son,e→area)
}

void mergeFathersGatheringEchograms(Echogram ech, Element f)
{
    ech→merge(father→gatheringEchogram)

    if (f→father!=NULL)
        mergeFathersGatheringEchograms(ech,f→father)
}

void mergeSonsGatheringEchograms(Echogram ech, Element s, double areaFactor )
{
    ech→merge(s→gatheringEchogram, s→area/areaFactor)

    foreach son in s→sons
        mergeSonsGatheringEchograms(ech,son,areaFactor)
}
```

Figure 9.9 - *Pseudo-code pour l'opération de «push-pull» sur les échogrammes. La fonction «merge» insert les échos contenus dans l'échogramme passé en paramètre dans l'échogramme appelant. Si un paramètre scalaire additionnel est donné, l'intensité de chaque écho est multipliée en conséquence avant l'insertion.*

Artefacts temporels et correction du délai de propagation

Un inconvénient du transport temporel hiérarchique, tel que nous l'avons défini, est que l'on risque de sous-estimer les temps de propagation. Considérons la Figure 9.10. L'élément E_1 échange de l'énergie avec l'élément E_2. Lors du «push-pull» cette énergie est transférée au niveau de l'élément E_3, père de E_2 avec le même temps d'arrivée. Lorsque celui-ci échange enfin son énergie avec E_4, le temps de parcours est inférieur à celui que l'énergie aurait eu si elle avait échangée entre E_2 et E_4. Afin de minimiser ce problème, on peut introduire au moment de l'opération de «pull» une compensation au niveau du temps de propagation. Pour cela on calcule la différence entre le temps de propagation de centre à centre de E_1 à E_3 puis de E_1 à E_2. La différence de ces deux temps est alors ajoutée au temps de chaque écho lors du «pull» pour compenser le temps de transfert plus court au niveau de hiérarchie plus élevé. De la même manière, on peut également compenser la largeur des échos qui, elle, va tendre à être sur-estimée.

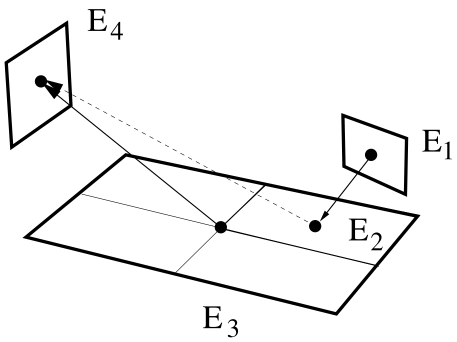


Figure 9.10 - *Artefacts temporels dûs au transport hiérarchique. Le transport à plus haut niveau dans la hiérarchie peut conduite à sous-estimer le temps de propagation entre éléments. Le temps de transport entre E_1 et E_2 puis entre E_3 et E_4 est plus court que le temps de transport entre E_1 et E_2 puis E_2 et E_4.*

9.4 Raffinement des échanges

Comme nous l'avons déjà vu au chapitre précédent, le point clé d'une approche hiérarchique est sans nul doute l'oracle de raffinement. En effet, c'est lui qui va permettre d'orienter les calculs de la manière la plus efficace possible (en temps comme en précision).

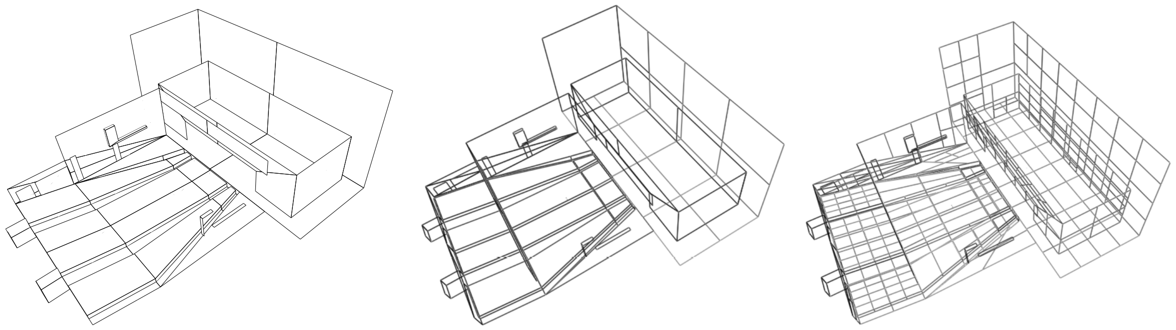
Une première possibilité de raffinement est un critère énergétique (*cf.* chapitre 8.2.2). On raffine lorsque la somme des intensités des échos à émettre multipliée par le facteur de forme est supérieure à un seuil prédéfini par l'utilisateur. L'intégrale du facteur de forme étant calculée par échantillonnage aléatoire des surfaces, on utilise la variance de l'intégrande plutôt que le facteur de forme lui-même. Cet oracle permet de subdiviser dans les zones où le facteur de forme point à point varie beaucoup.

Toutefois un tel oracle n'est pas suffisant puisque, la qualité des échanges dépend également de la bonne approximation des échanges temporels. Nous proposons donc d'utiliser également un oracle qui raffine lorsque la durée d'un échange entre facettes est trop importante. Cet oracle tend aussi à imposer que le temps de propagation point à point entre deux élément soit uniforme. Dans ce cas, on peut effectivement considérer que les interférences sont constructives et que l'on peut additionner les intensités des échos. Ce dernier point dépend bien évidemment de la fréquence. Nous avons donc choisi de définir notre critère en fonction de celle-ci. En conséquence, notre oracle raffine un lien si :

$$\tau_{ij} \geq T_\varepsilon \frac{c}{\lambda},$$

où τ_{ij} est la durée associée à l'échange et T_ε est un pourcentage fixé par l'utilisateur. $\frac{c}{\lambda}$ est la *période* de l'onde considérée.

Cet oracle étant purement géométrique, on peut commencer par raffiner les éléments en l'utilisant dans une phase de pré-traitement. La figure 9.11 en montre un exemple d'utilisation et le raffinement induit sur les surfaces. Comme on peut le constater, il raffine très rapidement les éléments lorsque la fréquence augmente.



(a) (b) (c)

Figure 9.11 - *Exemple de raffinement dépendant de la fréquence. La valeur du coefficient T_ε est 1. (a) Surfaces initiales. (b) Raffinement obtenu à 50 Hz. (c) Raffinement obtenu à 100 Hz.*

Une meilleure solution serait sans doute d'étendre les facteurs de forme pour prendre en compte une notion de «phase». Le facteur de forme serait alors une quantité complexe :

$$\hat{F}_{ij}^{\lambda} = \frac{1}{A_i} \int_{x \in E_i} \int_{y \in E_j} \frac{\cos\theta \cos\theta'}{\pi r^2} V(x,y) \mu^{\lambda}(r) e^{ikr} dxdy,$$

Cela permettrait de prendre en compte des interférences lors de l'échange. De plus on pourrait également comparer le facteur de forme F_{ij}^{λ} et le module de $\hat{F}_{ij}^{\lambda}$ et raffiner si la différence des deux quantités est trop importante. L'utilisation de ces facteurs de formes dans un oracle énergétique comme nous l'avons présenté précédemment permettrait enfin de combiner raffinement énergétique et raffinement fréquentiel dû à l'existence d'interférences.

Nous n'avons pas essayé de mettre en pratique l'utilisation de ces facteurs de forme complexes, négligeant de ce fait la prise en compte d'interférences au moment du transfert. Néanmoins, nous proposons à présent une méthode qui peut traiter des phénomènes d'interférence entre échos provenant de plusieurs éléments différents et incidents, à des temps voisins, sur un élément donné.

9.5 Limitation de la complexité dans le domaine temporel et théorie ondulatoire

Afin de limiter le nombre d'échos, qui croît de manière exponentielle, nous proposons un moyen de contrôler la complexité directement dans le domaine temporel, au moment de l'insertion des échos dans les échogrammes. L'approche est basée sur une combinaison des échos atteignant un élément à des instants voisins. Toutefois, nous ne proposons pas d'«additionner» directement l'intensité de ces échos, mais de prendre en compte leur retard relatif afin de pouvoir simuler d'éventuelles interférences.

9.5.1 Fusion d'échos et raffinement dans le domaine temporel

Le principe de notre approche est de chercher à fusionner les échos entre eux pour en diminuer le nombre. Chaque fois qu'un écho est inséré dans un échogramme, on vérifie si un autre écho a déjà été inséré à un temps voisin. Si tel est le cas nous les additionnons pour produire un écho unique.

Afin de pouvoir les combiner en prenant en compte d'éventuelles interférences, nous allons considérer chaque écho comme une représentation de la pression acoustique échangée entre les éléments.

Comme nous l'avons vu au Paragraphe 1.1, l'intensité acoustique varie comme le carré de la pression acoustique. On peut alors représenter la pression acoustique correspondant à l'écho considéré sous la forme :

$$\hat{P} = \sqrt{I}e^{i\omega T},$$

où I est l'intensité et T le temps de propagation de l'écho. On a également $\omega = \frac{2\pi c}{\lambda}$.

Considérant deux échos, on peut écrire :

$$\hat{P}_1 = \sqrt{I_1}e^{i\omega T_1}, \hat{P}_2 = \sqrt{I_2}e^{i\omega T_2}.$$

On a alors (*cf.* Figure 9.12) :

$$\hat{P}_1 + \hat{P}_2 = \hat{P}_1(1 + \frac{\hat{P}_2}{\hat{P}_1}) = \sqrt{I}e^{(\omega(T_1+dt))},$$

avec

$$I = I_1 + I_2 + \sqrt{I_1 I_2}\cos(\omega(T_2 - T_1))$$

et

$$dt = \frac{1}{\omega}\arctan(\frac{\sqrt{I_2}\sin(\omega(T_2 - T_1))}{\sqrt{I_1} + \sqrt{I_2}\cos(\omega(T_2 - T_1))}).$$

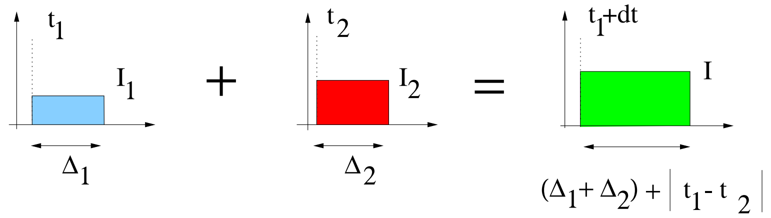


Figure 9.12 - *Fusion d'échos. Deux échos incidents sur un même élément à des instants voisins peuvent être combinés en un seul.*

Comme on peut le voir sur la Figure 9.12 nous attribuons une nouvelle durée aux échos ainsi générés de manière à couvrir la somme des durées des deux échos de départ.

Cet algorithme de fusion d'échos implique de maintenir les échogrammes triés suivant le temps de propagation et d'utiliser une structure de donnée permettant une recherche et une insertion très efficaces. Nous avons choisi pour cela d'implémenter nos échogrammes sous formes d'arbres binaires équilibrés permettant une recherche logarithmique en fonction du nombre d'échos.

9.5.2 Contrôle de la fusion

Nous proposons de contrôler le processus de fusion d'échos en utilisant un paramètre similaire à notre oracle dépendant de la fréquence. Nous combinons deux échos si le temps qui les sépare est inférieur à un certain pourcentage de la période de l'onde considéré que nous appellerons *facteur de fusion*.

La Figure 9.13 montre quatre échogrammes résultant de simulations réalisées avec des valeurs différentes du facteur de fusion. Comme on peut le constater, le nombre d'échos diminue mais malgré cela l'intensité n'a pas tendance à être localement sur-estimée.

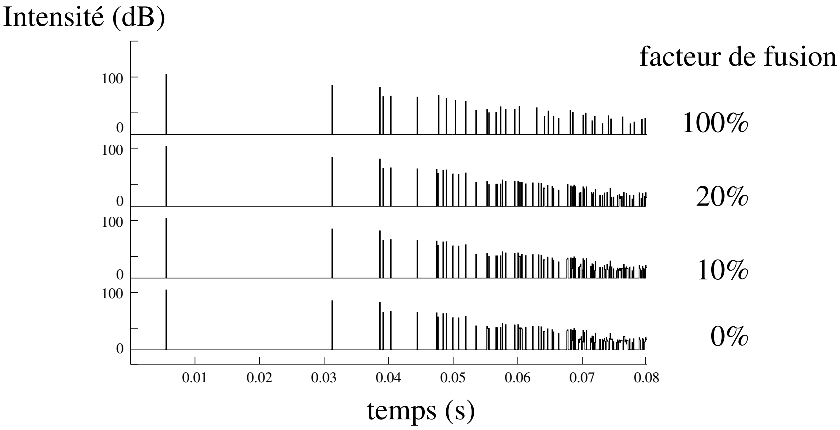


Figure 9.13 - *Influence de la fusion d'échos sur l'énergie. Comparaison de quatre échogrammes résultant de simulations réalisées avec différentes valeurs du facteur de fusion. On notera que l'énergie ne tend pas à être sur-estimée par le processus de fusion.*

La Figures 9.14 montre une comparaison du nombre d'échos générés pendant les six premières itérations d'une simulation. Lorsqu'on n'utilise pas la fusion, le nombre d'échos croît de manière exponentielle. Un facteur de fusion de 2% permet de réduire de manière significative le nombre d'échos. Avec un facteur de 20%, on gagne deux ordres de grandeur environ. La validité physique d'une telle approche demande cependant encore à être vérifiée. Cela implique de réaliser des comparaisons quantitatives avec des approches par éléments finis par exemple, que nous n'avons pas pu réaliser dans le cadre de cette étude.

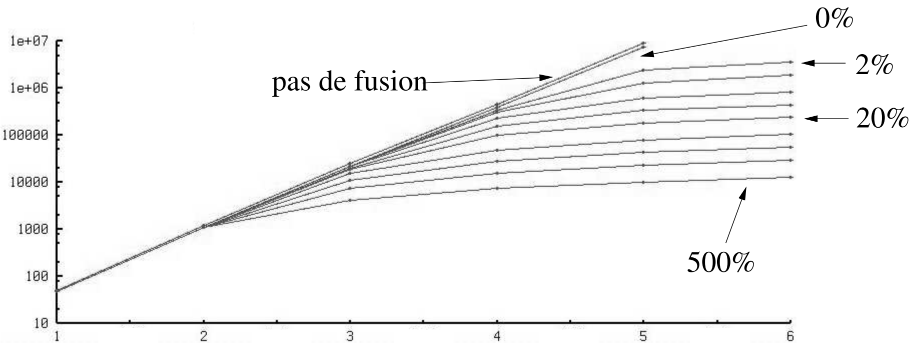


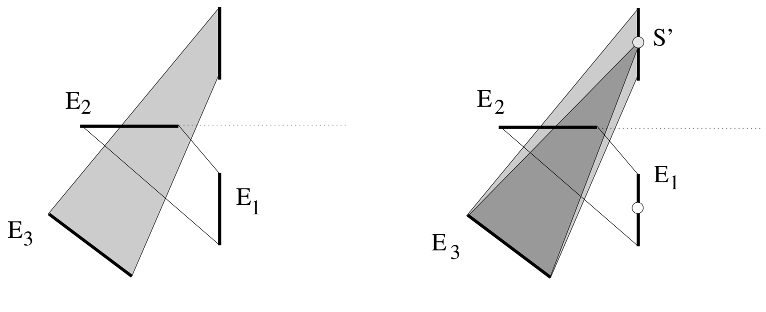
Figure 9.14 - *Comparaison du nombre d'échos en fonction de l'ordre de réflexion pour différentes valeurs du facteur de fusion. De 0% de la période (courbe du haut) à 500% de la période (courbe du bas).*

9.6 Traitement unifié des spécularités à l'aide de sources-images

Comme nous l'avons vu, les réflexions spéculaires sont un modèle privilégié dans les simulations acoustique. En effet, pour les longueurs d'ondes moyennes, les surfaces peuvent rapidement être assimilés à des réflecteurs parfaitement plans et rigides. Nous proposons d'étendre la méthode d'échanges radiatifs hiérarchiques que nous avons détaillée dans les paragraphes précédents à un modèle de réflexion des surfaces combinant une composante lambertienne idéale et une composante spéculaire idéale.

Pour cela nous proposons d'utiliser des sources-images. Ces sources-images permettent de représenter l'énergie échangée par réflexion spéculaire entre les éléments. Elles permettent également de représenter de manière exacte les chemins de propagation ne comportant que des réflexions spéculaires depuis les sources principales ponctuelles jusqu'aux récepteurs également ponctuels. Ces derniers correspondent aux sources-images classiques que nous avons déjà présentées et utilisées.

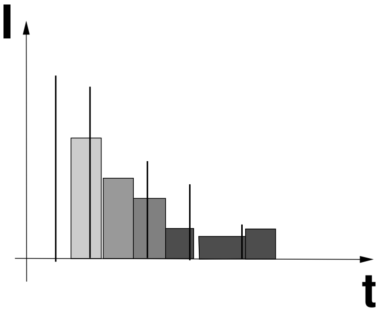
Afin de simplifier le problème, nous allons, pour les échanges spéculaires, assimiler les surfaces à leur centre comme on peut le voir sur la Figure 9.15. Cela nous permet d'utiliser un traitement unifié pour les sources ponctuelles et surfaciques, sachant que ces dernières peuvent être subdivisées.



(a) (b)

Figure 9.15 - *Utilisation du centre des éléments comme source-image pour le calcul des échanges spéculaires. Le centre de l'élément d'origine de l'écho E_1 est réfléchi par rapport à l'élément source E_2 et un facteur de forme étendu est calculé entre le point obtenu (S') et l'élément de destination E_3 .*

Afin de traiter les spécularités, on ajoute à la structure de donnée de l'écho que nous avons présenté dans la Figure 9.4 un point de provenance qui représente la source-image associée. Ces nouveaux échos sont échangés entre les surfaces de la même manière que précédemment.




```
struct Echo
{
    double      width
    double      time
    double      intensity
    Element     shooter
    int         frequencyBand
    Vector      origin
    double      gain
}
```

Figure 9.16 - *Extension de la structure de donnée d'un écho pour traiter les spécularités. Le point d'origine de l'écho est conservé permettant de construire une source-image. Dans le cas d'un élément surfacique, ce point est le centre de l'élément. Des échos particuliers de durée nulle vont être utilisés pour représenter de manière exacte les chemins correspondant uniquement à des réflexions spéculaires de l'onde depuis les sources primaires ponctuelles*

9.6.1 Construction des sources-images à la demande et échanges spéculaires

Le principe de la construction de ces sources-images est simple. Chaque fois qu'un élément reçoit les échos en provenance des éléments auxquels il est lié, un traitement est effectué si l'élément source est spéculaire.

Pour chaque écho on construit la source-image correspondant au symétrique de son point d'origine par rapport au plan de l'élément source. Un facteur de forme étendu est calculé entre cette source-image et l'élément récepteur. L'énergie correspondante est échangée (en utilisant le coefficient de réflexion spéculaire du réflecteur) sous forme d'un nouvel écho. Cette idée est similaire à celle du «transport à trois points» évoquée dans le paragraphe 8.3.2. Le calcul du facteur de forme est différent puisqu'il implique de prendre en compte la validité de la source-image. On le calcule par échantillonnage en vérifiant que les différents rayons entre échantillons intersectent bien l'élément source.

Si la surface est également diffusante, on insert un autre écho correspondant à l'échange diffus comme nous l'avons décrit dans les paragraphes précédents.

9.6.2 Traitement des chemins spéculaires

Nous allons également utiliser les échos pour représenter les chemins de propagation correspondant à un nombre arbitraire de réflexions uniquement spéculaires de la source sur les parois. Les échos correspondants à de tels chemins ont une largeur nulle puisqu'ils correspondent à des chemins point à point.

Afin de conserver l'aspect «indépendant» du point de réception de notre simulation nous allons propager ces échos afin de conserver sur les éléments toutes les réflexions des sources principales. Pour cela, lorsqu'un élément reçoit un écho de largeur nulle, il insert dans son échogramme de réception un nouvel écho de largeur nulle dont le point d'origine est, là encore, le symétrique de celui de l'écho initial par rapport à l'élément source. Il faut également prendre en compte dans ce cas la validité de la source-image avant de l'insérer. On peut remarquer que de telles sources images ne sont définies que pour les surfaces initiales du processus, c'est-à-dire les racines des hiérarchies de surfaces. Or, dans le cas où une surface est subdivisée, on crée pour chaque élément différentes instances de la même source image. Ce problème peut être résolu en modifiant notre critère de fusion d'échos de manière à fusionner les échos de durée nulle et provenant du même point. Ainsi, les différentes instances sont fusionnées au moment du «push-pull».

A ces échos particulier on associe, en plus de leur intensité, un gain correspondant simplement le produit des réflectances des parois rencontrées. Lorsqu'ils sont finalement reprojetés sur le récepteur ponctuel final, on peut alors calculer leur intensité exacte en fonction de leur temps d'arrivée et de ce produit.

```
void gatherLinks()
{
    foreach l in links
    {
        Element shooter = l→shooter
        Element gatherer = l→gatherer
        double FF = l→ff.formFactor
        double t = l→ff.delay
        double τ = l→ff.duration
        double A_s = shooter.getArea()
        double A_g = gatherer.getArea()
        foreach echo in shooter→shootingEchogram
        {
            if gatherer→isSpecular()
            {
                boolean valid = buildImageSource(echo.origin,t,ff)
                if (valid)
                {
                    double I = echo.intensity
                    double Δ = echo.width
                    gatherer→gatheringEchogram →insertEvent(I*FF*4π/A_g, t, Δ, reflectionCoeff*echo.gain)
                }
            }
            if gatherer→isDiffuse()
            {
                double I = echo.intensity
                double T = echo.time
                double Δ = echo.width
                gatherer→gatheringEchogram →insertEvent(I*FF*A_s/A_g, T+t, Δ+τ)
            }
        }
    }
}
```

Figure 9.17 - *Pseudo-code pour l'opération de «récolte» sur les échogrammes avec spécularités.*

9.7 Implémentation et résultats : le module *ECHO*

Nous donnons ici quelques détails supplémentaires concernant l'implémentation de notre approche. Nous présenterons également des résultats dans le cadre de la visualisation de la propagation d'ondes sonores.

9.7.1 Implementation de la radiosité hiérarchique sonore

Nous avons implémenté cette méthode en utilisant un «déraffinement progressif» semblable à l'approche décrite dans [SSSS98, Gra98]. C'est-à-dire que l'on va raffiner en utilisant un critère sur l'énergie restant à émettre. A chaque itération, correspondant à une réflexion supplémentaire, on va donc utiliser des liens de plus haut niveau, puisqu'il restera de moins en moins d'énergie à échanger. Les itérations se terminent lorsqu'il ne reste plus suffisamment d'énergie à échanger. Cette approche à l'inconvénient d'augmenter l'erreur à chaque itération. De plus, elle n'exploite pas un caractère global sur la durée de la simulation pour choisir les liens à raffiner. Toutefois elle permet une implémentation assez simple et directe.

On peut néanmoins utiliser un raffinement plus «global» sur toute la durée de la simulation. Pour cela, on commence par effectuer toutes les itérations jusqu'à convergence en utilisant les liens de plus haut niveau. On a donc déjà une bonne approximation des échanges sur toute la durée de la simulation. Ensuite, on sélectionne les liens à raffiner, on corrige la solution en conséquence et on ré-itère le processus d'échange jusqu'à nouvelle convergence. On poursuit cette approche jusqu'à ce qu'il ne reste plus de liens à raffiner.

L'avantage de cette approche est d'utiliser au moment du raffinement un résultat global des échanges. On va donc raffiner de manière plus appropriée. Par contre, lorsque l'on va raffiner un lien, il va falloir être capable d'éliminer toute l'énergie qui a transité le long de ce lien au cours du temps pour la remplacer par celle qui va transiter par les nouveaux liens.

On va donc commencer par reconstruire toute l'énergie reçue par l'élément raffiné au cours du temps. Cela est effectué par une étape de «push-pull» de tous les échogrammes historiques des éléments dans sa hiérarchie[2]. Puis, on commence par ré-émettre cette énergie le long de l'ancien lien avec un coefficient négatif. Cela revient à ré-insérer tous les échos correspondant avec des intensités négatives dans l'échogramme du récepteur. Enfin, on ré-émet encore une fois cette énergie, cette fois positive, le long des nouveaux liens crées. Un nouveau «push-pull» est alors nécessaire avant de reprendre le processus d'itération jusqu'à convergence.

9.7.2 Résultats

Une fois la solution de radiosité calculée sur les surfaces, on peut créer une animation de la propagation de l'énergie sur les parois. Les Figures 9.18 à 9.20 présentent des résultats de telles animations dans des environnement simples et dans le cas de surfaces lambertiennes et spéculaires idéales. Les temps de calculs pour les exemples présentés sont de l'ordre de la demi-minute sur une station SGI O_2 R5000 150 MHz.

2. on rappelle qu'on ne maintient pas en permanence une version hiérarchique cohérente de ces échogrammes.

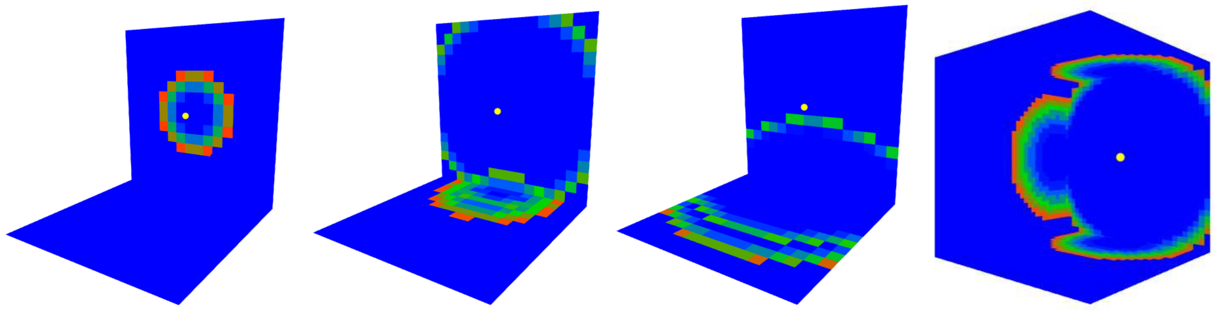
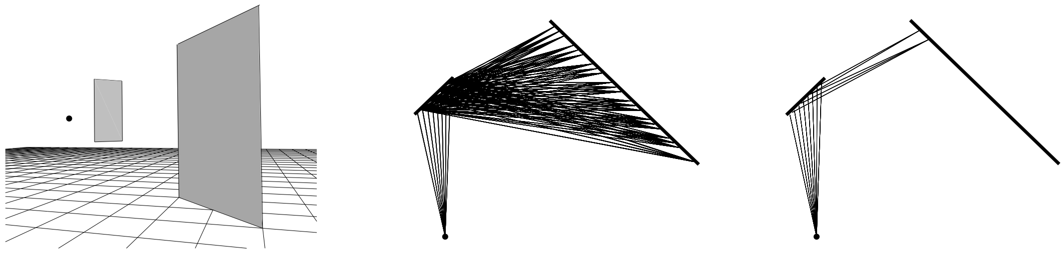


Figure 9.18 - *Visualisation de la propagation d'un front d'onde avec un ordre de réflexion spéculaire en affichant notre solution de radiosité au cours du temps. Le point blanc représente la source de l'onde. On notera sur la seconde et la troisième image, les deux fronts d'onde secondaires dûs à des réflexions spéculaires idéales sur les parois. Les solutions ont été obtenues avec 685 éléments et 132610 liens pour le dièdre. Dans le cas du cube 8193 éléments et 6188 liens ont été utilisés.*



(a) (b) (c)

Figure 9.19 - *Cas diffus et cas spéculaire. (a) Scène de test utilisée. Le point représente une source ponctuelle. (b) Vue du dessus montrant les liens établis par l'élément central dans le cas où les surfaces sont lambertiennes idéales. (c) Vue du dessus montrant les liens établis par l'élément central dans le cas où les surfaces sont spéculaires idéales. Seul un nombre réduit de liens va effectivement transporter de l'énergie.*

La Figure 9.21 montre une comparaison de la solution obtenue pour le premier ordre de réflexion dans une pièce parallélépipédique aux parois spéculaires, dont le plafond est également diffusant.

Nous présenterons dans le chapitre suivant des exemples dans des environnements plus complexes et des résultats plus quantitatifs de notre méthode appliquée à l'évaluation de la qualité acoustique de lieux d'écoute.

9.8 Auralisation de la solution de radiosité temporelle

Dans les paragraphes précédants, nous avons décrits les différentes caractéristiques de notre approche, ainsi que le déroulement de la simulation. La solution obtenue à la fin du processus est une distribution de l'intensité acoustique au cours du temps. Afin de pouvoir effectuer l'auralisation de cette solution, il est nécesaire de reconstruire la réponse impulsionnelle en pression.

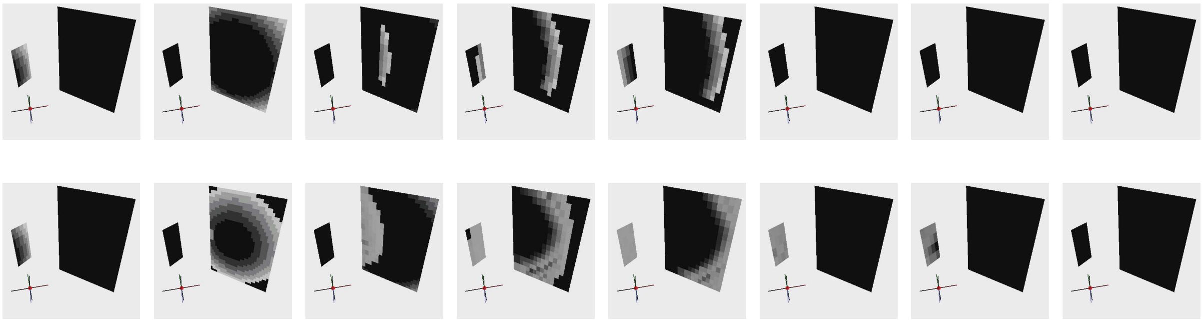


Figure 9.20 - *Comparaison entre deux solutions obtenues dans des cas spéculaires et diffus idéaux. La séquence d'image supérieure présente une animation de la solution de radiosité temporelle obtenue dans l'environnement décrit dans la Figure 9.19 avec des surfaces spéculaires idéales. La séquence du bas présente une solution à des instants voisins dans le même environnement mais avec des surfaces lambertiennes idéales. On remarquera la différence entre les fronts d'onde secondaires. On notera également dans le cas diffus, la présence de zones de plus faible intensité (dans le front d'onde réfléchi, 4ème vignette en bas) dues à des interférences.*

9.8.1 Reconstruction de la réponse impulsionnelle en pression

La reconstruction de la réponse en pression implique deux phases :

- reconstruire la réponse énergétique à partir des données des échos pour chaque bande de fréquence considérée;
- pour chacune des réponses énergétiques obtenues, reconstruire une réponse en pression.

Reconstruction de la réponse en intensité à partir des échos

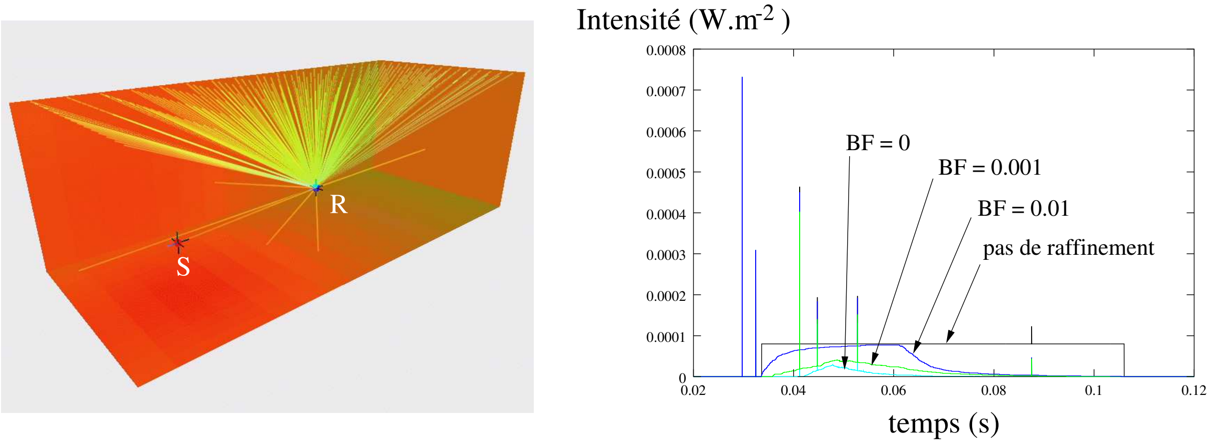
La première étape pour reconstruire la réponse en intensité est de choisir une fréquence déchantillonnage. L'avantage ici est que la représentation de la solution sous forme d'échos est indépendante de la fréquence d'échantillonnage choisie ce qui permet de ne la fixer qu'en tout dernier lieu. Une fois la fréquence choisie, les échos vont être convertis en une réponse discrète. On rappelle que dans le cas général de parois diffuses et spéculaires, la solution finale obtenue au point de réception considéré est une suite d'échos «ponctuels» (correspondant aux sources-images) et d'échos étendus (correspondant à des chemins comportant une ou plusieurs réflexions diffuses).

Les échos «ponctuels» ont une durée instantanée qui correspond à un échantillon à la fréquence considérée. Dans le cas d'échos étendus, on génère une suite d'échantillons dont la valeur est celle de l'intensité de l'écho. Le nombre et le temps de départ de la contribution sont déterminés par le temps et la largeur de l'écho.

Ce processus est répété pour chaque écho, leurs contributions étant ajoutées successivement à la réponse énergétique.

Génération de la réponse en pression

La réponse finale en pression est obtenue à partir de la réponse en intensité en prenant la racine carrée des valeurs des échantillons. En effet, nous rappelons que la pression acoustique varie comme le carré de l'intensité (*cf.* Chapitre 1.1 et Annexe A.1).



(a) (b)

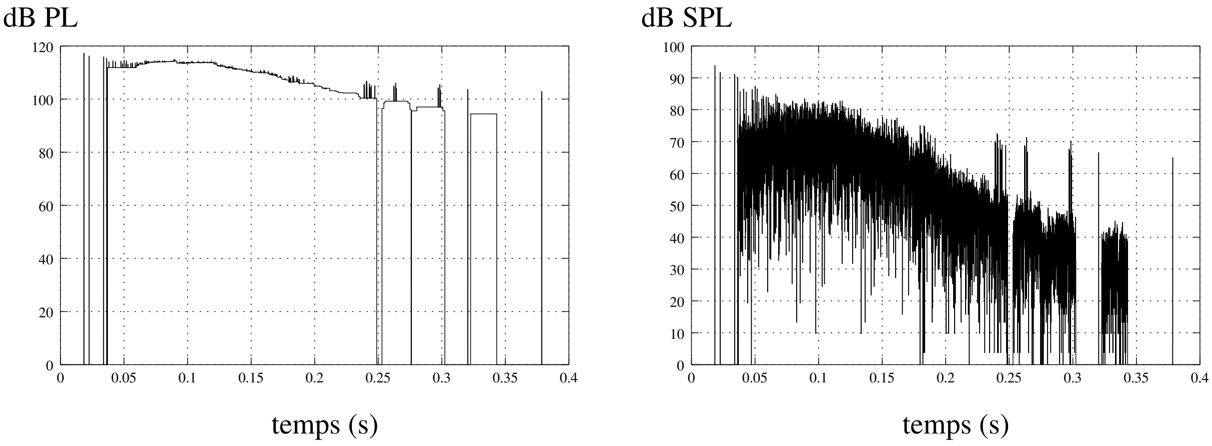
Figure 9.21 - *(a) La scène de test est une salle en forme de boite avec des murs spéculaires et un plafond diffusant (le plafond à également une composante spéculaire). La figure montre les liens ayant effectivement transporté de l'énergie après le premier ordre de réflexion. (b) Courbe décrivant l'intensité acoustique reçue par le récepteur au cours du temps pour le premier ordre de réflexion. Le processus de raffinement permet d'accroître la précision de la valeur de l'énergie ainsi que son temps d'arrivée. On remarquera que le début d'arrivée de l'énergie diffuse provenant du plafond tend à se décaler vers le temps d'arrivée de la réflexion spéculaire correspondante.*

De plus notre processus restant énergétique, nous n'avons pas d'informations sur la phase du signal. On ne connait donc pas la manière dont la pression va fluctuer au cours du temps. Pour reconstruire les échos étendus, nous allons donc faire l'hypothèse d'une phase aléatoire, comme dans [Kut91b]. En pratique, ceux-ci sont reconstruits en premier lieu. La réponse obtenue va ensuite être utilisée pour fenêtrer un bruit blanc. Les contributions spéculaires seront ajoutées en dernier lieu, puis on prendra la racine carrée de chaque échantillon. Un exemple de réponse énergétique et sa réponse équivalente en pression acoustique est donné en Figure 9.22. Dans le cas où la simulation est effectuée pour différentes bandes de fréquences, une réponse est générée pour chacune. Les réponses obtenues sont ensuite filtrées par le filtre passe-bande correspondant avant d'être ajoutées (*cf.* Paragraphe 4.1.3).

Une solution intéressante ici pourrait être de considérer les réflexions spéculaires de manière indépendante et de les reconstruire avec une approche en bandes fines comme nous l'avons vu au Chapitre 7. Cela permettrait sans doute d'obtenir de meilleurs résultats dans le cas des premières réflexions qui sont perceptivement importantes. La réverbération tardive ainsi que les réflexions diffuses peuvent être reconstruites avec une approche «larges bandes» puisque qu'elles contienent moins d'information de détail du point de vue perceptif.

9.8.2 Auralisation en environnements dynamiques

Nous avons également implémenté un moteur de rendu sonore similaire à celui présenté au Chapitre 7 dans le contexte de notre simulation d'échange radiatifs. Nous nous sommes intéressés en particulier au cas d'un auditeur en mouvement. Contrairement au système décrit dans le Chapitre 7, on ne va pas traîter la réverbération sous forme d'une somme de convolution avec des filtres courts. Ici, nous disposons de la réponse complète prenant en compte «toutes» les réflexions spéculaires et diffuses au cours du temps. Néanmoins le principe général reste le même. L'auditeur se déplace par pas de temps fixes. Pour chaque



(a) (b)

Figure 9.22 - *Exemple de réponse énergétique et la réponse en pression reconstruite. (a) Réponse énergétique (exprimée en puissance acoustique) reconstruite à partir des données contenues dans les échos. (b) Réponse en pression correspondante. La réponse énergétique obtenue pour les échos «larges» est utilisée pour fenêtrer un bruit blanc.*

nouveau pas de temps, on reprojette la solution stockée sur les surfaces au point d'écoute, on reconstruit la réponse impulsionnelle puis on filtre le signal émis par la source par blocs de durée égale au pas de temps d'animation considéré. Afin d'éliminer les artefacts dûs aux changements (non continus) dans les filtres au cours du temps, on calcule deux signaux, l'un avec le filtre au pas de temps précédant et l'un avec le filtre au pas de temps courant, puis on effectue un fondu enchainé entre les deux blocs d'échantillons obtenus. On notera bien ici que la taille des filtres peut être largement supérieure au pas de temps d'animation considéré. Il faut donc prendre garde à ajouter les «queues» des signaux générés à des pas de temps antérieurs au signal calculé au pas de temps courant.

9.9 Bilan et extensions

Dans ce chapitre nous avons présenté une extension de la technique de radiosité hiérarchique utilisée en simulation de l'éclairage adapté à la simulation de phénomènes temporels. Nous avons défini une nouvelle représentation hiérarchique de l'intensité d'une onde sonore au cours du temps sous la forme d'une liste d'«échos» stockés sur les éléments. Nous avons également proposé la définition de nouveaux paramètres pour décrire les échanges, représentés par des facteurs de forme étendus. Un nouvel oracle de raffinement à été introduit de manière à contrôler la simulation en fonction de la fréquence de l'onde. Pour limiter la complexité temporelle, nous avons décrit un opérateur de fusion des échos, prenant en compte des phénomènes d'interférence. Enfin, nous avons étendu cette approche pour prendre en compte des surfaces idéalement spéculaires avec un modèle de source-images. La solution obtenue nécessite une simple phase de reprojection au point d'écoute lorsque celui-ci se déplace. Cette phase est beaucoup plus rapide que la simulation elle-même. Cette méthode peut donc ouvrir la voie à des applications de «walkthroughs» où la solution est précalculée et reprojetée de manière interactive au point d'écoute. Les premiers résultats obtenus sont prometteurs. Néanmoins, de nombreuses directions peuvent être explorées pour améliorer l'approche.

Tout d'abord le facteur de forme devrait sans doute être étendu pour prendre en compte un terme de phase destiné à prendre en compte des interférences durant le transport de l'énergie. Cela conduirait sans doute à un nouveau critère de raffinement comme nous l'avons déjà évoqué. De même, il serait souhaitable détendre le terme de visibilité de manière à le rendre dépendant de la fréquence. Les approches évoquées dans le Chapitre 6 pourraient être une piste intéressante à explorer pour cette application. Des essais effectués en rajoutant ce terme uniquement pour les liens directs des sources aux récepteurs permettent d'obtenir des résultats plus réalistes.

On peut également penser à rajouter assez simplement des phénomènes de transmission en créant des liens à l'intérieur des parois. Là encore, un nouveau facteur de forme sera sans doute nécessaire, la vitesse de l'onde étant susceptible d'être modifiée.

Les techniques hiérarchiques en radiosité ont récemment été étendues de manière à traiter également des agglomérats d'objets [Sil95b]. Utiliser un tel modèle pour la propagation d'ondes sonores pourrait permettre d'accroître la complexité des scènes.

Enfin, l'utilisation d'une telle approche de manière interactive reste bien entendu un point particulièrement important à développer. Des approches similaires à [DS97] peuvent être envisagées.

Il serait également souhaitable d'envisager une approche à plusieurs passes où les sources images seraient toutes précalculées jusqu'à un certain ordre de réflexion, avant d'être utilisées dans les échanges énergétiques. Nous avons vu, en outre, que des méthodes très rapides permettaient de les mettre à jour dans le cas d'un auditeur mobile. Cela permettrait d'accélérer de façon drastique le processus de traitement des échanges spéculaires puisque les sources-images sont ici recalculées pour chaque écho.

Chapitre 10

Validation en acoustique des salles

Ce n'est pourtant pas ma faute si l'acoustique et moi n'avons pas pu nous entendre. Autant que cela m'a été possible, j'ai fait une cour assidue à cette science bizarre, parente du chien de Jean de Nivelle ; mais, malgré mes assiduités, je ne suis, au bout de quinze ans, guère plus avancé qu'au premier jour. [...] et je suis arrivé après toutes ces études à découvrir ceci : c'est qu'une salle pour être sonore et avoir un timbre agréable devait être longue ou large, être haute ou basse, être en bois ou en pierre, ronde ou carrée, avoir des parois rigides ou capitonnées, [...être] vide ou remplie de monde, sombre ou éclairée [...]

Charles Garnier, *Le nouvel Opéra de Paris,* 1878.

EVALUER la qualité acoustique d'un lieu d'écoute reste l'une des applications principales de l'acoustique virtuelle. Depuis l'époque où Charles Garnier concevait le «Palais Garnier» en proie aux contradictions d'une science de l'acoustique des salles balbutiante, les progrès effectués en psychoacoustique, acoustique géométrique et en auralisation permettent aujourd'hui de pouvoir évaluer objectivement et par l'écoute la qualité d'une salle avant même sa construction. Dans ce chapitre, nous présentons quelques résultats de notre technique de simulation par échanges radiatifs appliqués à l'évaluation des qualités acoustiques d'un lieu d'écoute. En particulier, nous comparerons notre approche à d'autres techniques de simulation géométriques et statistiques.

10.1 Évaluation de l'acoustique d'une salle

Plusieurs difficultés apparaissent d'emblée lorsque l'on veut quantifier la qualité acoustique d'un lieu d'écoute. Quels critères utiliser ? Comment les mesurer ?

10.1.1 Propriétés subjectives

Une première réponse au problème consiste à en considérer l'aspect subjectif. Là encore un obstacle apparait : trouver les termes subjectifs permettant de décrire la sensation d'écoute. Un vocabulaire commun à été mis au point [Ber96, Bar71, BM81, And85] pour permettre aux musiciens et acousticiens d'échanger leurs opinions sur la qualité des salles. Ainsi les principales notions utilisées sont la *réverbérance* (la salle est qualifiée de «vivante» ou de «sèche» suivant la longueur de la réverbération), la «sensation d'espace» (lié à la direction d'incidence des réflexions) et la «clarté» (qui évalue la capacité à distinguer les divers sons les uns des autres). Différentes études ont permis d'isoler des critères «orthogonaux» (statistiquement différents) à partir d'études sur un échantillon de sujets. Ainsi, Ando [And85] retient quatre critères : la «force», l'«intimité», la «réverbérance» et la différence entre le message sonore parvenant aux deux oreilles.

10.1.2 Étude de la réponse impulsionnelle et critères objectifs

A la suite de ces études psychoacoustiques, des critères «objectifs», fondés sur l'étude de la répartition de l'énergie dans la réponse impulsionnelle ont pu être dégagés [Ber96, And85, Bar71, BM81]. La difficulté est bien évidemment de savoir si ces critères sont orthogonaux et permettent de décrire complètement la qualité acoustique d'un lieu découte. On peut séparer les critères en trois familles :

– les critères de décroissance de l'énergie : *temps de réverbération (RT*[1]*)*, *temps de décroissance précoce (EDT*[2]*)*. Historiquement, sans doute les critères principaux de description de l'acoustique d'une salle. Ils mesurent la durée entre l'arrêt de l'émission d'un son dans la salle et le temps où celui-ci devient inaudible ;

– les critères de «clarté». Ces critères sont généralement fondé sur le rapport entre l'énergie précoce dans la réponse impulsionnelle et l'énergie tardive. L'idée directrice est de mesurer le rapport entre une énergie précoce «utile» et une énergie tardive considérée comme «nuisible» à la bonne perception du message sonore. Les principaux critères sont la *clarté (C80)*, la *définition (D50)*, le *temps central (Tc)* ;

– Enfin, la dernière catégorie apparue plus récemment est celle des critères *binauraux*. Ils sont liés à notre perception «stéréophonique». Ce sont plus particulièrement l' *efficacité latérale (LE*[3]*)* et le *coefficient de corrélation inter-aural (IACC*[4]*)*. Ils permettent d'évaluer la «sensation d'espace» et l'«enveloppement» ressentis par l'auditeur.

D'autres critères comme le *niveau de graves (BR*[5]*)* ou le *temps de séparation initial (ITDG*[6]*)* et la *force sonore (G)* sont reliés respectivement aux notions subjectives de «chaleur» et «intimité» du lieu. On trouvera des détails supplémentaires sur ces différents critères et la manière de les calculer en Annexe A.6.

10.2 Un cas simple

Dans un premier temps, nous avons choisi d'évaluer la méthode dans le cas simple d'un volume parallélépipédique. La géométrie du problème est donnée en Figure 10.1. Toutes les surfaces sont absorbantes, la proportion d'intensité absorbée étant fixée à 10%. Le plafond est également diffusant : 30% de l'intensité réfléchie est diffusée. Ces propriétés de matériaux sont choisies comme indépendantes de la fréquence et la source et le récepteurs sont supposés parfaits (ils ont une réponse plate en fréquence) et omnidirectionnels.

1. Reverberation Time.
2. Early Decay Time.
3. Lateral Efficiency.
4. Inter-Aural Cross Correlation.
5. Bass Ratio.
6. Initial-Time-Delay Gap.

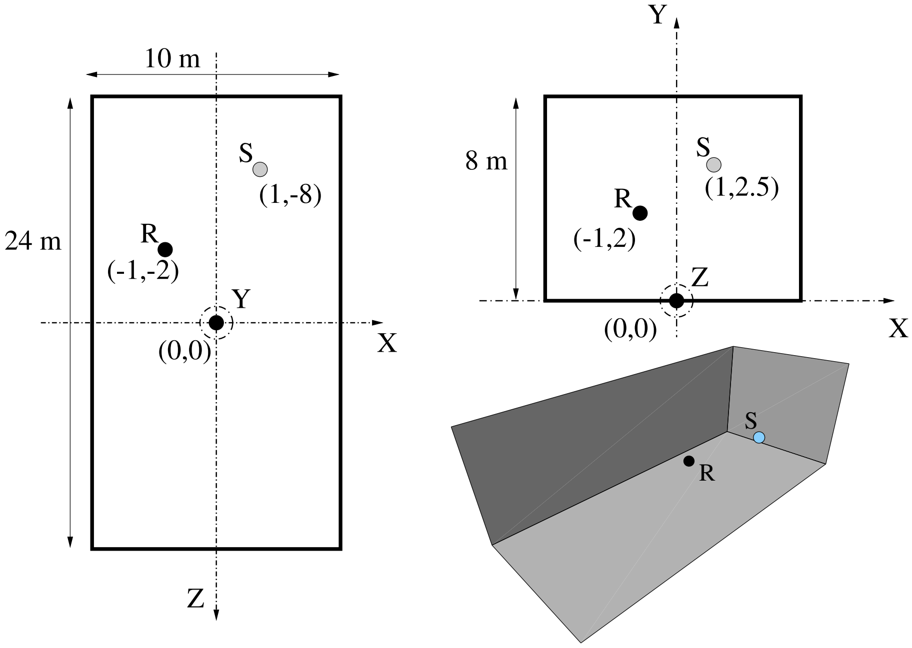


Figure 10.1 - *Configuration utilisée pour les tests comparatifs. La pièce est un parallélépipède dont toutes les parois sont des réflecteurs spéculaires. Le plafond possède également une composante diffuse.*

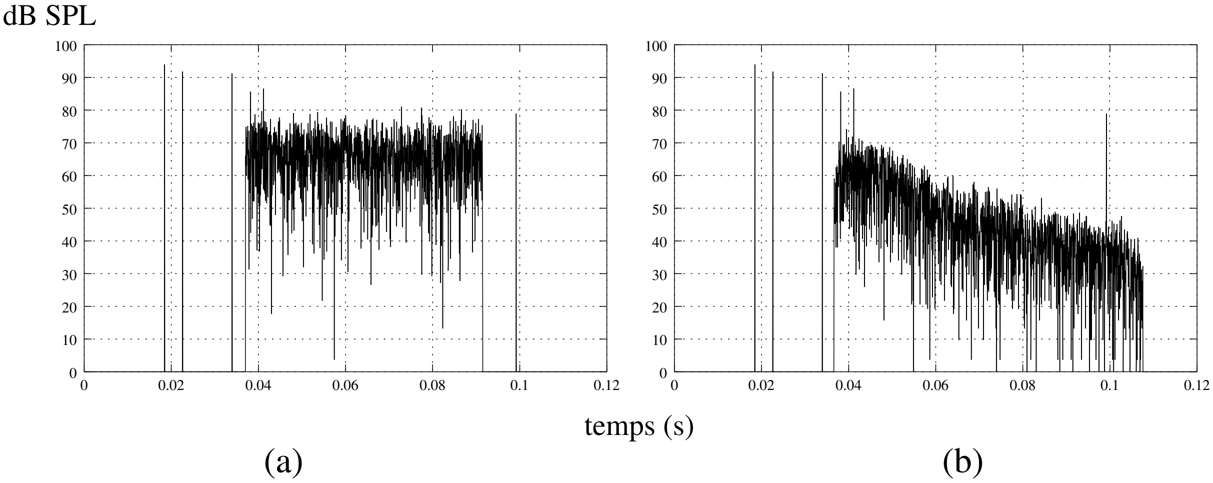


Figure 10.2 - *Comparaison de deux réponses obtenues avec et sans raffinement. (a) Pas de raffinement (b) Les liens ont tous été établis au plus bas niveau. La taille des éléments de surface est environ 30 cm*^{2}.

10.2.1 Comparaison avec une approche hybride lancer de cônes/statistique

Nous avons effectué des comparaisons avec une approche hybride géométrique et statistique développée par le CSTB dans le cadre du logiciel *Ebinaure* [MvMV93]. Cette approche est basée sur un modèle de sources-images pour les réflexions spéculaires précoces et une passe statistique pour les réflexions diffuses et la réflexions spéculaires tardives. Le calcul des sources-images est effectué à l'aide d'un lancer de cônes. La figure 10.3 montre un exemple comparatif des deux méthodes dans le cas où l'on ne prend pas en compte la diffusion *i.e.* on ne considère que les réflexions diffuses. Dans ce cas, les deux approches sont, bien entendu, équivalentes.

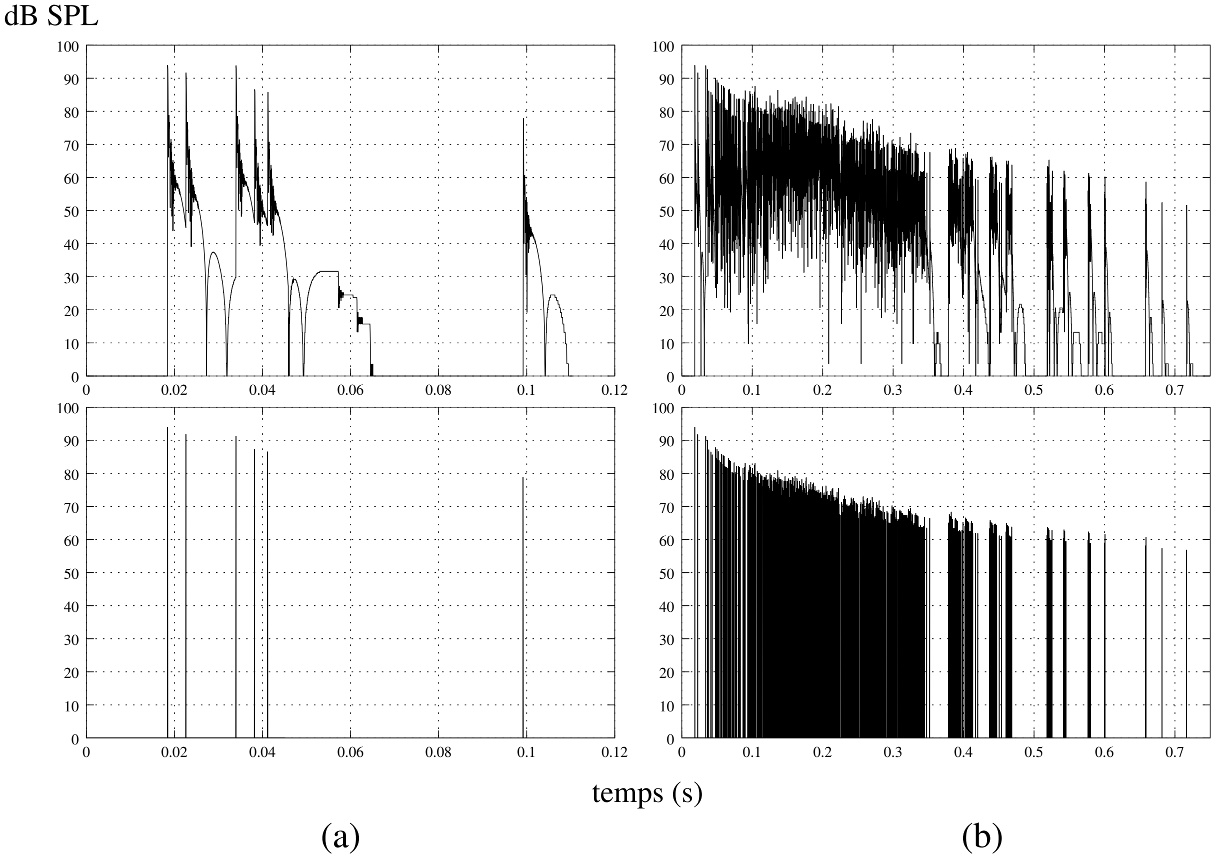


Figure 10.3 - *Comparaison des deux approches en omettant la diffusion. Les figures du haut correspondent au lancer de cônes et celles du bas à la radiosité temporelle hiérarchique avec sources-images. (a) Un ordre de réflexion. (b) Dix ordres de réflexion. Comme on pouvait s'y attendre dans ce cas les deux réponses sont pratiquement identiques. La différence provient ici du fait que nous n'avons pas filtré la réponse obtenu par notre approche avec les filtres passe-bande. Les niveaux de pression et la répartition temporelle des réflexions sont cohérents.*

La Figure 10.4 montre des résultats pour une simulation identique mais avec traitement des réflexion diffuses. Cet exemple est beaucoup plus interessant puisque les deux techniques de simulations sont ici très différentes. Dans le cas d'*Ebinaure*, la fenêtre temporelle utilisée pour générer le bruit blanc est déterminée en fonction de l'ordre de réflexion par des paramètres statistiques qui ne dépendent pas de la géométrie. Ainsi, comme on peut le constater, des contributions dues aux réflexions diffuses sont générées dès le temps d'arrivée du son direct, ce qui n'est pas cohérent avec un modèle de réflexion diffuse géométrique. En effet, on ne peut pas obtenir de contribution diffuse avant qu'il y ait eu une réflexion. De même, comme on peut le voir dans la Figure 10.4 (a), la fenêtre temporelle s'étend plus loin que le temps maximum possible. Cela peut conduire à une sur-estimation de l'énergie au début de la réponse impulsionnelle qui peut influer sur des critères de type clarté, défintion,... Notre approche, restant basée sur la géométrie, permet de générer une réponse cohérente avec le modèle choisi. Comme on peut le constater, les contributions dues aux réflexions diffuses n'interviennent qu'au moment des réflexions et sont fenêtrées de manière plus précise.

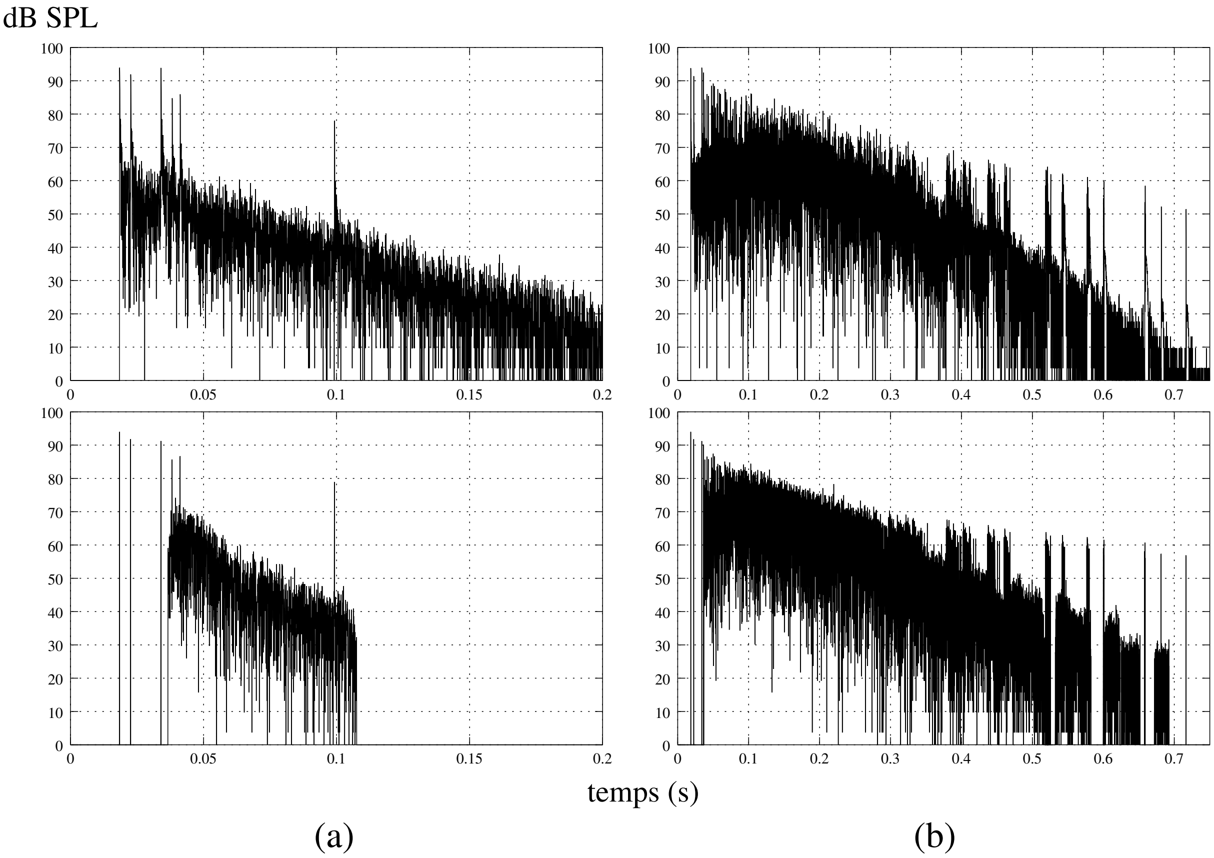


Figure 10.4 - *Comparaison des deux approches avec diffusion. Les figures du haut correspondent au lancer de cônes et celles du bas à la radiosité temporelle hiérarchique avec sources-images. (a) Un ordre de réflexion. (b) Dix ordres de réflexion. Les niveaux de pression restent cohérents mais la distribution temporelle des réflexions diffuses est différente.*

Nous avons également évalué différents critères acoustiques par les deux approches. Parmi les plus importants, les temps de réverbération, TR et EDT, permettent dévaluer la décroissance de l'énergie. Des différences notables sont apparues entre les deux approches puisque l'approche hybride lancer de cônes/statistique donne un TR de l'ordre de 2.9 s. et un EDT de 2.6 s. (pour 140 ordres de réflexion) dans le cas spéculaire pur. Avec la diffusion, le TR reste stable mais l'EDT décroit à 2.4 s. Notre approche prédit un EDT de l'ordre de 1.2 s. et un TR de l'ordre de 1.7 s., ce qui parait trop faible. Cela peut être en partie du au fait que ces critères ont été évalués après 40 ordres de réflections seulement (le temps de calcul étant déjà de 48 h!). Des différences notables ont également été constatées dans l'évaluation des critères de clarté. On enregistre un C80 de l'ordre de -1 dB dans les deux cas spéculaire et diffu avec le lancer de cônes. Avec notre approche, le C80 est plutôt de l'ordre de 2 dB.

10.2.2 Comparaison avec une approche de lancer de rayons de Monte-Carlo

Nous avons également effectué une comparaison avec une approche de type tracé de chemins de Monte-Carlo [Hod91]. Les résultats peuvent être trouvés en Figure 10.5. On remarque une étendue beaucoup plus restreinte des réflections diffuses, qui demeure inexpliquée.

Par contre, les critères évalués dans ce cas sont plus élevés que ceux obtenus par l'approche par lancer

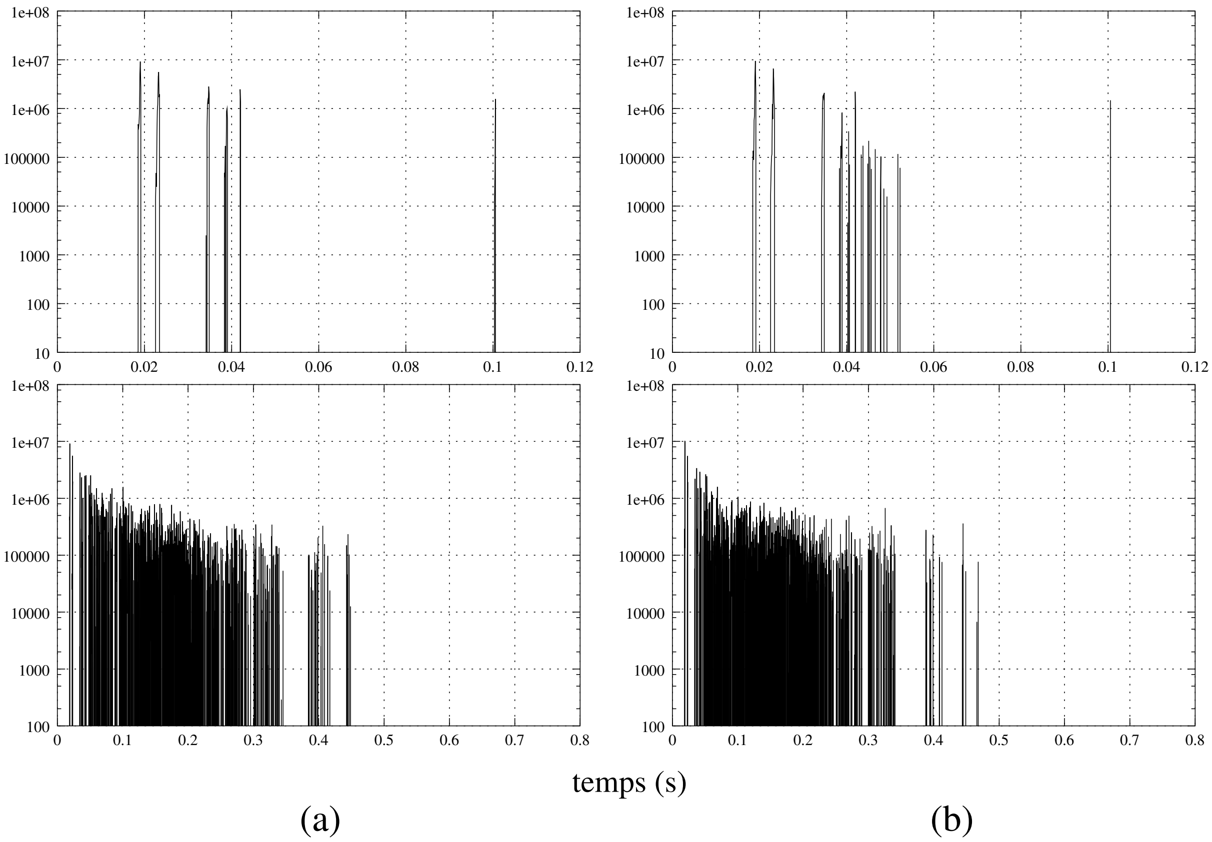


Figure 10.5 - *Comparaison avec un tracé de chemin de Monte-Carlo. (a) Un ordre de réflexion. (b) Dix ordres de réflexion.*

de cônes. Le TR étant de l'ordre de 3.8 s. et l'EDT de l'ordre de 2.85 s. (pour 140 ordres de réflection) dans le cas spéculaire. Dans le cas diffus, on constate une diminution des temps de réverbération avec un TR à 3.39 s. et un EDT à 2.71 s. La clarté est moins élevée dans les deux cas avec un C80 à -2.2 dB.

10.3 Bilan

Dans ce chapitre, nous avons présenté une évaluation de notre approche dans le cadre de l'évaluation de la qualité acoustique d'un lieu découte. Les premiers résultats sont satisfaisants. En particulier, notre méthode permet un fenêtrage temporel précis des réflexions diffuses. Des différences notables lors de l'évaluation des critères acoustiques suggèrent de poursuivre les expériences afin d'assurer la validation de l'approche. En particulier, des tests sur des salles existantes paraissent essentiels au vu des différences obtenues avec les différents approches. Un problème persiste néanmoins puisque la méthode reste néanmoins trop couteuse pour évaluer la réverbération tardive (il n'est guère envisageable de dépasser les 40 premiers ordres de réflexions dans un environnement simple de type boite parallélépipédique). Un moyen de résoudre ce problème serait sans doute d'utiliser une approche statistique [MOD96, MvMV93] pour calculer la réverbération tardive. Les données calculées sur chacune des surfaces pourraient sans doute être utilisées pour contrôler un processus statistique de manière plus fine en maintenant un lien avec la géométrie de l'environnement. Enfin, un point intéressant pourrait être d'utiliser notre approche dans un processus d'optimisation inverse [MOD98].

Conclusion

Les travaux effectuées dans le cadre de cette thèse apportent trois solutions originales à une partie des problèmes posés par la simulation acoustique dans des environnements virtuels.

La première concerne la modélisation efficace de la diffraction des ondes sonores par des obstacles entre un point source et un point de réception. Nous avons proposé un modèle qualitatif du phénomène d'occultation des ondes sonores à l'aide d'un terme de visibilité géométrique, prenant des valeurs réelles entre 0 et 1. Cette valeur est directement liée à l'occultation de régions de l'espace particulières qui sont dépendantes de la fréquence de l'onde sonore : les premiers ellipsoïdes de Fresnel. En conséquence, ce modèle permet de reproduire de manière réaliste les effets d'atténuation dûs aux obstacles, plus forts à haute fréquence.

De plus, nous avons montré qu'il était possible d'utiliser les capacités des cartes de rendu graphique câblé afin de réaliser les calculs nécessaires de manière souple et rapide. Cette approche ouvre la porte au traitement, jusqu'alors ignoré, des phénomènes complexes d'occultation sonore dans des applications interactives.

Nous avons également montré que l'on pouvait utiliser plus finement les capacités des processeurs graphiques en calculant en temps-réel une «carte de profondeur» des obstacles entre la source et le récepteur. Nous avons utilisé cette information pour calculer le champ sonore diffracté donné par la théorie de Fresnel-Kirchhoff. Cette théorie semi-quantitative de la diffraction, dérivée de l'optique, est fondée sur l'expression de la contribution d'un élément différentiel de surface du front d'onde au champ total reçu par le récepteur. La carte de profondeur nous permet à la fois de déterminer quelle partie du front d'onde sera occultée ainsi que les paramètres nécessaires au calcul. Des tests comparatifs, dans des cas simples, ont montré un très bon accord avec une méthode d'éléments finis de frontière.

La méthode reste limitée puisqu'elle ne considère que la partie des obstacles visible depuis la source. Elle peut permettre toutefois d'obtenir une première approximation dans des cas complexes où une simulation plus précise serait très coûteuse. Elle demande un traitement plus coûteux que la solution qualitative précédente mais on peut encore accélérer le processus. Le temps de calcul (pour une fréquence donnée) varie entre le centième de seconde et la minute suivant la précision choisie. L'utilisation de moteurs de rendu graphique câblé le rend toutefois peu dépendant de la complexité géométrique de l'environnement.

Nous avons présenté ensuite un système de rendu sonore permettant un rendu synchronisé du son dans un environnement d'animation 3D interactif. Cette approche est basée sur un concept de «prise de son virtuelle», où plusieurs sources et microphones peuvent être placés dans l'environnement virtuel. Les sources, à l'origine ponctuelles, peuvent être assemblées de manière à simuler des sources complexes. Microphones et sources sont également décrits par leurs réponses impulsionnelle directionnelle. Des dispositifs d'émission et de prise de son réalistes peuvent être simulés (des oreilles humaines par exemple). Les signaux captés par les microphones sont ensuite mixés en fonction du type de restitution choisi et de l'effet voulu par l'utilisateur. Le rendu du son utilise le concept de chemin sonore entre une source et un récepteur. Ces sources peuvent être également des sources-images (symétriques de la source principale par rapport au plan de la réflexion) permettant de traiter de manière souple des réflexions spéculaires idéales. Ces réflexions sont responsables des «échos» et de la réverbération du son. Pour chacun des chemins, une réponse impulsionnelle courte associée est calculée et utilisée pour filtrer une texture sonore. Enfin, ce système utilise directement l'approche décrite précédemment pour le calcul des occultations sonores. Nous avons montré à cette occasion qu'elle peut être très facilement étendue aux chemins réfléchis, les opérations géométriques supplémentaires étant, là encore, prises en charge par le processeur graphique.

Les applications principales d'un tel système sont la génération de bande son pour des films de synthèse et la réalité virtuelle. Le coût des calculs limite encore cette dernière application à des environnements simples. On notera cependant que le traitement du signal, effectué de manière logicielle, monopolise une large part des ressources en temps de calcul. Ce système est actuellement en cours d'intégration pour simuler un dispositif de prise de son/restitution dans le cadre d'une application de téléconférence.

L'utilisation de sources-images limite également la simulation à de faibles ordres de réflexions. On peut néanmoins songer à l'étendre, dans un premier temps, en utilisant des approches de type «réverbération artificielle». Les réflexions tardives seraient calculée par filtrage du signal avec un bruit coloré de décroissance exponentielle variable en fonction de la fréquence.

Enfin la dernière contribution de ce travail de thèse est une méthode originale de simulation par échanges radiatifs hiérarchiques. Cette méthode est fondée sur les approches multi-échelles de plus en plus utilisées en synthèse d'image pour accélérer les temps de calculs des simulations de l'éclairage. Nous avons, dans cet esprit, proposé une adaptation de la technique de radiosité hiérarchique, de manière à pouvoir prendre en compte l'aspect temporel de la propagation du son. Nous avons également présenté une solution originale pour contrôler la complexité du processus dans le domaine temporel tout en permettant de traiter les interférences. Enfin, nous avons étendu cette approche développée dans le cadre de surfaces lambertiennes, pour pouvoir utiliser des miroirs acoustiques, puisque la réflexion du son comporte généralement une forte composante spéculaire. Notre approche permet de simuler des phénomènes complexes et de les visualiser, dans des temps raisonnables mais non interactifs. En outre, la solution obtenue peut être facilement reprojetée en n'importe quel point de réception dans l'environnement. Cela ouvre la voie à des applications de «walkthrough», dans lesquelles la solution est tout d'abord précalculée sur les surfaces puis «reprojetée» de manière interactive lorsque le point d'écoute se déplace.

Des techniques de radiosité hiérarchique permettant des mises à jour incrémentales en cas de modification de l'environnement existent déjà en synthèse d'image. On pourrait dans l'avenir songer à appliquer les solutions proposées à notre problème. On pourrait également pousser plus avant l'aspect hiérarchique en utilisant des regroupements de surfaces, comme les approches récentes basées sur le «clustering» en radiosité. De plus, notre approche demande encore à être validée plus finement d'un point de vue tant quantitatif que perceptif.

Enfin, la fusion des trois approches décrites dans cette thèse au sein d'un système unique permettrait d'obtenir un simulateur acoustique de qualité capable de traiter des environnements dynamiques généraux et de produire une bande-son cohérente avec une simulation visuelle associée. Il est aussi envisageable d'utiliser nos techniques de diffraction et de propagation temporelle hiérarchique pour d'autres types d'ondes comme les ondes radioélectriques. Des applications s'ouvriraient alors dans le cadre de la téléphonie mobile ou des réseaux sans fil qui sont également des domaines en intense développement.

Nous avons montré dans cette thèse qu'il est possible d'accélérer les calculs géométriques nécessaires aux simulation acoustiques. Toutefois, la majeure partie du temps de calcul nécessaire à l'auralisation d'un signal sonore est consacrée aux seules opérations de convolution et de traitement du signal. Bien entendu, il existe des cartes spécialisées pour ces taches permettant d'effectuer des convolutions en temps réel. Mais elles restent encore réservées à un marché professionnel. A l'heure actuelle, elles n'offrent pas le contrôle haut niveau que l'on pourrait en attendre. De plus, leur capacités paraissent presque dérisoires en comparaison de celles de cartes graphiques déjà disponibles sur le marché grand public. On ne peut qu'espérer que le développement de l'acoustique virtuelle invitera les constructeurs à développer de nouveaux systèmes. Enfin, un manque crucial a combler, probablement avant même le développement de cartes spécialisées, est celui d'une «librairie» permettant la description et le rendu d'une scène virtuelle auditive. Cette librairie pourrait être un homologue d'«*OpenGL*», standard du graphique 3D. Elle définirait des fonctions de description des données acoustiques et géométriques associées à la scène virtuelle auditive. Elle permettrait également de réaliser à très haut niveau le rendu d'un chemin sonore ou le calcul de sources-images par exemple. Les fonctions pourraient être ensuite implémentées de manière câblée sur des cartes spécialisées. Ces dernières pourraient même partager le traitement des opérations géométriques avec un moteur de rendu graphique associé. Cette approche est sans doute bien plus prometteuse que de chercher à intégrer quelques fonctions audio simplistes dans des librairies graphiques, pour les rendre multi-média. En particulier, lorsque l'on sait que la puissance de calcul qui serait nécessaire à un bon traitement audio n'est généralement pas au rendez-vous.

Comme l'optique et l'acoustique physique partagent les mêmes fondements, l'acoustique virtuelle et la synthèse d'image utilisent des outils informatiques communs. S'il parait peu probable d'appliquer directement les approches développées pour la simulation lumineuse à la synthèse du son, il est clair que bon nombre d'idées peuvent être ré-utilisées d'un domaine à l'autre. Cet aspect peut être profitable aux deux parties. On rappellera, pour conclure, que des formulations de type échanges radiatifs ont été utilisées, en acoustique, dès 1971, pour n'apparaître en synthèse d'image qu'en 1984, soit plus de dix ans après[7]...

7. le fait que la synthèse d'images soit une science récente n'explique pas tout ; les premières conférences internationales de synthèse d'images ont eu lieu dès le début des années 70...

Cinquième partie

Annexes

Annexe A

Quelques chiffres et formules

DANS cette annexe, nous présentons quelques détails supplémentaires et données numériques diverses concernant en particulier les relations entre pressions, intensités et puissances acoustique et les différents coefficients liés à la propagation des ondes sonores. On trouvera également explicité le calcul de la contribution d'un élément différentiel de surface du front d'onde pour la théorie de la diffraction de Fresnel-Kirchhoff, des valeurs numériques concernant les ellipsoïdes de Fresnel, ainsi que des informations supplémentaire sur le calcul des critères objectifs de qualité d'un lieu d'écoute.

A.1 Niveaux, pressions, intensités et puissances acoustiques

Pour une source ponctuelle harmonique, générant une onde sphérique, la relation suivante lie la puissance acoustique P à l'intensité I dans la direction de déplacement de l'onde sonore :

$$I = \frac{P}{4\pi r^2}, \tag{A.1}$$

où I est exprimée en $W.m^{-2}$, P en W et r est la distance à la source en m.

Dans le cas d'une onde plane harmonique, on a également la relation suivante[1] :

$$I = \overline{pv} = \frac{\overline{p^2}}{\rho_0 c}, \tag{A.2}$$

où p est la pression en $N.m^{-2}$ (ou Pa), v la vitesse particulaire en $m.s^{-1}$, $\bar{\star}$ dénote la moyenne temporelle de la variable $\star$ et $\rho_0 c$ est une constante, appelée *l'impédance caractéristique* du milieu. Pour l'air aux conditions normales on a $\rho_0 c = 414\ kg.m^{-2}.s^{-1}$.

Étant donné que nous sommes capables de percevoir environ 13 ordres de magnitude en intensité acoustique, on préfère généralement utiliser un gain logarithmique pour décrire les grandeurs acoustiques. Cette unité de gain est le décibel (dB).

On pourra donc exprimer :

– les intensités acoustiques :

$$G_{IL} = 10 \log_{10}\left(\frac{I}{I_0}\right),$$

1. Cette relation est également valide en champ lointain ($r \gg \lambda$) d'une source ponctuelle (onde sphérique). En pratique c'est cette expression qui sera utilisée dans les applications qui nous intéressent ici.

où I et I_0 sont exprimée en $W.m^{-2}$ et G est l'intensité exprimée en $dB\ IL$[2]. $I_0 = 1.10^{-12}\ W.m^{-2}$ est une intensité de référence correspondant au seuil d'audition.

– les pressions acoustiques :

$$G_{SPL} = 20\log_{10}\left(\frac{\tilde{p}}{\tilde{p}_0}\right),$$

où $\tilde{p}$ et $\tilde{p}_0$ sont exprimées en $N.m^{-2}$ et G est la pression exprimée en $dB\ SPL$[3]. $\tilde{p}_0 = 2.10^{-5}\ N.m^{-2}$ est une pression de référence correspondant au seuil d'audition à 1000 Hz. Ici, $\tilde{p}$ et $\tilde{p}_0$ sont des *pressions quadratiques moyennes* i.e. la racine carrée de la moyenne temporelle du carré de la pression.

Pour une onde plane les deux écritures sont équivalentes. Dans l'hypothèse d'une onde sphérique et pour un niveau à 1 m, les deux expressions sont également équivalentes. D'un manière plus générale on a, à une distance r de la source :

$$G_{SPL} = G_{IL} - 20\log_{10}(r)$$

– les puissances acoustiques :

$$G_{PL} = 10\log_{10}\left(\frac{P}{P_0}\right),$$

où P et P_0 sont exprimée en $W.m^{-2}$ et G est la puissance exprimée en $dB\ PL$[4]. $P_0 = 1.10^{-12}\ W$ est choisi comme référence.

Dans le cas d'une onde sphérique, il n'y pas équivalence directe à 1 m et on a la relation suivante à la distance r de la source :

$$G_{PL} = G_{SPL} + 20\log_{10}(r) + \underbrace{10\log_{10}\left(\frac{4\pi\tilde{p}_0{}^2}{\rho_0 c P_0}\right)}_{\approx 11}$$

A.2 Propagation du son

A.2.1 Vitesse de propagation

proportionnelle a la densité On suppose la vitesse du son, c, constante par rapport au temps et à l'espace.

$$c = 331.4 + 0.6\ \theta\ \ m.s^{-1}, \tag{A.3}$$

où θ est le température en $°C$ [Kut91b].

2. Intensity Level
3. Sound Pressure Level
4. Power Level

A.2.2 Atténuation atmosphérique

La propagation d'un signal harmonique dans l'air sur une distance r s'accompagne d'une décroissance exponentielle de l'amplitude de la pression acoustique p à partir de sa valeur initiale p_i. Cette décroissance est due aux phénomènes d'absorption atmosphérique et s'exprime pour des ondes planes en champ libre par :

$$p = 1.10^{-(0.1151\ r\ \mu^{\lambda}_{T,pa,h})} p_i = \mu^{\lambda}(r) p_i. \tag{A.4}$$

On a donc pour l'intensité acoustique :

$$I = (\mu^{\lambda}(r))^2 I_i. \tag{A.5}$$

Le coefficient d'atténuation atmosphérique *par mètre* $\mu^{\lambda} = 1.10^{-(0.1151\ \mu^{\lambda}_{T,pa,h})}$, exprimé en $dB\ SPL/m$, peut être calculée en utilisant les formules de la norme ISO 9613 [ISO93] :

$$\begin{aligned} \mu^{\lambda}_{T,p_a,h} = \quad & 8.686\, f^2 \left\{ \left[1.84.10^{-11} \left(\frac{p_a}{p_r}\right)^{-1} \left(\frac{T}{T_0}\right)^{1/2} \right] \right. \\ & \left. + \left(\frac{T}{T_0}\right)^{5/2} \left(0.01275\, e^{\frac{-2239.1}{T}} \left[f_{r_0} + \left(\frac{f^2}{f_{r_O}}\right) \right]^{-1} + \ 0.1068\, e^{\frac{-3352}{T}} \left[f_{r_N} + \left(\frac{f^2}{f_{r_N}}\right) \right]^{-1} \right) \right\} \end{aligned}$$

où :

$f_{r_O} = \frac{p_a}{p_r}\left(24\ +\ 4.04.10^4 h \frac{0.02+h}{0.391+h}\right)$	est la fréquence de relaxation de l'oxygène
$f_{r_N} = \frac{p_a}{p_r}\left(\frac{T}{T_0}\right)^{1/2} \left(9 + 280h\, e^{-4.170\left[\left(\frac{T}{T_0}\right)^{1/3} - 1\right]}\right)$	est la fréquence de relaxation de l'azote,
$f = \lambda/c$ (Hz)	est la fréquence de l'onde sonore,
p_a (kPa)	est la pression atmosphérique,
$p_r = 101.325$ kPa	est la pression atmosphérique de référence,
T (K)	est la température atmosphérique,
$T_0 = 293.15$ K	est la température atmosphérique de référence
et h (%)	est la fraction molaire de vapeur d'eau.

La fraction molaire de vapeur d'eau en pourcentage, h, peut se calculer à partir de l'humidité relative h_r, de la pression p_a et de la température T :

$$h = h_r \left(\frac{p_{sat}}{p_r}\right) \left(\frac{p_a}{p_r}\right),$$

où $p_{sat} = p_r.10^{-\ 6.8346\ (T_{01}/T)^{1.261} +\ 4.6151}$ est la pression de vapeur saturante et $T_{01} = 273.16$ K est la température isotherme au point triple.

Les valeurs de l'atténuation atmosphérique pour des bandes d'octaves sont déduites de celles pour les sons purs (harmoniques) appliquées aux fréquences centrales des bandes de fréquence.
Pour information, on donne quelques valeurs du coefficient d'atténuation pour une fréquence de 1000 Hz et une pression atmosphérique de 101.325 kPa :

- pour 20° C et 50% d'humidité : $\mu^{\lambda}_{T,p_a,h} = 0.004665$ et $(\mu^{\lambda})^2 = 0.99753$.

– pour 25° C et 50% d'humidité : $\mu^{\lambda}_{T,p_a,h} = 0.005677$ et $(\mu^{\lambda})^2 = 0.99699$.

– pour 10° C et 30% d'humidité : $\mu^{\lambda}_{T,p_a,h} = 0.006769$ et $(\mu^{\lambda})^2 = 0.99642$.

A.2.3 Coefficients d'absorption et de réflexion

Dans cette section, nous récapitulons quelques détails relatifs aux coefficients d'absorption par les parois et l'audience. Pour plus de détails à ce propos, nous suggérons de se reporter aux références [Kut91b, Ber96]. En particulier, on trouvera dans ces ouvrages de nombreuses valeurs numériques de coefficients d'absorption pour différents matériaux, les spectateurs,*etc.*

Comme nous l'avons évoqué au paragraphe 1.5.1, une onde sonore atteignat une paroi va être partiellement réfléchie, une partie de son intensité étant absorbée. En acoustique, étant donné les grandeurs relatives des longueurs d'ondes traitées et des aspérités des parois, on utilise généralement un modèle de réflection spéculaire. La grandeur physique utilisée pour décrire le comportement acoustique d'une paroi est son *impédance normale* :

$$Z = \left(\frac{p}{v_n}\right)_{\text{surface}}.$$

Note : on utilise également son *impédance acoustique spécifique*, définie par :

$$\zeta = \frac{Z}{\rho_0 c}.$$

Dans le cas général, l'impédance est *dépendante de la fréquence et de la direction d'incidence de l'onde*. Toutefois, on se place généralement dans le cas particulier de parois «rigides», dites *à réaction localisée*. Dans ce cas, la paroi elle-même (ou l'espace derrière elle) ne peut propager d'ondes ou de vibrations dans une direction parallèle à sa surface. L'impédance est alors indépendante de la direction d'incidence. Comme nous allons le voir, ce n'est pas pour autant le cas du coefficient d'absorption.

Absorption en incidence normale et oblique

Lors de la réflection, la pression réflechie peut s'exprimer comme le produit de la pression incidente et d'un coefficient de réflection :

$$\hat{R}^{\lambda}(\theta,\phi) = R^{\lambda}(\theta,\phi)e^{i\chi^{\lambda}(\theta,\phi)}.$$

Une fraction $1 - |\hat{R}^{\lambda}(\theta,\phi)|^2$ de l'intensité de l'onde est donc perdue lors de la réflection, définissant le coefficient d'absorption de la paroi, $\alpha^{\lambda}(\theta,\phi)$.

Dans le cas où la direction d'incidence est normale à la paroi, on montre que [Kut91b] :

$$R^{\lambda} = \frac{\zeta - 1}{\zeta + 1},$$

d'où :

$$\alpha^{\lambda} = \frac{4\mathrm{Re}(\zeta)}{|\zeta|^2 + 2\mathrm{Re}(\zeta) + 1}.$$

Dans le cas d'une incidence oblique on a :

$$R^{\lambda}(\Theta) = \frac{\zeta\cos\Theta - 1}{\zeta\cos\Theta + 1},$$

où $\cos\Theta = \mathbf{v}.\mathbf{n}$, Θ est l'angle d'incidence de l'onde par rapport à la normale de la paroi.
On a donc :

$$\alpha^{\lambda}(\Theta) = \frac{4\mathrm{Re}(\zeta)\cos(\Theta)}{|\zeta\cos(\Theta)|^2 + 2\mathrm{Re}(\zeta)\cos(\Theta) + 1}.$$

Coefficient d'absorption en champ diffus

Dans le cas d'un champ diffus, on peut définir un coefficient d'absorption en incidence aléatoire, un grand nombre d'ondes provenant de toutes les directions simultanément. Le coefficient d'absorption, pour des surfaces à réaction localisé est alors donné par la *formule de Paris* [Kut91b] :

$$\alpha_{uni} = \frac{8}{|\zeta|^2}\cos\beta\left[|\zeta| + \frac{\cos 2\beta}{\sin\beta}\arctan\left(\frac{|\zeta|\sin\beta}{1+|\zeta|\cos\beta}\right) - \cos\beta\ln(1+2|\zeta|\cos\beta + |\zeta|^2)\right],$$

avec $\beta = \arctan\left(\frac{\mathrm{Im}\zeta}{\mathrm{Re}\zeta}\right)$.

Pour conclure ce rappel a propos des coefficients d'absorption, on notera que les mesures et la normalisation de tels coefficients est encore extrêmement difficile en acoustique. La plupart des coefficients disponibles sont valables dans le cadre d'une utilisation avec des formules de calcul du temps de réverbération, telles celles d'Eyring ou Sabine. Ces coefficients sont-ils directement applicables au sein d'une approche géométrique où «chaque» réflexion est considérée et non pas toutes les réflexions dans leur ensemble. Le même problème se pose dans une moindre mesure pour les coefficients de diffusion. Il n'est pas évident de savoir si l'on doit utiliser les mêmes coefficients au sein d'approches géomètriques de la diffusion (comme la rdiosité ou le lancer de rayon de Monte-Carlo) et d'approches statistiques.

A.3 Ellipsoïdes et zones de Fresnel

Nous donnons ici quelques détails supplémentaires relatifs aux ellipsoïdes et zones de Fresnel (*cf.* paragraphe 6.2)

On rappelle que les ellipsoïdes de Fresnel dont les foyers sont une source (S) et un récepteur (R) sont définis par l'ensemble des points $M \in \mathbb{R}^3$ vérifiant l'équation :

$$\|\vec{SM}\| + \|\vec{MR}\| = \|\vec{SR}\| + \frac{k\lambda}{2}, k \in \mathbb{N}^+,$$

où λ est la longueur d'onde de l'onde sonore.

Si on pose $r = \|\vec{SR}\|$, on a donc pour longueur du grand axe b :

$$b = r + \frac{k\lambda}{2}$$

et pour longueur du petit axe a :

$$a = \sqrt{k\lambda\left(r + \frac{k\lambda}{4}\right)}$$

Aire des zones de Fresnel

Cas d'une onde plane

Dans le cas d'une onde plane, les zones de Fresnel sont des couronnes circulaires exceptée la première zone en forme de disque (*cf.* Figure A.1(a)). L'aire de la première zone à distance x de la source est donnée par :

$$S_1(x) = \pi h^2$$

où h est le rayon du disque qui coupe le $k^{\text{ème}}$ ellipsoïde de Fresnel à distance x de la source.

Ce rayon peut être calculé en considérant les triangles SMP et PMR. On a alors :

$$\begin{aligned} &h^2 + x^2 = d^2 \text{ et} \\ &h^2 + (r-x)^2 = (b-d)^2, \end{aligned}$$

où b est le grand axe de l'ellipsoïde.

La solution du système est :

$$\begin{aligned} &d = \tfrac{2rx - r^2 + b^2}{2b} \text{ et} \\ &h = \sqrt{d^2 - x^2}. \end{aligned}$$

L'aire de la $k^{\text{ème}}$ zone à distance x de la source est donné par la différence entre les aires de deux disques :

$$S_k(x) = \pi(h_k^2 - h_{k-1}^2),$$

où h_k est le rayon du disque qui coupe le $k^{\text{ème}}$ ellipsoïde de Fresnel à distance x de la source.

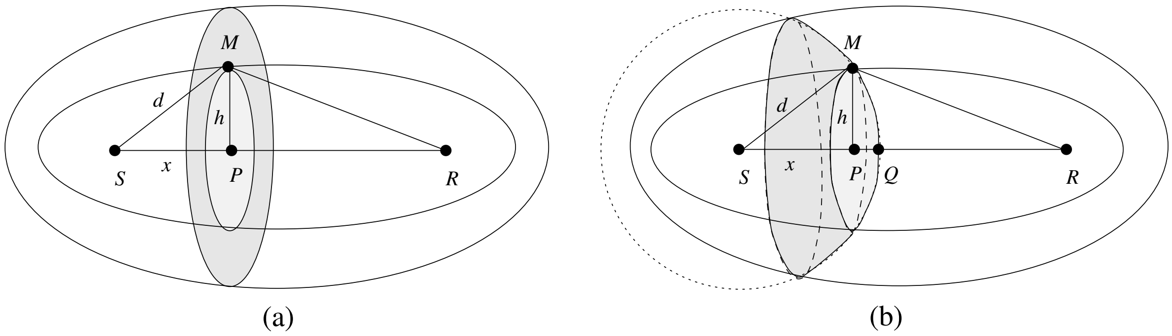


Figure A.1 - *Notations pour le calcul d'aire des zones de Fresnel. (a) Cas d'une onde plane (b) Cas d'une onde sphérique*

Cas d'une onde sphérique

Dans la cas d'une onde sphérique (*cf.* Figure A.1(b)), la première zone est une calotte sphérique d'aire :

$$S_1(d) = \pi((d-x)^2 + 2h^2),$$

où d est le rayon du front d'onde sphérique considérée, x est la distance (projetée sur l'axe) à laquelle la sphère coupe l'ellipsoïde et h est la hauteur de la calotte à la distance x.

Les valeurs de x et h peuvent être calculées comme dans le cas de l'onde plane en considérant les triangles SMP et PMR :

$$x = \tfrac{2bd+r^2-b^2}{2r} \text{ et}$$
$$h = \sqrt{d^2-x^2},$$

L'aire de la $k^{\text{ème}}$ zone à distance d de la source est donc donnée par :

$$\begin{aligned} S_k(d) = &\ \pi(2d(x_{k-1}-x_k)+h_{k-1}^2+h_k^2 \\ = &\ \pi(2d(x_{k-1}-x_k)+d_{k-1}^2-x_{k-1}^2+h_k^2-x_k^2 \\ = &\ \pi(4d^2-(x_k+d)^2-(x_{k-1}-d)^2) \end{aligned}$$

A.4 Théorie de Fresnel-Kirchhoff : contribution d'un élément de surface du front d'onde

Comme nous l'avons déjà vu au paragraphe 3.1.2, on peut résoudre l'équation d'Helmholtz à l'aide du théorème de Green. Dans le cas de l'équation d'Helmholtz homogène en régime continu, on obtient une solution de la forme [Hec87] (*cf.* Figure 3.1 (a)):

$$\hat{P}(M) = \frac{1}{4\pi}\left[\iint_S \frac{e^{ikr}}{r}\nabla\hat{\mathbf{P}}(U)\cdot\mathbf{dS} - \iint_S \hat{P}(U)\nabla\left(\frac{e^{ikr}}{r}\right)\cdot\mathbf{dS}\right], \tag{A.6}$$

où $\mathbf{dS} = \mathbf{n}dS$, avec $\|\mathbf{n}\| = 1$.

Appliquons cette formulation à une source ponctuelle, créant le champ de pression :

$$\hat{P}(t,\rho) = \frac{P_o}{\rho}e^{i(k\rho-\omega t)} = \hat{P}_u(\rho)e^{-i\omega t},$$

où P_o est l'amplitude de la pression émise par la source.

Dans ce cas l'équation A.6 devient:

$$\hat{P}(M) = \frac{1}{4\pi}\iint_S\left[\frac{e^{ikr}}{r}\frac{\partial}{\partial\rho}\left(\frac{P_o e^{ik\rho}}{\rho}\right)(\mathbf{n}\cdot\mathbf{v}) - \frac{P_o e^{ik\rho}}{\rho}\frac{\partial}{\partial r}\left(\frac{e^{ikr}}{r}\right)(\mathbf{n}\cdot\mathbf{r})\right]dS,$$

avec $\|\mathbf{v}\| = \|\mathbf{r}\| = \|\mathbf{n}\| = 1$.

On a également

$$\frac{\partial}{\partial\rho}\left(\frac{e^{ik\rho}}{\rho}\right) = e^{ik\rho}\left(\frac{ik}{\rho}-\frac{1}{\rho^2}\right)$$

d'où on tire

$$\begin{aligned} \hat{P}(M) &= \frac{P_o}{4\pi}\iint_S\left[\frac{e^{ik(\rho+r)}}{r}\left(\frac{ik}{\rho}-\frac{1}{\rho^2}\right)(\mathbf{n}\cdot\mathbf{v}) - \frac{e^{ik(\rho+r)}}{\rho}\left(\frac{ik}{r}-\frac{1}{r^2}\right)(\mathbf{n}\cdot\mathbf{r})\right]dS \\ &= \frac{P_o}{4\pi}\iint_S\left[\frac{e^{ik(\rho+r)}}{\rho r}\left\{\left(ik-\frac{1}{\rho}\right)(\mathbf{n}\cdot\mathbf{v}) - \left(ik-\frac{1}{r}\right)(\mathbf{n}\cdot\mathbf{r})\right\}\right]dS \end{aligned}$$

Si l'on choisit à présent la surface S comme la région doublement connectée comme sur la Figure A.2, l'intégrale porte alors sur $S_1 \cup S_2$. Toutefois on peut faire croître S_2 jusqu'a l'infini et on montre alors que sa contribution à l'intégrale tend alors vers 0 [Hec87]. L'intégrale ne porte plus que sur S_1 qui est une sphère centrée sur S et correspond donc obligatoirement au front d'onde à un instant donné. Il est également clair que dans ce cas on a : $(\mathbf{n} \cdot \mathbf{v}) = -1$ et $(\mathbf{n} \cdot \mathbf{r}) = -(\mathbf{v} \cdot \mathbf{r})$.

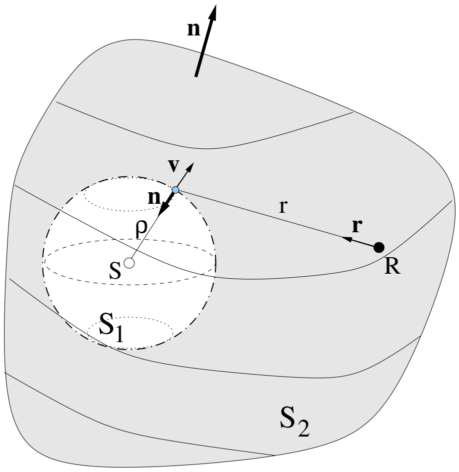


Figure A.2 - *Évaluation du théorème d'Helmholtz-Kirchhoff sur une région doublement connectée entourant le récepteur (d'après [Hec87]). La surface S_1 est une sphère centrée sur la source. La surface S_2 entoure entièrement le récepteur et S_1. Pour* $\rho = 0$ *la pression crée par la source à une singularité et ce point est donc exclu du volume entre S_1 et S_2.*

Finalement on obtient :

$$\hat{P}(M) = -\frac{P_o}{4\pi} \iint_S \left[\frac{e^{ik(\rho+r)}}{\rho r} \left(ik\,(1 - (\mathbf{v} \cdot \mathbf{r})) - \frac{1}{\rho} + \frac{\mathbf{v} \cdot \mathbf{r}}{r} \right) \right] dS$$

La contribution $d\hat{P}(M)$ d'un élément différentiel de surface du front d'onde s'exprime donc comme[5] :

$$d\hat{P}(M) = -\frac{P_o}{4\pi} \frac{e^{ik(\rho+r)}}{\rho r} \left(ik\,(1 - (\mathbf{v} \cdot \mathbf{r})) - \frac{1}{\rho} + \frac{\mathbf{v} \cdot \mathbf{r}}{r} \right)$$

A.5 Convergence de l'évaluation itérative du délai de propagation

On rappelle que l'on peut exprimer le délai de propagation $\tau(t)$ entre une source S et un récepteur R mobiles par la formule récursive suivante :

$$\tau(t) = \frac{1}{c} \|\mathbf{R}(t) - \mathbf{S}(t - \tau(t))\|,$$

qui revient à considérer à l'instant t le récepteur immobile et la source en mouvement. Dans l'hypothèse où *la vitesse de de la source est strictement inférieure à la vitesse du son c*, on montre que l'on peut évaluer le délai par un processus itératif décrit dans le paragraphe 7.4.2, qui converge vers une limite unique.
En effet, soit la suite (U_n) définie par :

$$U_{n+1} = f(U_n), \text{ avec } f \text{ l'application telle que } f(x) = \frac{1}{c} \|\mathbf{R}(t) - \mathbf{S}(t - x)\|, (t,x) \in \mathbb{R}^2.$$

5. dans les paragraphes 3.1.2 et 6.3.1, cette contribution est exprimée avec un vecteur **r** dans le sens opposé.

On a alors $\forall(x,y) \in \mathbb{R}^2$:

$$\begin{aligned} \|\mathbf{R}(t)-\mathbf{S}(t-x)\| &= \|(\mathbf{R}(t)-\mathbf{S}(t-y))+(\mathbf{S}(t-y)-\mathbf{S}(t-x))\| && \text{(A.7)} \\ &\leq \|\mathbf{R}(t)-\mathbf{S}(t-y)\|+\|\mathbf{S}(t-y)-\mathbf{S}(t-x)\|. && \text{(A.8)} \end{aligned}$$

et également :

$$\begin{aligned} -\|\mathbf{R}(t)-\mathbf{S}(t-y)\| &= -\|(\mathbf{R}(t)-\mathbf{S}(t-x))+(\mathbf{S}(t-x)-\mathbf{S}(t-y))\| && \text{(A.9)} \\ &\geq -\|\mathbf{R}(t)-\mathbf{S}(t-x)\|-\|\mathbf{S}(t-x)-\mathbf{S}(t-y)\|. && \text{(A.10)} \end{aligned}$$

On suppose sans perte de généralité que $x \geq y$.
D'après l'équation A.8 on tire :

$$\|\mathbf{R}(t)-\mathbf{S}(t-x)\|-\|\mathbf{R}(t)-\mathbf{S}(t-y)\| \leq \underbrace{\|\mathbf{S}(t-y)-\mathbf{S}(t-x)\|}_{<c\ (x-y) \text{ par hypothèse}} \tag{A.11}$$

D'après l'équation A.10 on tire :

$$\underbrace{-\|\mathbf{S}(t-x)-\mathbf{S}(t-y)\|}_{>-c\ (x-y) \text{ par hypothèse}} \leq \|\mathbf{R}(t)-\mathbf{S}(t-x)\|-\|\mathbf{R}(t)-\mathbf{S}(t-y)\| \tag{A.12}$$

On obtient finalement en combinant A.11 et A.12 :

$$-c\ (x-y) < \|\mathbf{R}(t)-\mathbf{S}(t-x)\|-\|\mathbf{R}(t)-\mathbf{S}(t-y)\| < c\ (x-y),$$

ce qui implique que :

$$\frac{1}{c}\left|\ \|\mathbf{R}(t)-\mathbf{S}(t-x)\|-\|\mathbf{R}(t)-\mathbf{S}(t-y)\|\ \right| < |x-y|$$

On a donc $|f(x)-f(y)| < |x-y|\ \forall(x,y) \in \mathbb{R}^2$. L'application f est donc *contractante* ce qui prouve la convergence de la suite (U_n) et donc du processus de calcul du délai.

A.6 Critères objectifs d'évaluation de la qualité acoustique d'un lieu d'écoute

On rappelle dans cette section, les formules permettant d'évaluer les différents critères les plus courants pour la mesure de la qualité acoustique d'un lieu d'écoute. On rappelle que de plus amples renseignements sur l'évaluation des qualités acoustiques des lieux d'écoute peuvent être trouvés dans [Ber96].

Temps de réverbération (RT) et temps de décroissance précoce (EDT)

Le temps de réverbération en un point du lieu d'écoute est le temps qui s'écoule entre le moment où l'on coupe l'émission d'un signal et celui où le niveau sonore à décru de 60 dB en dessous du niveau initial.

En général ce critère est calculé après intégration de l'énergie dans la réponse impulsionnelle $h(t)$, qui donne sa courbe de décroissance $E(t)$:

$$E(t) = \int_t^{+\infty} h^2(u)du/E(0)$$

On calcule ensuite la pente de la droite de régression linéaire sur les valeurs obtenues entre deux seuils (-5dB et -35dB par exemple).

Le temps de décroissance précoce est le temps que met l'énergie à décroitre de 60dB si l'on considère qu'elle décroit à la même vitesse que pendant les 10 premiers dB.

Définition (D50)

La définition est le *pourcentage* d'énergie arrivant à l'auditeur dans les 50 premières millisecondes sur l'énergie totale. La définition est un indice d'intelligibilité de la parole.

$$D_{50}(\%) = 100\frac{\int_0^{0.05} h^2(t)dt}{\int_0^{+\infty} h^2(t)dt}$$

Clarté (C80)

La clarté est, comme la définition, fondée sur un rapport de lénergie précoce sur l'énergie tardive. Elle s'applique plutôt à un message musical.

$$C_{80}(dB) = 10\log\left(\frac{\int_0^{0.08} h^2(t)dt}{\int_{0.08}^{+\infty} h^2(t)dt}\right)$$

Temps central (Tc)

Il fournit le même type d'information que les deux critères précédants. Il correspond au moment d'ordre 1 de la réponse impulsionnelle.

$$Tc(s) = \frac{\int_0^{+\infty} th^2(t)dt}{\int_0^{+\infty} h^2(t)dt}$$

Temps de séparation initial (ITDG)

Le temps de séparation initial est le délai entre l'arrivée du son direct à l'auditeur et celui de la première réflexion.

Force sonore (G)

La force sonore caractérise l'énergie totale parvenant à l'auditeur.

$$G(dB) = 10\log\left(\frac{\int_0^{+\infty} h^2(t)dt}{\int_0^{+\infty} h^2_{\text{ref}}(t)dt}\right),$$

où h_{ref} est la réponse impulsionnelle mesurée à dix mètres de la même source (omnidirectionnelle) dans un environnement anéchoïque (on ne considère que le son direct). Cet indice permet de caractériser l'homogénéité du niveau sonore dans la salle.

Niveau de graves (BR)

Le niveau de graves est défini comme le rapport du temps de réverbération pour les fréquences basses sur celui des fréquences moyennes :

$$BR = \frac{RT_{125} + RT_{250}}{RT_{500} + RT_{1000}},$$

où les temps de réverbération sont calculés pour les fréquences apparaissant en indices.

Éfficacité latérale (LE)

L'efficacité latérale se rencontre aussi sous la dénomination *fraction d'énergie latérale* ou «Lateral Energy Fraction» en anglais (LF). Elle mesure le rapport entre l'énergie mesurée avec microphone en «figure de 8» [HW94, MR92] (orienté orthogonalement à la direction source-microphone de manière à annuler la réception dans la direction de la source) et celle mesurée avec un microphone omnidirectionnel.

$$LE = \frac{\int_{0.005}^{0.08} h_8^2(t)dt}{\int_0^{0.08} h^2(t)dt}$$

La borne inférieure de 5 ms est introduite pour être sûr d'éliminer le son direct.

Coefficient de corrélation inter-aural (IACC)

Cet indice caractérise la relation entre le son reçu par les deux oreilles d'un auditeur. La valeur de cet indice est comprise entre 0 et 1. Il est minimal en champ diffus.

$$IACC = \max_{|\tau|<0.001} \left| \frac{\int_{-\infty}^{+\infty} h_L(t) h_R(t+\tau)dt}{\sqrt{\int_{-\infty}^{+\infty} h_L^2(t)dt \int_{-\infty}^{+\infty} h_R^2(t)dt}} \right|,$$

où h_L et h_R sont les réponses impulsionnelles mesurées aux entrées des canaux auditifs gauches et droits. Le choix de l'intervalle de variation de τ est lié aux dimensions de la tête.

Annexe B

Calcul des facteurs de forme étendus

CETTE annexe récapitule les définitions des facteurs de forme étendus aux phénomènes temporels que nous utilisons dans notre technique de simulation des échanges diffus. Après une présentation des différents types de facteurs de forme, nous présentons comment on peut les calculer par échantillonnage des surfaces dans le cas où aucune expression analytique n'est disponible (c'est le cas dès que la visibilité entre éléments est prise en compte).

B.1 Les facteurs de forme étendus

Nous rappelons que le facteur de forme «étendu» $\mathcal{F}_{ij}^{\lambda}$ entre deux éléments de surface E_i et E_j sont composés de trois termes :

- un *terme énergétique* qui correspond au facteur de forme classique F_{ij}^{λ} ; il correspond à la *proportion de l'énergie* quittant E_i et arrivant sur E_j *durant la durée de l'échange*,
- le *temps de parcours minimal* T_{ij} qui correspond à la durée minimum du trajet d'un point de l'élément i vers un point de l'élément j,
- la *durée d'échange* τ_{ij} qui caractérise la durée de l'échange radi atif entre les deux éléments.

Si on appelle $t(x,y)$ le temps de propagation d'un point x de E_i au point y de E_j, on a donc les relations suivantes :

$$T_{ij} = \min_{x \in E_i, y \in E_j} t(x,y) \tag{B.1}$$

$$\tau_{ij} = \max_{x \in E_i, y \in E_j} t(x,y) - \min_{x \in E_i, y \in E_j} t(x,y) \tag{B.2}$$

Le terme énergétique F_{ij}^{λ} est calculé de manière similaire au facteur de forme classique auquel on adjoint l'atténuation atmosphérique dépendant de la distance et de la longueur d'onde $\alpha^{\lambda}(r)$ (*cf.* figure B.1) :

$$F_{ij}^{\lambda} = \frac{1}{A_i} \int_{x \in E_i} \int_{y \in E_j} \frac{\cos\theta \cos\theta'}{\pi r^2} V(x,y) \alpha^{\lambda}(r) dx dy, \tag{B.3}$$

où $V(x,y)$ est le terme de visibilité entre les points x et y.

B.2 Cas particuliers

Nous décrivons ici quatre cas particuliers de facteur de forme, faisant intervenir un élément dégénéré ponctuel.

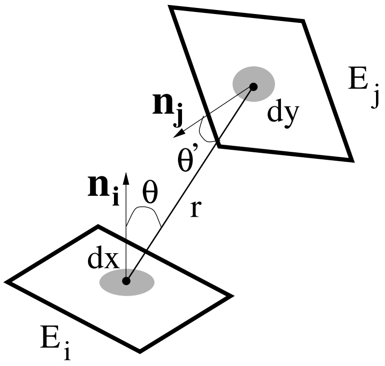


Figure B.1 - *Notations pour le facteur de forme surface/surface*

B.2.1 Facteur de forme source ponctuelle/récepteur ponctuel

Dans ce cas, les deux éléments sont dégénérés. C'est ce facteur de forme qui intervient entre une source sonore et un microphone. Dans ce cas on a :

$$F_{ij}^{\lambda} = \frac{\alpha^{\lambda}(r)V_{ij}}{r^2}, \tag{B.4}$$

où r est la distance de la source au microphone et V_{ij} le terme de visibilité entre les deux éléments.

Les autres termes du facteur de forme étendu restent inchangés et on a : $T_{ij} = \tau_{ij} = r/c$. Dans ce cas, on a également la relation $\mathcal{F}_{ij} = \mathcal{F}_{ji}$.

B.2.2 Facteur de forme entre élément ponctuel et polygone

Ces facteurs de forme interviennent pour les échanges entre source et surfaces ou surfaces et récepteurs. On a alors :

$$F_{ij}^{\lambda} = \frac{1}{A_j}\int_{y\in E_j}\frac{\cos\theta}{4\pi r^2}V_i(y)\alpha^{\lambda}(r)dy, \tag{B.5}$$

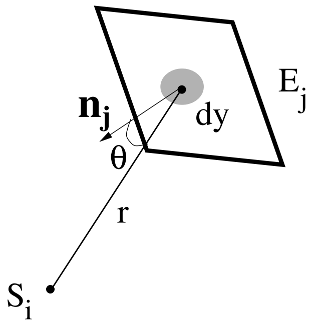


Figure B.2 - *Notations pour le facteur de forme point/surface*

Facteur de forme source image/éléments

Ce cas est une extension du facteur de forme source ponctuelle/polygone ou source ponctuelle/récepteur ponctuel mais la source est ici une source virtuelle, image miroir de la source réelle par rapport au polygone réflecteur P. On rajoute donc un terme de validité $\xi_i^P(y)$ tel que $\xi_i^P(y) = 1$ si le rayon joignant la source image i et le point y intersecte P et 0 sinon. Le terme de visibilité $V^P(x,y)$ doit également être calculé en prenant en compte en compte l'image miroir de la scène par rapport au plan de P. Dans le cas d'un facteur de forme entre source-image et surface, on obtient :

$$F_{ij}^{\lambda} = \frac{1}{A_j} \int_{y \in E_j} \frac{\cos\theta}{4\pi r^2} \xi_i^P(y) V_i^P(y) \alpha^{\lambda}(r) dy, \tag{B.6}$$

En pratique, on utilisera le facteur de visibilité des facteurs de forme liés à l'échange pour limiter le coût des calculs.

B.3 Échantillonner l'intégrale du facteur de forme

Nous évaluons les intégrales nécessaires au calcul du facteur de forme en utilisant un échantillonnage des surfaces. Pour chaque paire d'éléments, on commence par générer des points d'échantillonnage sur la surface. Ensuite on évalue la contribution élémentaire point à point nécessaire à lévaluation de l'intégrale. On évalue également la longueur des trajets point à point pour le calcul des délais de propagation.

Nous avons implémenté deux techniques d'échantillonnage :

- échantillonnage récursif [SP94] ;
- échantillonnage de Halton [Kel97].

L'échantillonage récursif consiste à utiliser comme points d'échantillonage des points placés sur une grille que l'on va subdiviser récursivement comme on subdivise les surfaces pour le raffinement. On peut ainsi garantir que chaque point d'échantillonage va représenter une proportion donnée de la surface de départ. Pour éviter d'obtenir un échantillonage trop régulier, on peut perturber aléatoirement les points obtenues par un déplacement aléatoire dans le plan de la surface.

L'échantillonage de Halton permet toutefois de selectionner de manière plus efficace un nombre de points aléatoires sur la surface qui sont «bien répartis». Pour plus de détails sur cette technique déchantillonage et le pseudo-code nécessaire au tirage aléatoire, nous renvoyons le lecteur à [Kel97].

Annexe C

Transformée de Hilbert et applications

DANS cette annexe, nous nous interessons à la transformée de Hilbert [OS75] et deux de ses applications qui nous ont intéressées dans le cadre de nos travaux : la reconstruction d'un filtre à phase minimale à partir des données de son spectre uniquement et la réalisation d'un encodeur permettant de générer un signal de type *Dolby Surround©* permettant de recréer des effets spatiaux simples à partir d'un signal stéréophonique.

C.1 Calcul de filtres à phase minimale

Bien souvent, on ne dispose pas dans la pratique des différentes réponses impulsionnelles des sources, récepteurs et de la réponse des matériaux. La plupart des valeurs disponibles correspondent à des spectres donnés en bandes d'octaves. Néanmoins, il peut être nécessaire de disposer d'un réponse impulsionnelle pour des opérations de filtrage. Une méthode, basée sur la transformée de Hilbert, permet de reconstruire à partir de la donnée du spectre, une réponse impulsionnelle de même spectre avec une phase minimale. De toutes les réponses possibles, elle est la plus courte. L'interêt d'une telle représentation vient du fait que les réponses impulsionnelles, en particulier dans le cas d'enceintes électroacoustiques peuvent souvent être considérées á phase minimale.

Dans un système à phase minimale, la phase est reliée au logarithme du spectre du signal par une transformée de Hilbert. L'algorithme se déroule en deux étapes :

- calcul de la transformée de Hilbert du logarithme (népérien) du spectre d'entrée ;
- à partir de cette phase, reconstruction d'un signal dont le spectre est identique au spectre d'origine.

On commence donc par calculer le logarithme du spectre d'entrée. On choisi une phase nulle et on calcule la transformée de Fourier inverse du ce signal complexe. La transformée de Hilbert se calcule en fenêtrant le signal par la suite suivante :

$$U_n = \begin{cases} 0 & n < 0 \\ 1 & n = 0 \\ 2 & n > 0 \end{cases} \tag{C.1}$$

Dans une deuxième phase, on calcule la transformée de Fourier du signal obtenu. La partie imaginaire du résultat est la phase minimale recherchée, la partie réelle est le logarithme spectre d'entrée. A partir de ces deux informations on peut reconstituer le signal complexe voulu. Enfin, une dernière transformée de Fourier inverse nous donne la réponse impulsionnelle recherchée.

L'algorithme général est décrit dans la Figure C.1. De plus amples détails sur cette opération peuvent être trouvés dans [Dig79].

```
void minimalPhase(float *spectrum,int n,float *result)
{
    complex tab[n]
    for int i = 0 to n-1
    {
        tab[i].re = ln(spectrum[n])
        tab[i].im = 0.0
    }
    inverseFFT(tab,n)
    for int i = 1 to n-1
    {
        tab[i+n/2].re = tab[i+n/2].im = 0.0
        tab[i].re *= 2.0
        tab[i].im *= 2.0
    }
    FFT(tab,n)
    for int i = 0 to n-1
    {
        float r = exp(tab[i].re)
        tab[i].re = r * cos(tab[i].im)
        tab[i].im = r * sin(tab[i].im)
    }
    inverseFFT(tab,n)
    for int i = 0 to n-1
        result[i] = tab[i].re
}
```

Figure C.1 - *Pseudo-code pour le calcul de filtres à phase minimale.*

C.2 Matriçage de signaux «surround»

La transformée de Hilbert permet également de réaliser un encodage des signaux de type *Dolby Surround*. Ce format, développé par les laboratoires *Dolby* pour les bandes-son cinématographiques (puis également pour des applications domestiques), permet de coder dans un signal stéréophonique, l'information relative à quatre canaux : gauche, droit, centre et arrière. Le dispositif de restitution associé comprend généralement cinq enceintes (trois à l'avant et deux à l'arrière, *cf.* Figure C.2). Les deux enceintes droite et gauche correspondent à de la stéréophonie classique. L'enceinte centrale permet de limiter l'effet de précédance pour des auditeurs situés sur les côtés de la zone découte. Utilisée principalement pour les dialogues, elle permet de maintenir leur provenance au centre de l'écran (situé généralement entre les enceintes gauche et droite). Les deux enceintes arrières diffusent le même signal monophonique, permettant de recréer des effets d'ambiance simples.

Des détails sur ce format d'encodage peuvent être trouvés dans [Dre88]. Ce format sonore est particulièrement intéressant puisqu'il permet de coder une information sonore spatiale (simple) dans un signal stéréophonique classique. De plus, les décodeurs sont de plus en plus répendus.

On peut réaliser un encodage satisfaisant de manière simple a partir de quatre canaux d'entrée : gauche, droit, centre, arrière. Les canaux droit et gauche sont répartis directement sur les pistes droite et gauche du signal stéréophonique. Le signal central est equi-réparti sur les deux pistes (son amplitude

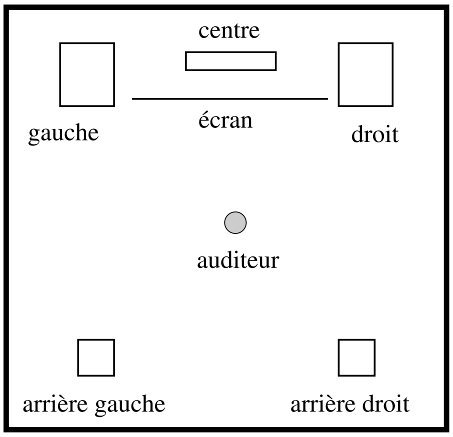


Figure C.2 - *Système de restitution «surround» classique.*

étant au préalable divisée par deux). Le signal arrière va également être réparti sur les pistes droite et gauche mais nécessite un traitement particulier. Il doit être déphasé de $+90^o$ et -90^o respectivement avant d'être ajouté aux pistes gauche et droite (son amplitude est également divisée par deux). Pour réaliser le déphasage, on utilise une transformée de Hilbert de manière similaire au cas précédent. On commence par effectuer la transformée de Fourier du signal d'entrée. Puis on multiplie le résultat par la suite U_n (*cf.* équation C.1) avant d'en prendre la transformée de Fourier inverse. Le signal recherché est la *partie imaginaire* du signal résultant. On peut alors ajouter le signal obtenu à la piste gauche et son opposé à la piste droite (on a donc bien le déphasage de 180^o recherché entre les deux pistes).

Annexe D

Articles

NOUS incluons dans cette annexe plusieurs articles en langue anglaise concernant une partie de ce travail correspondant aux chapitres 6, 7, 9 de ce mémoire. Tout d'abord, on trouvera un article concernant la réalisation de notre système de rendu intégré image-son avec traitement qualitatif de la diffraction par des obstacles :

– **Soundtracks for Computer Animation:**
  Sound Rendering in Dynamic Environments with Occlusions
  N. Tsingos and J.D. Gascuel, in proceedings of Graphics Interface'97,
  Kelowna B.C., Canada, May 1997.

L'article suivant décrit l'aspect plus quantitatif et adaptatif de notre approche basée sur une technique de radiosité hiérarchique:

– **Acoustic simulation using hierarchical time-varying radiant exchanges**
  unpublished, January 1998.

Une version résumée peut être trouvée dans les «*visual proceedings*» de Siggraph 1997 [TG97a].

Enfin, on trouvera une publication détaillant notre traitement semi-quantitatif de la diffraction par la théorie de Fresnel-Kirchhoff:

– **Fast Rendering of Sound Occlusion and Diffraction Effects**
  for Virtual Acoustic Environments
  N. Tsingos and J.D. Gascuel, Preprint no. 4699 (P4-7), 104th AES convention,
  Amsterdam, The Netherlands, May 1998.

Table des figures

Liste des tableaux

Bibliographie

[AB79] J.B. Allen et D.A. Berkley. – Image method for efficiently simulating small room acoustics. *Journal of the Acoustical Society of America*, vol. 65, n° 4, 1979.

[AH93] L. Aupperle et P. Hanrahan. – A hierarchical illumination algorithm for surfaces with glossy reflections. *ACM Computer Graphics, SIGGRAPH'93 Proceedings*, pp. 155–162, août 1993.

[Ama84] John Amanatides. – Ray tracing with cones. *ACM Computer Graphics, SIGGRAPH'84 Proceedings*, vol. 18, n° 3, pp. 129–135, juillet 1984.

[And85] Y. Ando. – *Concert Hall Acoustics*. – Springer-Verlag, 1985.

[Arv95] J. Arvo. – Application of irradiance tensors to the simulation of non lambertian phenomena. *ACM Computer Graphics, Annual Conference Series, SIGGRAPH'95 Proceedings*, août 1995.

[Ast95] Peter Astheimer. – Acoustic simulation for visualisation and virtual reality. *EUROGRAPHICS*, vol. 3, n° 23, juillet 1995.

[Bar71] M. Barron. – The subjective effects of first reflections in concert halls - the need for lateral reflections. *Journal of Sound and Vibration*, vol. 15, n° 4, pp. 475–494, 1971.

[BdV98] A.J. Berkhout et D. de Vries. – Simulation of impulse responses in enclosed spaces. *Proc. 104th Audio Engineering Society Convention*, preprint n° 4714 (P8-7), Amsterdam, Netherlands, mai 1998.

[Beg91] D.R. Begault. – Challenges to the successful implementation of 3D sound. *Journal of the Audio Engineering Society*, vol. 39, n° 11, pp. 864–870, novembre 1991.

[Beg94] Durand R. Begault. – *3D Sound for Virtual Reality and Multimedia*. – Academic Press Professional, 1994.

[Bel95] M. Bellanger. – *Traitement numérique du signal : théorie et pratique (5 edition)*. – Masson ed., 1995.

[Ber86] Leo L. Beranek. – *Acoustics*. – Published for the Acoustical Society of America through the American Institute of Physics, 1986.

[Ber88] Henry L. Bertoni (coordonnateur), IEEE Vehicular Technology Society Committee on Radio Propagation. – Coverage prediction for mobile radio systems operating in the 800/900 MHz frequency range. *IEEE Transactions on Vehicular Technology (Special Issue on Mobile Radio Propagation)*, vol. 37, n° 1, février 1988.

[Ber96] Leo L. Beranek. – *Concert and opera halls: how they sound*. – Published for the Acoustical Society of America through the American Institute of Physics, 1996.

[Bla83] J. Blauert. – *Spatial Hearing : The Psychophysics of Human Sound Localization*. – M.I.T. Press, Cambridge, MA, 1983.

[BM81] M. Barron et A.H. Marshall. – Spatial impression due to early lateral reflections in concert halls: the derivation of a physical measure. *Journal of Sound and Vibration*, vol. 77, n° 2, pp. 211–232, 1981.

[Bor84] J. Borish. – Extension of the image model to arbitrary polyhedra. *Journal of the Acoustical Society of America*, vol. 75, n° 6, 1984.

[Bul77] Kenneth Bullington. – Radio propagation for vehicular communication. *IEEE Transactions on Vehicular Technology*, vol. VT-26, n° 4, pp. 295–308, novembre 1977.

[Bur92a] D.A. Burgess. – *Real-Time Audio Spatialization With Inexpensive Hardware*. – Rapport technique n° GIT-GVU-92-20, Georgia Institute of Technology, 1992.

[Bur92b] D.A. Burgess. – *Techniques for Low Cost Spatial Audio*. – Rapport technique n° GIT-GVU-92-09, Georgia Institute of Technology, 1992.

[BV93a] D.A. Burgess et Jonke C. Verlinden. – *An architecture for Spatial Audio servers*. – Rapport technique n° GIT-GVU-93-34, Georgia Institute of Technology, 1993.

[BV93b] D.A. Burgess et Jonke C. Verlinden. – *A first experience with Spatial Audio in a Virtual Environment*. – Rapport technique n° GIT-GVU-93-31, Georgia Institute of Technology, 1993.

[Can65] M. Canac. – *L'acoustique des théâtres antiques*. – Bulletin du Groupe d'Acoustique Musicale n° 13, Université de Paris VI, 1965.

[CB89] Duane H. Cooper et Jevald L. Bauck. – Prospects for transaural recording. *Journal of the Audio Engineering Society*, vol. 37, n° 1/2, pp. 3–19, janvier/février 1989.

[CCWG88] M.F. Cohen, S.E. Chen, J.R. Wallace et D.P. Greenberg. – A progressive refinment approach to fast radiosity image generation. *ACM Computer Graphics, SIGGRAPH'88 Proceedings*, vol. 22, n° 4, 1988.

[Che90] Shenchang E. Chen. – Incremental radiosity: an extension of progressive radiosity to an interactive image synthesis system. *ACM Computer Graphics, SIGGRAPH'90 Proceedings*, vol. 24, n° 4, pp. 135–144, août 1990.

[Cho71] J.M. Chowning. – The simulation of moving sound sources. *Journal of the Audio Engineering Society*, vol. 19, n° 19, pp. 2–6, janvier 1971.

[Col68] Coleman. – Dual role of frequency spectrum in determination of auditory distance. *Journal of the Acoustical Society of America*, vol. 44, n° 2, pp. 631–632, 1968.

[Cot97] Francis Cottet. – *Traitement des signaux et acquisition de données*. – Dunod ed., 1997.

[CSSD96] P.H. Christensen, E. Stollnitz, D.H. Salesin et T.D. DeRose. – Global illumination of glossy environments using wavelets and importance. *ACM Transactions on Graphics*, vol. 15, n° 1, pp. 37–71, janvier 1996.

[CW93] Michael F. Cohen et John R. Wallace. – *Radiosity and Realistic Image Synthesis*. – Academic Press Professional, 1993.

[CW95] Michael Cohen et Elisabeth M. Wenzel. – *The Design of Multidimensional Sound Interfaces*. – Rapport technique n° 95-1-004, Human Interface Laboratory, The University of Aizu, 1995.

[Dal96] Bengt-Inge L. Dalenbäck. – Room acoustic prediction based on a unified treatment of diffuse and specular reflection. *Journal of the Acoustical Society of America*, vol. 100, n° 2, pp. 899–909, août 1996.

[Dau78] A. Daumas. – Etude de la diffraction par un écran mince posé sur le sol. *Acustica*, vol. 40, n° 4, 1978.

[Dau96] D. Dauger. – Simulation and study of Fresnel diffraction for arbitrary two-dimensional apertures. *Computer in Physics*, vol. 10, n° 6, pp. 591–604, novembre/décembre 1996.

[Def96] Jerôme Defrance. – *Méthode Analytique pour le calcul de propagation de bruit exterieur*. – Thèse de doctorat, Université du Maine & Centre Scientifique et Technique du Bâtiment, 1996.

[Der97] Philippe Derogis. – *Rapport sur l'implémentation des panpots ambisonics dans le Spatialisateur*. – Rapport technique, Paris, Institut de Recherche and Coordination Acoustique/Musique (IRCAM), juillet 1997.

[dEV92] A.W.M. Van den Enden et N.A.M. Verhoeckx. – *Traitement numérique du signal : une introduction*. – Masson ed., 1992.

[Dig79] Digital Signal Processing committee, IEEE Acoustics, Speech and Signal Processing Society (coordonnateur). – *Programs for Digital Signal Processing*. – IEEE Press, 1979.

[dP83] G. de Poli. – A tutorial on digital sound synthesis technique. *Computer Music Journal*, vol. 7, n° 4, Winter 1983.

[DQ96] David Declercq et André Quinquis. – *Le signal aléatoire*. – Hermès, Paris, 1996.

[Dre88] R. Dressler. – *Dolby PROLOGIC Surround Decoder Principles of Operation*. – Rapport technique n° S89/8540/8624, Dolby Laboratories Inc., 1988.

[DS97] G. Drettakis et F.X. Sillion. – Interactive update of global illumination using a line-space hierarchy. *ACM Computer Graphics, SIGGRAPH'97 Proceedings*, pp. 57–64, août 1997.

[Dur77] J. Durkin. – Computer prediction of service areas for VHF and UHF land mobile radio services. *IEEE Transactions on Vehicular Technology*, vol. VT-26, n° 4, pp. 323–327, novembre 1977.

[DW83] A.P. Dowling et J.E. Ffowcs Williams. – *Sound and Sources of Sound*. – Ellis Horwood publishers, 1983.

[Eme94] Toni Emerson. – *Virtual Audio Bibliography*. – Rapport technique n° B-94-2, University of Washington, Human Interface Technology Lab, 1994. http://www.hitl.washington.edu/kb/audio-bib.html.

[Eme95] Marc Emerit. – *Simulation Binaurale de l'Acoustique de Salles de Concert*. – Thèse de doctorat, Centre Scientifique et Technique du Bâtiment, 1995.

[ETA98] Michael J. Evans, Anthony I. Tew et James A.S. Angus. – Relative spatialisation performance of ambisonic and transaural speech. *Proc. 104th Audio Engineering Society Convention*, preprint n° 4668 (P4-6), Amsterdam, Netherlands, mai 1998.

[Eve94] F. Alton Everest. – *The Master Handbook of Acoustics (3rd edition)*. – TAB Books (Imprint of McGraw-Hill), 1994.

[FCE+98] T. Funkhouser, I. Carlbom, G. Elko, G. Pingali, M. Sondhi et J. West. – A beam tracing approach to acoustic modeling for interactive virtual environments. *ACM Computer Graphics, Annual Conference Series, SIGGRAPH'98 Proceedings*, pp. 21–32, juillet 1998.

[Fil94] P.J.T. Filippi (coordonnateur). – *Acoustique générale*. – Les Editions de Physique, 1994.

[Fla93] Patrick Flandrin. – *Temps-fréquence*. – Hermès, Paris, 1993.

[Fla95] W. Flageollet. – Le son du réel : Le son du son au cinéma. *Proc. of Rencontres Musicales Pluridisciplinaires, Le Son & L'Espace*, pp. 45–52, Lyon, mars 1995.

[FR91] N.H. Fletcher et T.D. Rossing. – *The Physics of Musical Instruments*. – Springer Verlag, 1991.

[FST92] Thomas A. Funkhouser, Carlo H. Séquin et Seth J. Teller. – Management of large amounts of data in interactive building walkthroughs. *Computer Graphics (Proc. ACM Symposium on Interactive 3D graphics*, pp. 11–20, Cambridge, MA, mars 1992.

[Fur90] Y. Furue. – Sound propagation from the inside to the outside of a room through an aperture. *Applied Acoustics*, vol. 31, n° 1, pp. 133–146, 1990.

[FVFH90] Foley, VanDam, Feiner et Hughes. – *Computer graphics, principles and practice*. – Addison Wesley, 1990.

[FWT91] S.H. Foster, E.M. Wenzel et R.M. Taylor. – Real-time synthesis of complex environments. *Proc. of the ASSP (IEEE) Workshop on Application of Signal Processing to Audio and Acoustics*, 1991.

[Gar95] Bill Gardner. – *Transaural 3D audio*. – Rapport technique n° 342, M.I.T. Media Lab Perceptual Computing, juillet 1995.

[Gas94] Jean-Dominique Gascuel. – Fabule : un environnement de recherche pour l'animation et la simulation. *Troisième groupe de travail «Animation et Simulation». Thème : Les simulateurs.*, Lille, octobre 1994.

[GM94] Bill Gardner et Keith Martin. – *HRTF Measurements of a KEMAR Dummy-Head Microphone*. – Rapport technique n° 280, M.I.T. Media Lab Perceptual Computing, mai 1994.

[Gra98] Xavier Granier. – *Illumination de scènes complexes par méthode de radiosité avec mémoire limitée*. – Rapport de DEA, Institut National Polytechnique de Grenoble, septembre 1998.

[GSCH93] Steven J. Gortler, Peter Schroder, Michael F. Cohen et Pat Hanrahan. – Wavelet radiosity. *ACM Computer Graphics, SIGGRAPH'93 Proceedings*, pp. 221–230, 1993.

[GSG90] D.W. George, F.X. Sillion et D.P. Greenberg. – Radiosity redistribution for dynamic environments. *IEEE Computer Graphics and Applications*, pp. 26–34, juillet 1990.

[GSH94] Reid Gershbein, Peter Schröder et Pat Hanrahan. – Textures and radiosity : Controlling emission and reflexion with texture maps. *ACM Computer Graphics, SIGGRAPH'94 Proceedings, Annual Conference Series*, pp. 51–58, 1994.

[GTGB84] C.M. Goral, K.E. Torrance, D.P. Greenberg et B. Battaile. – Modeling the interaction of light between diffuse surfaces. *ACM Computer Graphics, SIGGRAPH'84 Proceedings*, vol. 18, n° 3, 1984.

[Har83] W.M. Hartmann. – Localization of sound in rooms. *Journal of the Acoustical Society of America*, vol. 74, n° 5, pp. 1380–1391, novembre 1983.

[HCWH91] D.C. Hothersall, S.N. Chandler-Wilde et M.N. Hajmirzae. – Efficiency of single noise barriers. *Journal of Sound and Vibration*, vol. 146, n° 2, pp. 303–322, 1991.

[Hec87] E. Hecht. – *Optics, second edition*. – Addison Wesley, 1987.

[Hei93] R. Heinz. – Binaural room simulation based on an image source model with addition of statistical methods to include the diffuse sound scattering of walls and to predict the reverberant tail. *Applied Acoustics*, vol. 38, pp. 145–159, 1993.

[HF97] C.H. Haan et F.R. Fricke. – An evaluation of the importance of surface diffusivity in concert halls. *Applied Acoustics*, vol. 51, n° 1, pp. 53–69, 1997.

[HFGL95] J. Hahn, H. Fouad, L. Gritz et J. Wong Lee. – Integrating sounds and motions in virtual environments. *Sound for Animation and Virtual Reality, Siggraph'95 Course #10 Notes*, août 1995.

[HGL$^+$95] J. Hahn, J. Geigel, J. Wong Lee, L. Gritz, T. Takala et S. Mishra. – An integrated approach to motion and sound. *Journal of Visualisation and Computer Animation*, vol. 28, n° 2, juillet 1995.

[HH84] P. Heckbert et P. Hanrahan. – Beam tracing polygonal objects. *ACM Computer Graphics, SIGGRAPH'84 Proceedings*, vol. 18, n° 3, pp. 119–127, juillet 1984.

[HKVH94] J. Huipaniemi, M. Karjalainen, V. Valimaki et T. Huotilainen. – Virtual instruments in virtual rooms - a realtime binaural room simulation environment for physical models of musical instruments. *Proc. of ICMC*, pp. 455–462, 1994.

[Hod91] M. Hodgson. – Evidence of diffuse surface reflections in rooms. *Journal of the Acoustical Society of America*, vol. 89, n° 2, pp. 765–771, février 1991.

[Hod96] Murray Hodgson. – When is diffuse-field theory applicable? *Applied Acoustics*, vol. 38, n° 3, pp. 197–207, 1996.

[Hol91] Tomlinson Holman. – New factors in sound for cinema and television. *Journal of the Audio Engineering Society*, vol. 39, n° 7/8, pp. 529–539, 1991.

[Hol97] Tomlinson Holman. – *Sound for Film and Television*. – Focal Press, 1997.

[HSA91] P. Hanrahan, D. Saltzman et L. Aupperle. – A rapid hierarchical radiosity algorithm. *ACM Computer Graphics, SIGGRAPH'91 Proceedings*, vol. 25, n° 4, juillet 1991.

[HW94] C. Hugonnet et P. Walder. – *Theorie et Pratique de la Prise de Son Stereophonique*. – Eyrolles ed., 1994.

[IC86] David S. Immel et Michael F. Cohen. – A radiosity method for non-diffuse environments. *ACM Computer Graphics, SIGGRAPH'86 Proceedings*, vol. 20, n° 4, pp. 133–142, 1986.

[Ing53] U. Ingard. – A review of the influence of meteorological conditions on sound propagation. *Journal of the Acoustical Society of America*, vol. 25, n° 3, 1953.

[ISO93] Acoustique - Atténuation du son lors de sa propagation à l'air libre. *Norme internationale, ISO 9613-1:1993(F)*, 1993.

[Jea98] P. Jean. – A variational approach for the study of outdoor sound propagation and application to railway noise. *Journal of Sound and Vibration*, vol. 212, n° 2, pp. 275–294, 1998.

[JLW95] J.M. Jot, V. Larcher et O. Warusfel. – Digital signal processing issues in the context of binaural and transaural stereophony. *Proc. 98th Audio Engineering Society Convention*, preprint n° 3980 (I6), Paris, 1995.

[Jon86] D.S. Jones. – *Acoustic and Electromagnetic waves*. – Oxford science publications, 1986.

[Jot92a] Jean-Marc Jot. – *Etude et réalisation d'un spatialisateur de sons par modèles physique et perceptifs*. – Thèse de doctorat, Ecole Normale Supérieure des Télécommunications, Paris, 1992.

[Jot92b] J.M. Jot. – An analysis/synthesis approach to real-time artificial reverberation. *Proc. of ICASSP*, 1992.

[Joy75] W.B. Joyce. – Sabine's reverberation time and ergodic auditoriums. *Journal of the Acoustical Society of America*, vol. 58, n° 3, pp. 643–655, septembre 1975.

[JW95] J.M. Jot et O. Warusfel. – Le spatialisateur. *Proc. of Rencontres Musicales Pluridisciplinaires, Le Son & L'Espace*, pp. 103–108, Lyon, mars 1995.

[Kaj86] J.T. Kajiya. – The rendering equation. *ACM Computer Graphics, SIGGRAPH'86 Proceedings*, vol. 20, n° 4, 1986.

[Kal93] Roy S. Kalawsky. – *The science of virtual reality and virtual environments*. – Addison Wesley, 1993.

[Kaw81] T. Kawai. – Sound diffraction by a many sided barrier or pillar. *Journal of Sound and Vibration*, vol. 79, n° 2, pp. 229–242, 1981.

[KDS93] M. Kleiner, B.I. Dalenbäk et P. Svensson. – Auralization - an overview. *Journal of the Audio Engineering Society*, vol. 41, n° 11, pp. 861–875, novembre 1993.

[Kel62] J.B. Keller. – Geometrical theory of diffraction. *Journal of the Optical Society of America*, vol. 52, n° 2, pp. 116–130, 1962.

[Kel97] Alexander Keller. – Instant radiosity. *ACM Computer Graphics, Annual Conference Series, SIGGRAPH'97 Proceedings*, pp. 49–54, août 1997.

[KGB97] Mendel Kleiner, Hans Gustafsson et Joakim Backman. – Measurement of directional scattering coefficients using near-field acoustic holography and spacial transformation of sound field. *Journal of the Audio Engineering Society*, vol. 45, n° 5, pp. 331–346, mai 1997.

[KL95] S. Kopuz et N. Lalor. – Analysis of interior acoustic fields using the finite element method and the boundary element method. *Applied Acoustics*, vol. 45, pp. 193–210, 1995.

[Klu91] A. Kludszuweit. – Time iterative boundary element method (TIBEM) - a new numerical method of four-dimensional system analysis for the calculation of the spatial Impulse Response. *Acustica*, vol. 75, pp. 17–27, 1991. – En allemand.

[KOK93] M. Kleiner, R. Orlowski et J. Kirszenstein. – A comparison between results from a physical scale model and a computer image source model for architectural acoustics. *Applied Acoustics*, vol. 38, pp. 245–265, 1993.

[KP74] Robert G. Kouyoumjian et Prabhakar H. Pathak. – A uniform geometrical theory of diffraction for an edge in a perfectly conducting surface. *Proceedings of IEEE*, vol. 62, pp. 1448–1461, novembre 1974.

[Kut91a] H. Kuttruff. – On the audibility of phase distorsions in rooms and its significance for sound reproduciton and digital simulation in room acoustics. *Acustica*, vol. 74, n° 1, 1991.

[Kut91b] Heinrich Kuttruff. – *Room Acoustics (3rd edition)*. – Elseiver Applied Science, 1991.

[Kut93] K.H. Kuttruff. – Auralization of impulse responses modeled on the basis of ray-tracing results. *Journal of the Audio Engineering Society*, vol. 41, n° 11, pp. 876–880, novembre 1993.

[Lam92] Alexis Lamouret. – *Utilisation des techniques d'infographie pour simuler la propagation d'ondes porteuses du radiotéléphone*. – Rapport de stage, Société Française de Radiotéléphones, 1992.

[Lam96] Y.W. Lam. – The dependance of diffusion parameters in a room acoustics prediction model on auditorium sizes and shapes. *Journal of the Acoustical Society of America*, vol. 100, n° 4, pp. 2193–2203, octobre 1996.

[LB92] H. Lehnert et J. Blauert. – Principles of binaural room simulation. *Applied Acoustics*, vol. 36, pp. 259–291, 1992.

[Lee90] William C. Y. Lee. – Estimate of channel capacity in Rayleigh fading environment. *IEEE Transactions on Vehicular Technology*, vol. 39, n° 3, pp. 187–189, août 1990.

[Lee98] Martin J. Leese. – Ambisonic surround sound FAQ (version 2.8), 1998. http://www.omg.unb.ca/ mleese/faq_latest.html.

[Leh93] H. Lehnert. – Systematic errors of the ray-tracing algorithm. *Applied Acoustics*, vol. 38, 1993.

[Lei78] E. Leipp. – *La qualité acoustique des lieux d'écoute*. – Bulletin du Groupe d'Acoustique Musicale n° 99, Université de Paris VI, novembre 1978.

[Lew93] T. Lewers. – A combined beam tracing and radiant exchange computer model of room acoustics. *Applied Acoustics*, vol. 38, 1993.

[Lor94] David I. Lorenson. – *Indoor Radio Channel Propagation Modelling by Ray Tracing Techniques*. – Thèse de doctorat, University of Edinborough, 1994.

[LS97] J. Lavergnat et M. Sylvain. – *Propagation des ondes radioelectriques - introduction*. – Masson ed., 1997.

[LSG94] Dani Lischinski, Brian Smits et Donald P. Greenberg. – Bounds and error estimates for radiosity. *ACM Computer Graphics, SIGGRAPH '94 Proceedings*, pp. 67–74, juillet 1994.

[LST94] B. Laurel, R. Strickland et R. Tow. – Placeholder : Landscape and narrative in virtual environments. *ACM Computer Graphics*, vol. 28, n° 2, pp. 118–127, May 1994.

[Mae68] Z. Maekawa. – Noise reduction by screens. *Applied Acoustics*, vol. 1, 1968.

[Mal86] Christian Malcurt. – *Simulations informatiques pour prédire les critères de qualification acoustique des salles. Comparaison des valeurs mesurées et calculées dans une salle à acoustique variable*. – Thèse de doctorat, Laboratoire Acoustique, Métrologie, Instrumentation, Toulouse, 1986.

[MF91] C.J. McCabe et D.J. Furlong. – Special stereo surround sound pan-pot. *Proc. 90th Audio Engineering Society Convention*, preprint n° 3067 (L-4), Paris, février 1991.

[MOD96] M. Monks, B.M. Oh et J. Dorsey. – Acoustic simulation and visualisation using a new unified beam tracing and image source approach. *Proc. Audio Engineering Society Convention*, 1996.

[MOD98] M. Monks, B.M. Oh et J. Dorsey. – *Audioptimization: goal based acoustic design*. – Rapport technique n° MIT/LCS/TM-588, M.I.T. Laboratory for Computer Science, septembre 1998.

[Møl89] Henrik Møller. – Reproduction of artificial-head recordings through loudspeakers. *Journal of the Audio Engineering Society*, vol. 37, n° 1/2, pp. 30–33, janvier/février 1989.

[Møl92a] Henrik Møller. – Fundamentals of binaural technology. *Applied Acoustics*, vol. 36, pp. 171–218, 1992.

[Møl92b] Henrik Møller. – Interfacing room simulation programs and auralisation systems. *Applied Acoustics*, vol. 38, pp. 333–347, 1992.

[Moo79] J.A. Moorer. – About this reverberation business. *Computer Music Journal*, vol. 23, n° 2, 1979.

[Moo83] F.R. Moore. – A general model for spatial processing of sounds. *Computer Music Journal*, vol. 7, n° 3, pp. 6–15, Fall 1983.

[MR92] T. McCormick et F. Rumsey. – *Sound and Recording : An Introduction*. – Butterworth-Heinemann, 1992.

[MvMV93] J. Martin, D. van Maercke et J.P. Vian. – Binaural simulation of concert halls : A new approach for the binaural reverberation process. *Journal of the Acoustical Society of America*, vol. 94, pp. 3255–3263, décembre 1993.

[Mye95] T. Myers. – Sound in live action and animation : A comparative look. *Sound for Animation and Virtual Reality, Siggraph'95 Course #10 Notes*, août 1995.

[Nay93] J.M. Naylor. – Odeon - another hybrid room acoustical model. *Applied Acoustics*, vol. 38, n° 1, pp. 131–143, 1993.

[NDM93] J. Neider, T. Davis et W. Mason. – *OpenGL Programming Guide*. – Addison-Wesley, 1993.

[NDR95] J. Nimeroff, J. Dorsey et H. Rushmeier. – A framework for global illumination in animated environments. *Sixth Eurographics Workshop on Rendering*, pp. 223–236, Dublin, Ireland, juin 1995.

[NDR96] J. Nimeroff, J. Dorsey et H. Rushmeier. – Implementation and analysis of an image-based global illumination framework for animated environments. *IEEE Transactions on Visualization and Computer Graphics*, vol. 2, n° 4, décembre 1996.

[NMT93] Kiyoshi Nakagawa, Toru Miyajima et Yasuhiko Tahara. – An improved geometrical sound field analysis in rooms using scattered sound and an audible room acoustics simulator. *Applied Acoustics*, vol. 38, pp. 115–129, 1993.

[NT95] H. Noser et D. Thalmann. – Synthetic vision and audition for digital actors. *Proc. of EUROGRAPHICS'95*, vol. 14, n° 3, pp. 326–336, 1995.

[OR98] Eyal Ofek et Ari Rappoport. – Interactive reflections on curved objects. *ACM Computer Graphics, SIGGRAPH'98 Proceedings*, pp. 333–342, juillet 1998.

[OS75] Alan V. Oppenheim et Ronald W. Schafer. – *Digital Signal Processing*. – Prentice-Hall Inc., 1975.

[Pet92] J. Petykiewicz. – *Wave optics*. – Kluwer Academic Publishers, 1992.

[PF93] Stephen Travis Pope et Lennart E. Fahlén. – The use of 3D audio in a synthetic environment: An aural renderer for a distributed virtual reality. *ICMC Proceedings*, pp. 146–149, 1993.

[PM77] D.R. Perrot et A.D. Musicant. – Minimum auditory movement angle : Binaural localization of moving sound sources. *Journal of the Acoustical Society of America*, vol. 62, n° 6, pp. 1463–1466, 1977.

[Pol93] J.D. Polack. – Playing billiards in the concert hall: The mathematical foundations of geometrical room acoustics. *Applied Acoustics*, vol. 38, pp. 235–244, 1993.

[PPD98] Eric Paquette, Pierre Poulin et George Drettakis. – A light hierarchy for fast rendering of scenes with many lights. *EUROGRAPHICS'98*, vol. 17, n° 3, 1998.

[PS69] R. Plomp et H. J. M. Steeneken. – Effect of phase on the timbre of complex tones. *Journal of the Acoustical Society of America*, vol. 46, n° 2 (part 2), pp. 409–421, 1969.

[PST+96] R. Pausch, J. Snoddy, R. Taylor, S. Watson et E. Haseltine. – Disney's aladdin: 1st steps toward storytelling in virtual reality. *Proc. of SIGGRAPH'96 conference. Orlando. Florida.*, août 1996.

[PTVF92] William Press, Saul Teukolsky, William Vetterling et Brian Flannery. – *Numerical Recipes in C, second edition*. – New York, USA, Cambridge University Press, 1992.

[RAOR91a] A. J. Rustako, Noach Amitay, G. J. Owens et R. S. Roman. – Attenuation and diffraction effects from truck blockage of an 11-GHz line-of-sight microcellular mobile radio path. *IEEE Transactions on Vehicular Technology*, vol. 40, n° 1, pp. 211–215, février 1991.

[RAOR91b] A. J. Rustako, Noach Amitay, G. J. Owens et R. S. Roman. – Radio propagation at microwave frequencies for line-of-sight microcellular mobile and personal communications. *IEEE Transactions on Vehicular Technology*, vol. 40, n° 1, pp. 203–210, février 1991.

[Rin93] J.H. Rindel. – Modelling the angle-dependant pressure reflection factor. *Applied Acoustics*, vol. 38, pp. 223–234, 1993.

[RM95] A. Reilly et D. McGrath. – Convolution processing for realistic reverberation. *Proc. 98th Audio Engineering Society Convention*, preprint n° 3977 (I3), Paris, février 1995.

[RPV93] H. Rushmeier, C. Patterson et A. Veerasamy. – Geometric simplification for indirect illumination calculations. *Proceedings of Graphics Interface'93*, pp. 227–236, Toronto, 1993.

[RSX98] Intel© Realistic Sound Experience (3D RSX), 1998. http://developer.intel.com/ial/rsx/index.htm.

[Rub91] Per Rubak. – Headphone signal processing system for out-of-the head localization. *Proc. 90th Audio Engineering Society Convention*, preprint n° 3063 (K-5), Paris, février 1991.

[SAF+87] G. Steinke, W. Ahnert, P. Fels, W. Hoeg et F. Steffen. – True directional sound system orientated to the original sound of diffuse sound structures - new applications of the Delta Stereophony System (DSS). *Proc. 82nd Audio Engineering Society Convention*, preprint n° 2427 (B-2), London, mai 1987.

[Sal95] C. Salès. – Processeur d'espace sonore Roland RSS-10. *Proc. of Rencontres Musicales Pluridisciplinaires, Le Son & L'espace, Lyon*, mars 1995.

[SAWG91] F.X. Sillion, J.R. Arvo, S.H. Westin et D.P. Greenberg. – A global illumination solution for general reflectance distributions. *ACM Computer Graphics, SIGGRAPH'91 Proceedings*, vol. 25, n° 4, juillet 1991.

[SB95] Holger Strauss et Jens Blauert. – Virtual auditory environments. *Proc. FIVE Conference'95*, pp. 123–131, London, 1995.

[Sch62] M.R. Schroeder. – Natural sounding artificial reverberation. *Journal of the Audio Engineering Society*, vol. 10, n° 3, pp. 219–223, 1962.

[SG89] Adam Stettner et Donald P. Greenberg. – Computer graphics visualisation for acoustic simulation. *ACM Computer Graphics, SIGGRAPH'89 Proceedings*, vol. 23, n° 3, pp. 195–206, juillet 1989.

[SH93] Peter Schröder et Pat Hanrahan. – On the form factor between two polygons. *ACM Computer Graphics, SIGGRAPH'93 Proceedings*, pp. 163–164, août 1993.

[SH94] Jesper Sandvad et Dorte Hammershøi. – Binaural auralization. Comparison of FIR and IIR filter representation of HIRs. *Proc. 96th Audio Engineering Society Convention*, preprint n° 3862 (P4.5), Amsterdam, Netherlands, février 1994.

[Sha73] Robert S. Shankland. – Acoustics of Greek theatres. *Physics Today*, pp. 30–35, octobre 1973.

[Sha97] Erin Shaw. – Hierarchical radiosity for dynamic environments. *Computer Graphics Forum*, vol. 16, n° 2, pp. 107–118, 1997.

[Sil95a] F.X. Sillion. – *Simulation Efficace des Echanges Radiatifs dans des Environnements Complexes*. – Mémoire d'habilitation, Université Joseph Fourier (Grenoble I), décembre 1995.

[Sil95b] F.X. Sillion. – A unified hierarchical algorithm for global illumination with scattering volumes and object clusters. *IEEE Transactions on visualization and computer graphics*, vol. 1, n° 3, septembre 1995.

[SK95] Uwe Stevenson et Ulf Kristiansen. – Pyramidal beam tracing and time dependent radiosity. *Proc. of 15th International Congress on Acoustics*, pp. 657–660, Trondheim, Norway, juin 1995.

[Son69] M. M. Sondhi. – Reconstruction of objects from their sound-diffraction patterns. *Journal of the Acoustical Society of America*, vol. 46, n° 5 (part 2), pp. 1158–1164, 1969.

[SP89] F.X. Sillion et C. Puech. – A general two pass method integrating specular and diffuse reflection. *ACM Computer Graphics, SIGGRAPH'89 Proceedings*, vol. 23, n° 3, juillet 1989.

[SP94] François X. Sillion et C. Puech. – *Radiosity and Global Illumination*. – Morgan Kaufmann Publishers inc., 1994.

[SS93] J.P. Smets-Solanes. – Animated sound textures with multimedia. *Proc. of 4th Eurographics Workshop on Computer Animation and Simulation*, septembre 1993.

[SSSS98] M. Stamminger, H. Schirmacher, Ph. Slusallek et H.P. Seidel. – Getting rid of links in hierarchical radiosity. *EUROGRAPHICS'98*, vol. 17, n° 3, 1998.

[Ste89] Adam Stettner. – *Computer Graphics for Acoustic Simulation and Visualisation*. – Thèse de Masters, Cornell University, janvier 1989.

[Str98] Holger Strauss. – Implementing doppler shifts for virtual auditory environments. *Proc. 104th Audio Engineering Society Convention*, preprint n° 4687 (P4-2), Amsterdam, Netherlands, mai 1998.

[SZEG93] J. Shi, A. Zhang, J. Encarnação et M. Göbel. – A modified radiosity algorithm for integrated visual and auditory rendering. *Computers and Graphics*, vol. 17, n° 6, 1993.

[Tel92] S.J. Teller. – *Visibility Computations in Densely Occluded Polyhedral Environments*. – Thèse de doctorat, Computer Science Division (EECS), University of California, Berkley, 1992.

[TG97a] N. Tsingos et J.D. Gascuel. – A general model for the simulation of room acoustics based on hierarchical radiosity. *technical sketch, in visual proceedings of SIGGRAPH'97, Los Angeles, USA*, août 1997.

[TG97b] Nicolas Tsingos et Jean-Dominique Gascuel. – Soundtracks for computer animation: sound rendering in dynamic environments with occlusions. *Proc. of Graphics Interface'97*, 1997.

[TG98] Nicolas Tsingos et Jean-Dominique Gascuel. – Fast rendering of sound occlusion and diffraction effects for virtual acoustic environments. *Proc. 104th Audio Engineering Society Convention*, preprint n° 4699 (P4-7), Amsterdam, Netherlands, mai 1998.

[TH92] Tapio Takala et James Hahn. – Sound rendering. *ACM Computer Graphics, SIGGRAPH'92 Proceedings*, vol. 28, n° 2, juillet 1992.

[TNKH98] Takashi Takeushi, P.A. Nelson, O. Kirkeby et H. Hamanda. – Influence of individual head related transfer function on the performance of virtual acoustic imaging systems. *Proc. 104th Audio Engineering Society Convention*, preprint n° 4700 (P4-3), Amsterdam, Netherlands, mai 1998.

[TS91] Seth J. Teller et Carlo H. Séquin. – Visibility preprocessing for interactive walkthroughs. *ACM Computer Graphics, SIGGRAPH'91 Proceedings*, vol. 25, n° 4, pp. 61–69, juillet 1991.

[Vau94] R. Vauzelle. – *Un Modèle de Diffraction en 3D dans le Premier Ellipsoide de Fresnel*. – Thèse de doctorat, Université de Poitiers - Laboratoire de signaux et communication, 1994.

[VD94] V.G. Voinov et V.E. Dorofeyev. – Integral equation of convolution type: a new algorithm for analytical deconvolver computing. *Computer Physics Communications*, vol. 81, pp. 351–356, 1994.

[VG97] Eric Veach et Leonidas J. Guibas. – Metropolis light transport. *ACM Computer Graphics, SIGGRAPH'97 Proceedings*, pp. 65–76, août 1997.

[vMM93] D. van Maercke et J. Martin. – The prediction of echograms and impulse responses within the Epidaure software. *Applied Acoustics*, vol. 38, n° 1, pp. 93–114, 1993.

[Vor89] M. Vorländer. – Simulation of the transient and steady-state sound propagation in rooms using a new combined ray-tracing/image-source algorithm. *Journal of the Acoustical Society of America*, vol. 86, n° 1, pp. 172–178, 1989.

[VS97] Alberto Villarreal et John A. Scales. – Distributed three-dimensional finite-diffrence modeling of wave propagation in acoustic media. *Computers in Physics*, vol. 11, n° 4, pp. 388–399, juillet/août 1997.

[WA91] Ren Wentang et Keith Attenborough. – Prediction of sound fields in rooms using statistical energy analysis. *Applied Acoustics*, vol. 34, pp. 207–220, 1991.

[WC76] R.V. Waterhouse et R.K. Cook. – Diffuse sound fields: Eigenmode and free-wave models. *Journal of the Acoustical Society of America*, vol. 59, n° 3, pp. 576–581, mars 1976.

[Wer94a] Josie Wernecke. – *The Inventor Mentor*. – Addison Wesley, 1994.

[Wer94b] Josie Wernecke. – *The Inventor Toolmaker*. – Addison Wesley, 1994.

[WF90] E.M. Wenzel et S.H. Foster. – Realtime digital synthesis of virtual acoustic environments. *Computer Graphics (Proc. ACM Symposium on Interactive 3D Computer Graphics)*, vol. 24, n° 2, pp. 139–140, mars 1990.

[WH91] U. Wilms et R. Heinz. – In-situ measurement of the complex reflection coefficient of the surfaces of walls. *Acustica*, vol. 75, pp. 28–39, 1991. – En allemand.

[Win65] Fritz Winkler. – *Influence de l'Acoustique des Salles sur l'evolution du style musicale*. – Bulletin du Groupe d'Acoustique Musicale n° 11, Université de Paris VI, mars 1965.

[Win70] Fritz Winkler. – *Architecture and Musique Spaciale*. – Bulletin du Groupe d'Acoustique Musicale n° 51, Université de Paris VI, novembre 1970.

[WK89] F. Wightman et D.J. Kistler. – Headphone simulation of free-field listening (part 1 and 2). *Journal of the Acoustical Society of America*, vol. 85, n° 2, pp. 858–878, février 1989.

[Wri95] J.R. Wright. – An exact model of acoustic radiation in enclosed spaces. *Journal of Audio Engineering Society*, vol. 43, n° 10, pp. 813–820, octobre 1995.

[Wri97] J.R. Wright. – Fundamentals of diffraction. *Journal of the Audio Engineering Society*, vol. 45, n° 5, pp. 347–356, mai 1997.

[Zeg97] Eric Zeghers. – *Radiosité spectrale et sources directionnelles : application à l'éclairage des tunnels routiers*. – Thèse de doctorat, Universtité de Rennes 1, SIAMES, IRISA, 1997.

[ZF81] E. Zwicker et R. Feldtkeller. – *Psychoacoustique*. – Masson ed, 1981.

[Zuc89] Irwin Zucker. – Reproducing architectural acoustical effects using digital soundfield processing. *Proc. AES 7th International Conference*, pp. 227–232, 1989.

Index

Résumé

Ce travail porte sur la simulation de champs sonores de haute qualité pour des applications graphiques interactives. Dans ce cadre, nous nous sommes intéressé à trois problèmes : le calcul interactif des effets de l'occultation des ondes sonores par des obstacles, l'intégration du son dans un système d'animation et de réalité virtuelle et la simulation adaptative des réflexions du son dans des environnements réverbérants. Nous présentons une méthode originale permettant d'approcher les effets des obstacles sur la propagation des ondes sonores. Cette méthode qualitative est fondée sur le calcul de l'obstruction des premiers ellipsoïdes de Fresnel. Pour cela nous utilisons le rendu câblé des cartes graphiques spécialisées pour effectuer un calcul interactif entre une source et un récepteur ponctuels dans des environnements généraux. Une extension plus quantitative, basée sur la théorie de Fresnel-Kirchhoff est également détaillée. Nous décrivons également un système interactif de simulation acoustique. Il permet le rendu synchronisé du son et de l'image dans le cadre d'applications d'animation de synthèse et de réalité virtuelle. Nous y avons intégré notre approche de traitement des occultations sonores. Nous présentons comment d'autres effets, comme les réflexions spéculaires du son ou l'effet Doppler sont également pris en compte. Enfin, nous introduisons une technique originale de simulation adaptative fondée sur un formalisme proche de la radiosité hiérarchique utilisée en synthèse d'images. Elle permet de prendre en compte efficacement des réflexions globales spéculaires et diffuses dans le cadre d'échanges énergétiques dépendants du temps. La solution obtenue est indépendante du point d'écoute et ouvre la porte à des applications de parcours interactifs de l'environnement virtuel. En outre, la complexité du processus peut être contrôlée, permettant des applications plus quantitatives, comme la prévision des qualités acoustiques de lieux d'écoute. Ces trois contributions peuvent permettre de réaliser un système de simulation complet d'une scène sonore virtuelle pouvant être utilisable dans une variété d'applications. Celles-ci ne se limitent toutefois pas à l'acoustique, mais peuvent être étendues à la simulation de propagation d'ondes radioélectriques pour la téléphonie mobile ou les réseaux sans fils.

Mots clef : Réalité virtuelle, acoustique virtuelle, multimédia, diffraction, auralisation, simulation adaptative, radiosité hiérarchique, acoustique des salles.

Abstract

This work is aimed at simulating high quality sound fields for interactive 3D graphics applications. We focused our efforts on three main problems: interactive calculation of sound occlusion by obstacles, integration of sound simulation in an interactive 3D animation system and adaptive simulation of sound reflections in reverberant environments. We first present an original method that allows for approximating the effects of obstacles on sound waves propagation. This qualitative method is based on the occlusion of the first Fresnel ellipsoids. We make advantage of the use of 3D graphics hardware to achieve interactive computation rates of the attenuation between a point source and a point receiver in general environments. A more quantitative method based on the Fresnel-Kirchhoff theory of diffraction is also described. Then, we describe an interactive system for integrated sound and graphics rendering in the context of computer animation or virtual reality. This system integrates the previous interactive occlusion rendering technique. We will show how other effects such as sound specular reflections and Doppler shifting are also taken into account. Eventually, we introduce a new adaptive simulation technique based on a hierarchical radiosity-like approach as used in lighting simulations. It allows for taking into account global specular and diffuse reflections in the context of time-varying energy exchanges. The obtained solution is independent of the listening position which makes the approach well suited to walkthrough applications. Moreover, the complexity of the process can be tuned to reach more quantitative results, making it usable for room acoustic quality prediction. These three contributions may allow to design a complete simulation system for rendering a virtual sound scene which could be used in a wide range of applications. These applications, however, do not limit to acoustic simulations but can also be extended to study radiowave propagation in the context of mobile communications or wireless networks.

Keywords: Virtual reality, virtual acoustics, multimedia, diffraction, auralization, adaptive simulation, hierarchical radiosity, room acoustics.