



Mémoire de fin d'étude

Synthèse du champ réverbéré en Audio Orienté Objet :
application au cas de la WFS

Directeur interne : **Alan Blum**

Directeur externe : **Pierre-Olivier Boulant**

Responsable universitaire : **Corsin Vogel**

Rapporteurs : **Laurent Millot** et **Sylvain Lambinet**

Résumé

Ce mémoire propose d'étudier quelques possibilités qu'apporte l'audio orienté objet dans la synthèse de champ réverbéré. On portera un intérêt particulier aux métadonnées de position des objets sonores qui offrent de nouvelles opportunités pour mettre en espace les sons. On décrit différentes approches existantes concernant la réverbération orientée objet, basées sur des méthodes de réverbération bien connues comme la convolution, l'acoustique informatique ou les algorithmes à réseaux de retard bouclés.

La partie pratique de ce travail consiste en la réalisation d'un réverbérateur paramétrique orienté objet, c'est à dire un réverbérateur qui prend en compte les paramètres d'espace des sources et de(s) l'auditeur(s). Ce prototype exploitera la technologie de *Scattering Delay Network*, méthode hybride entre l'acoustique informatique et les réseaux de retards bouclés. La spatialisation de ce réverbérateur sera étudiée dans le cadre spécifique de la Synthèse de Front d'Onde (WFS).

Mots clés Réverbération, Champ Réverbéré, Audio Orienté Objet, Métadonnées, Synthèse de Front d'Onde, WFS, Feedback Delay Network, Scattering Delay Network, Acoustique Informatique.

Table des matières

Introduction	5
1 Le phénomène de réverbération	8
1.1 Définition et phénomènes physiques	8
1.2 La réponse impulsionnelle d'une salle	10
1.3 Dépendance fréquentielle des phénomènes ondulatoires	12
1.4 Perception psychoacoustique de la réverbération	14
1.4.1 Perception des premières réflexions	14
1.4.2 Perception du champ diffus	16
2 L'audio orienté objet	17
2.1 Définition et principales caractéristiques de l'audio orienté objet	17
2.2 Exemple d'utilisation de l'audio orienté objet	18
3 La synthèse de champ réverbéré	20
3.1 Les méthodes acoustiques et mécaniques	21
3.2 La réverbération par convolution	21
3.3 L'acoustique informatique	22
3.4 Les méthodes utilisant des réseaux de retards audionumériques	23
3.4.1 Première initiative de Schroeder	23
3.4.2 Les Feedback Delay Network	25
3.5 Architecture générale d'un réverbérateur numérique et paramétrisation	28
4 Synthétiser le champ réverbéré en audio orienté objet	30
4.1 Une première approche de la réverbération en audio orienté objet : l'approche "multicanale"	30
4.2 L'approche "paramétrique"	32
4.2.1 Méthode basée sur l'analyse d'une réponse impulsionnelle	32
4.2.2 Interpolation de réponses impulsionnelles	35
5 Réalisation de notre réverbérateur orienté objet	36

5.1	Cahiers des charges	36
5.2	Choix d'un logiciel de travail	37
5.3	Présentation du réverbérateur	37
5.4	Choix technologiques	39
5.5	Module de champ diffus	41
5.5.1	Scattering Delay Network	41
5.5.2	Le choix de la matrice de mélange	43
5.5.3	Implémentation pratique du module de champ diffus	44
5.6	Module de premières réflexions	47
6	Attentes relatives à la partie pratique	48
	Conclusion	50

Introduction

L'audio orienté objet connaît un essor particulier depuis quelques années avec un fort développement industriel¹, institutionnel² et open-source³. L'audio orienté objet permet de produire un programme audio unique et de le diffuser facilement sur toutes sortes de système de diffusion (5.1, stéréo, binaural etc) en respectant les intentions artistiques originales grâce à la transmission d'informations de positionnement des sons appelées *métadonnées*. Il offre de nombreuses possibilités techniques et créatives pour créer des espaces sonores immersifs lorsqu'il est utilisé avec des techniques de restitution sonore comme la WFS, l'Ambisonic ou le binaural par exemple.

Si l'audio orienté objet tend à se démocratiser, l'intégration de réverbération dans ce contexte est encore une question ouverte puisqu'il n'existe aucune norme à ce propos et que nombreuses approches sont envisageables, résultant en une grande diversité de pratique chez les professionnels du son "immersif". En effet, la réverbération est un phénomène complexe qui met en jeu de nombreux phénomènes physiques et psychoacoustiques et l'audio orienté objet apporte de nouvelles opportunités dans l'utilisation et la synthèse de champ réverbéré. On étudiera ainsi l'intérêt des métadonnées de position des objets ainsi que les possibilités de spatialisation qu'offrent des techniques de sonorisation orientées objet.

Des initiatives de réverbération en audio orienté objet ont déjà été proposées. La plupart de ces méthodes reposent sur l'analyse d'une réponse impulsionnelle pour ensuite synthétiser un champ réverbéré adapté au système de diffusion. Deux des points faibles de ces méthodes sont la difficulté à rendre les mouvements des sources au sein de l'espace et la "rigidité" de la réponse impulsionnelle qui ne décrit qu'une seule position au sein de la salle.

1. Les constructeurs de matériel de sonorisation proposent presque tous un système de diffusion orienté objet comme *L-Acoustics* avec son processeur *L-ISA*, *d&b* avec *Soundscape*, *Amadeus* avec son système *Holophonix* ou *Sonic Emotion* avec le *Sonic Wave I*.

2. Des normes comme l'Audio Definition Model (ADM) de l'ITU ou le MPEG-H 3D Audio ont été mises en place pour standardiser la pratique de l'audio orienté objet.

3. De nombreuses initiatives open-sources existent telles que *WFS Collider*, *WONDER*, *Soundscape Renderer* ou *WFS DIY* pour n'évoquer que de la WFS.

L'enjeu pratique de ce mémoire est la réalisation d'un réverbérateur orienté objet qui permet la prise en compte des positions des sources dans l'espace ainsi que leur éventuels déplacements. Sans prétendre proposer un logiciel capable de fonctionner en conditions réelles, il s'agira de réaliser d'un prototype visant à faire la démonstration de différents résultats de ce mémoire. On s'appuiera pour cela sur deux méthodes de synthèse de champ réverbéré numérique bien connues que sont les algorithmes de réverbération à réseaux de retards bouclés (*Feedback Delay Network*) et l'acoustique informatique. Le réverbérateur utilisera une technologie hybride entre ces deux méthodes appelée *Scattering Delay Network* qui permet de répondre aux différents enjeux cités précédemment.

Le modèle proposé dans ce mémoire ne dépendra pas dans son principe d'un système de reproduction particulier. Pour la partie pratique du mémoire et la démonstration du prototype on choisira cependant la WFS (*Wave Field Synthesis*), car c'est une des techniques de sonorisation "immersive" les plus populaires dans le milieu du spectacle vivant. Un des enjeux de la partie pratique sera de tirer profit des possibilités de spatialisations spécifiques à la WFS pour placer nos points de réverbération de façon la plus pertinente possible.

Notices préalables

Afin de faciliter la compréhension de ce mémoire et de fluidifier sa lecture, quelques remarques d'ordre lexical sont nécessaires. L'essentiel de la littérature sur la réverbération et l'audio orienté objet étant en anglais, on utilisera les termes anglophones lorsqu'il n'existe pas de traduction évidente et/ou usuelle. On précisera toujours la signification d'un sigle ainsi que son éventuelle traduction lors de la première occurrence du terme. On précise dès maintenant la signification des quatre sigles les plus couramment utilisés au cours de ce mémoire :

- OBA pour *Object-Based Audio* ou *Audio Orienté Object*.
- WFS pour *Wave Field Synthesis* ou *Synthèse de Front d'Onde*.
- FDN pour *Feedback Delay Network* que l'on pourrait traduire par *Réseau de Retards Bouclés*.
- SDN pour *Scattering Delay Network* que l'on pourrait traduire par *Réseau de Retards Dispersés*.

1 Le phénomène de réverbération

1.1 Définition et phénomènes physiques de la réverbération

La réverbération correspond à la persistance d'un son dans un espace clos ou semi-clos après interruption de la source sonore. Elle représente la "signature acoustique" d'un lieu qui rend immédiatement identifiable à l'oreille la sonorité d'une église, d'un tunnel ou d'une salle de bain. Elle est le résultat de l'interaction de la source sonore avec le milieu dans laquelle elle est plongée.

L'explication la plus naïve du phénomène de réverbération est la superposition d'un grand nombre d'échos résultants des multiples "rebonds" du son sur les obstacles environnants.

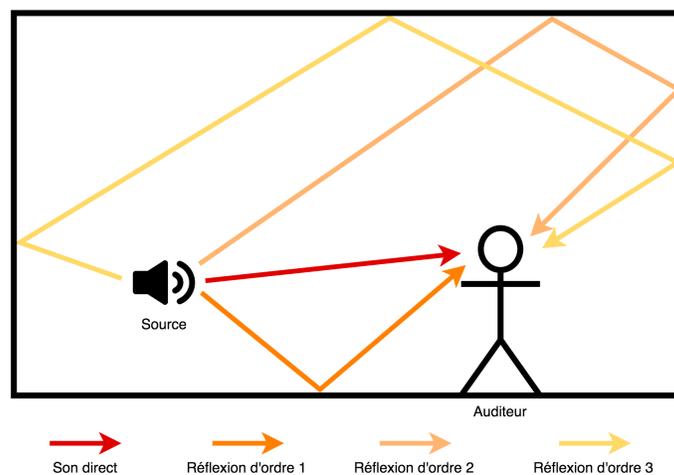


FIGURE 1 – Exemple en 2D de quelques trajets du son possibles

La partie du son qui suit le trajet rectiligne de la source vers l'auditeur est appelé de *champ direct* et correspond à la partie non-réverbérante du son. Les "rebonds" du son sur les obstacles correspondent au phénomène physique de *réflexion* des ondes sonores. L'ordre d'une réflexion quantifie le nombre de "rebonds" qu'elle a subit avant de parvenir à l'auditeur.

Si la réflexion est le phénomène qui permet de comprendre le plus facilement la réverbération, les ondes sonores subissent au sein d'une salle de nombreux phénomènes physiques ondulatoires [1] :

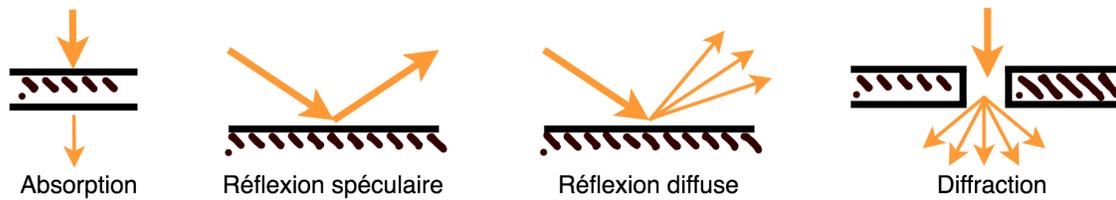


FIGURE 2 – Différents phénomènes physiques subit par une onde sonore au contact d'un obstacle.

- Un phénomène d'absorption qui se traduit par une diminution de l'intensité sonore lors de la traversée d'un obstacle. Ce phénomène est dû à la transformation de l'énergie acoustique en énergie thermique par frottement avec le milieu rencontré. L'air absorbe lui aussi les ondes sonores, cela se traduit par une atténuation du son proportionnelle au carré de la distance parcourue.
- Des réflexions sur les obstacles rencontrés. Elles sont dites spéculaires lorsque le son "rebondit" dans une unique direction, prévue par la loi de Snell-Descartes. En fonction des longueurs d'onde mises en jeu et des aspérités de l'obstacle, le son peut repartir dans de multiples directions, ces réflexions sont alors dites diffuses.
- Au contact d'une discontinuité sur un obstacle tel qu'un trou ou une arrête, le son peut subir de la diffraction, c'est à dire être envoyé dans de multiples directions à partir de cette discontinuité.

Des effets atmosphériques peuvent également influencer la trajectoire du son. Par exemple, le taux d'humidité de l'air et la température affectent la célérité du son et un gradient de température dans l'air vient courber sa trajectoire de propagation. Nous ne considérerons pas ces effets dans le cadre de notre travail car leurs effets ne sont significatifs que pour de très grandes distances (dans le cadre de grands concerts en plein air par exemple), ce qui est rarement le cas concernant l'acoustique des salles.

Le terme de "diffusion" est fréquemment employé par les professionnels du son (on retrouve d'ailleurs très souvent un paramètre de "diffusion" sur les réverbérateurs). Ce terme renvoie à la notion de physique ondulatoire du même nom qui décrit la déviation par un obstacle d'une onde dans de multiples directions. Une bonne diffusion sonore renvoie à une répartition

homogène et "aléatoire" du champ réverbéré dans une salle et est perçue comme quelque chose de positif. On utilise d'ailleurs des *diffuseurs* dans les studios pour répartir le son dans la pièce et éviter des phénomènes gênants comme des effets de *flutter*⁴ [2].

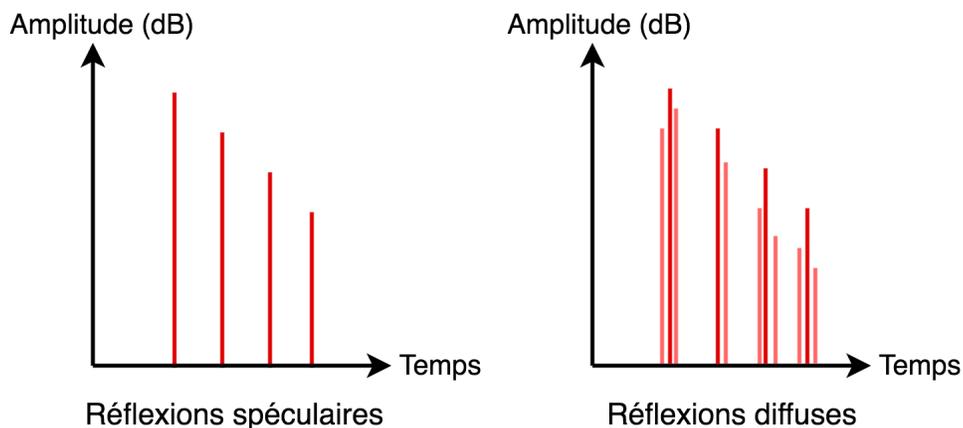


FIGURE 3 – Exemple de diffusion sur des premières réflexions.



FIGURE 4 – Exemple d'un diffuseur de Schroeder bi-dimensionnel. Les différences de niveau entre les éléments favorisent le phénomène de diffraction pour les longueurs d'onde concernées.

1.2 La réponse impulsionnelle d'une salle

Pour mieux comprendre le phénomène de réverbération, on peut étudier la réponse impulsionnelle d'une salle, appelée *RIR* pour *Room Impulse Response* dans la littérature.

4. on parle de *flutter* ou d'écho flottant lorsque le son rebondit de nombreuses fois entre deux surfaces parallèles proches. Cet effet est particulièrement présent dans des petites pièces ou des couloirs et colore le son d'une manière considérée le plus souvent comme désagréable.

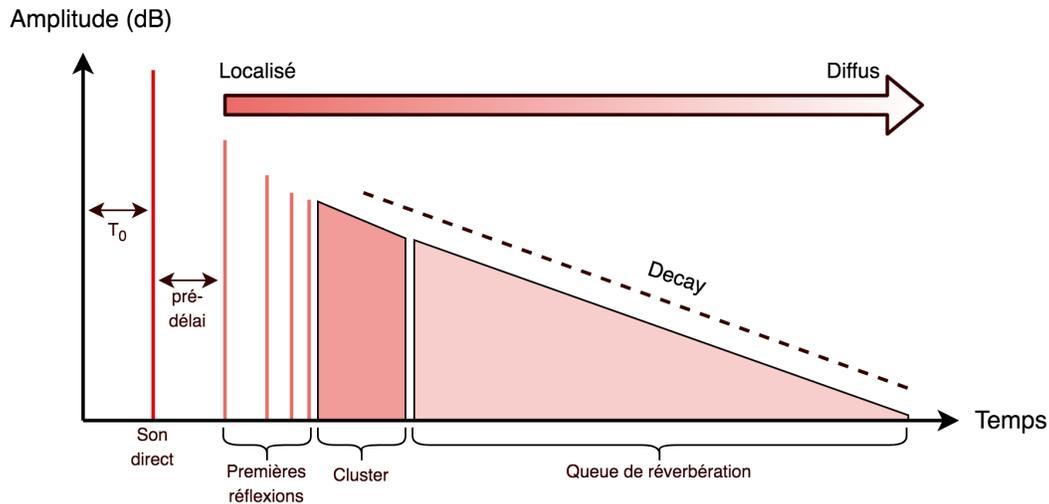


FIGURE 5 – Réponse impulsionnelle générique d’une salle.

On sépare généralement la réponse impulsionnelle d’une salle, qui vont du plus localisé au plus diffus :

- Les *premières réflexions*, très identifiables et peu nombreuses qui correspondent aux réflexions du premier ordre sur les murs, le sol ou le plafond. Elles parviennent chacune à l’auditeur avec une direction bien précise. Elles jouent un rôle important dans l’estimation de la taille de la pièce ainsi que l’estimation de la distance à la source.
- Le *cluster* qui contient les réflexions d’ordre supérieur, la densité d’échos commence à augmenter et les réflexions individuelles deviennent difficiles à identifier.
- La *queue de réverbération* pour laquelle le son a subit tellement de réflexion et de diffusion que les réflexions individuelles ne sont plus identifiables. Le son ne provient plus d’une direction précise mais adopte un comportement statistique. La queue de réverbération suit une décroissance (*Decay*) exponentielle.

On peut dégager plusieurs informations importantes d’une réponse impulsionnelle de salle :

- Le *pré-délai* qui correspond à la différence de temps entre le son direct et la première réflexion. Ce pré-délai est dû à la différence de marche entre le son direct et la première réflexion, c’est à dire l’écart de distance parcourue par ses deux ondes. Il donne un indice sur la

- taille de la salle ainsi que sur la distance entre la source et l'auditeur.
- La durée de la réverbération, mesurée le plus souvent via le T_{60} qui correspond au temps nécessaire pour que le niveau de la réverbération diminue de 60 dB par rapport au son direct.

1.3 Dépendance fréquentielle des phénomènes physiques ondulatoires

En plus d'interagir avec la salle, les ondes sonores peuvent interagir avec d'autres ondes sonores sous la forme d'*interférences*. En particulier les diverses réflexions d'un même son peuvent interagir pour former des motifs d'interférences complexes. Ces interférences se traduisent par un phénomène de *filtrage en peigne* des diverses réflexions et du son direct. Le filtrage en peigne est considéré le plus souvent comme indésirable par l'ingénieur du son (on parle de "problème de phase") mais il fait partie intégrante du phénomène de réverbération. Cet effet est particulièrement marqué pour les premières réflexions qui viennent se superposer au son direct et participe à définir la couleur de la pièce. En ce qui concerne la queue de réverbération, la densité d'échos étant très importante, le filtrage en peigne n'est pas vraiment perçu comme tel mais il altère la réponse en phase et favorise la sensation de diffusion.

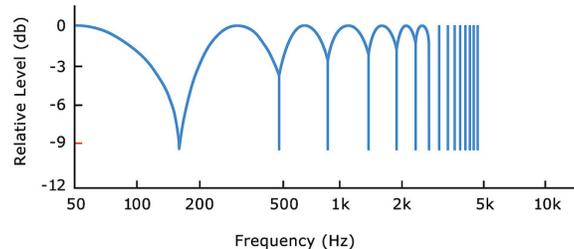


FIGURE 6 – Réponse en fréquence d'un filtre en peigne.

Le son est un phénomène physique ondulatoire qui met en jeu une large plage de longueurs d'onde qui vont de la dizaine de mètres pour les basses fréquences au centimètre pour les très hautes fréquences. Or les différents phénomènes physiques que nous venons de citer dépendent de la longueur d'onde du son considéré et de la longueur caractéristique des obstacles rencontrés. Par exemple, un mur de 20 cm d'épaisseur n'est pas suffisant pour atténuer les basses fréquences, car cet obstacle est trop petit en com-

paraison avec les longueurs d'onde mises en jeu (de l'ordre de plusieurs mètres). Pour cette raison on peut séparer le comportement des ondes sonores dans une salle en deux catégories définies par la fréquence de Schroeder $f_{Schroeder}$:

- Les fréquences supérieures à $f_{Schroeder}$, pour lesquelles les phénomènes de réflexion et de diffraction seront assez significatifs pour diffuser le son dans la salle. On parle alors de mode réverbérant.
- Pour les fréquences inférieures à $f_{Schroeder}$ pour lesquelles leur longueur d'onde est de l'ordre des dimensions de la pièce, on observe un comportement dit *nodal*. Ces ondes vont "passer au travers" et donc ignorer la plupart des obstacles. Elles vont former des ondes stationnaires et présenter des ventres et des noeuds de vibration dans l'espace.

Schroeder propose une formule pour estimer cette fréquence limite :

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

Où T_{60} désigne le temps de réverbération de la salle et V son volume. Lors de la conception d'une salle dédiée au son, on désire abaisser au plus possible la fréquence de Schroeder afin d'être en mode réverbérant sur une plage de fréquence la plus large possible et repousser les ondes stationnaires dans le bas du spectre. La formule ci-dessus nous permet de retrouver deux principes bien connus par les ingénieurs du son :

- Les petites salles posent plus de problème que les grandes. En effet les espaces de petit volume, de part leurs dimensions, présentent un comportement nodal dans le bas du spectre et dans le bas médium.
- Il faut une salle relativement absorbante afin de limiter son T_{60} . C'est pour cela que l'on retrouve des matériaux absorbants dans de nombreuses régies de mixage ou cabines de prise de son par exemple. Cependant une pièce trop absorbante telle qu'une chambre anéchoïque peut se révéler peu naturelle voire désagréable à l'écoute. A ce sujet, on peut penser au concept de *Live End Dead End*, assez populaire pour le design de régie de mixage, qui consiste à rendre très absorbante l'avant de la pièce tout en laissant l'arrière réfléchissant dans

le but d'obtenir un rendu naturel mais contrôlé dans les basses fréquences.

La réverbération est donc un phénomène complexe car c'est un phénomène tri-dimensionnel impliquant de nombreux phénomènes physiques ondulatoires eux même dépendants de la fréquence du signal considéré.

On parle souvent par abus de langage de "la réverbération d'une salle" car on considère en première approximation le champ réverbéré homogène dans l'espace. En réalité la réverbération dépend du placement de la source et du placement de l'auditeur. Le changement d'une de ces positions entraîne une modification de la géométrie des premières réflexions et donc de la réverbération dans son ensemble. Cette notion de double dépendance de position (source et auditeur) est essentielle pour la développement d'un réverbérateur orienté objet puisque en connaissant la position des sources et éventuellement de l'auditeur via les métadonnées de position, on est en mesure de prendre en compte ces subtilités dans le champ réverbéré.

1.4 Perception psychoacoustique de la réverbération

Le système auditif est fait pour capter la voix et les informations sonores importantes comme le bruit d'un danger imminent. De ce fait le cerveau humain est doué dans le fait d'isoler les informations sonores importantes en ignorant celles de second plan comme le bruit de fond ou la réverbération. On peut penser à ce sujet à l'effet "cocktail party" qui désigne la capacité de notre cerveau à isoler la voix de notre interlocuteur parmi un flux très dense d'autres conversations.

Le système auditif n'est donc à priori pas fait pour analyser la réverbération des lieux qui nous entoure, il montre d'ailleurs vite des limites dans cet exercice. La réverbération peut même se révéler comme néfaste pour la compréhension de la parole, cependant notre cerveau est capable de tirer certaines informations de spatialisation des signaux réverbérés qu'il reçoit.

1.4.1 Perception des premières réflexions

Les premières réflexions sont très corrélées au son direct et parviennent à l'auditeur juste après celui-ci avec des directions différentes. Cela donne lieu

à un *effet de précedence* qui regroupe plusieurs phénomènes de perception et de localisation d'un son accompagnés de ses réflexions [3]. L'intensité de l'effet de précedence dépend fortement du pré-délai ainsi que du type de son concerné.

De nombreuses études ont été menées avec un son direct et une réflexion unique simulés via des enceintes délayées dont voici les conclusions générales :

- Pour des valeurs de pré-délai extrêmement faibles (inférieures à 1 ms), les deux sons sont perçus comme un événement sonore unique provenant d'une direction moyenne entre les deux sons. On parle alors de sommation de localisation (*summing localization*).
- Pour des valeurs de pré-délai faibles (entre 1 et 5 ms), les deux sons sont perçus comme un événement sonore unique localisé sur le premier son. On parle alors de *loi du premier front d'onde* car c'est le premier son qui parvient à l'auditeur qui domine la localisation perçue par l'auditeur.
- Lorsque le pré-délai augmente et franchi le *seuil d'écho*, les deux événements sonores sont perçus indépendamment sous la forme d'un écho. La valeur de ce seuil peut être très faible pour des sons très brefs et percussifs (jusqu'à 2 ms d'après certaines études) ou assez grande pour des sons résonnants (jusqu'à 50 ms).

Les études soulignent également le fait que même si les deux sons sont perçus comme un événement sonore unique, cette perception n'est pas équivalente à celle du son direct seul. La sommation du son direct et de ses réflexions a un effet sur le timbre (filtrage en peigne), sur l'enveloppe du son (superposition de plusieurs transitoires qui viennent "lisser" l'enveloppe) et peuvent rendre la source perceptivement plus large. Ces divers effets facilitent la compréhension de l'espace acoustique par notre cerveau [4].

Dans le cadre de salle de taille usuelle, la différence de marche entre le son direct et les premières réflexions les placent très souvent dans le cadre de la loi du premier front d'onde (50 ms de délai, le seuil d'écho le plus haut correspond à une différence de marche de 17 mètres). Des réflexions avec un retard supérieur à 50 ms dégradent fortement le timbre et diminuent l'intelligibilité du son. Un excès d'énergie dans cette plage de délai est très

souvent considéré comme néfaste par l'ingénieur du son.

1.4.2 Perception du champ diffus

Le champ diffus est par définition moins localisé, notre cerveau n'arrive plus à discerner les réflexions individuelles ce qui limite sa capacité à en tirer des informations. Ainsi la queue de réverbération donne peu d'informations de localisation mais permet d'évaluer la taille et les propriétés de l'espace.

La perception du champ diffus est d'ailleurs inhibée par la présence de son direct et sa perception redevient maximale environ 150 ms après l'arrêt de celui-ci [5].

Si le champ diffus se révèle assez peu important dans la localisation du son, il a une grande implication esthétique dans le travail de mixage. Les deux critères les plus importants concernant le champ diffus dans le choix d'une réverbération, outre le rapport champ direct sur champ diffus sont :

- son temps de réverbération, souvent signifiée via son T_{60} .
- sa couleur spectrale

Ces deux critères doivent être adaptés au style musical ou à l'image.

Il a été montré qu'un champ diffus peut être reproduit en WFS par la superposition 8 à 10 ondes planes [6] [7]. Cette superposition d'onde plane permet d'assurer la condition d'isotropie du champ diffus, c'est à dire qu'il présente les mêmes propriétés dans toutes les directions.

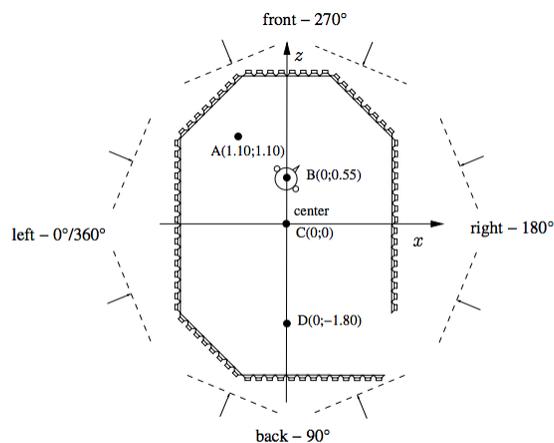


FIGURE 7 – L’auditeur est placé à l’intérieur d’un anneau WFS de 104 enceintes et le champ diffus est reproduit via 8 ondes planes, issu de [7]

Si le sujet est autorisé à se déplacer, les rotations du champ diffus peuvent être entendus est il faut alors augmenter le nombre d’ondes planes jusqu’à 14.

2 L’audio orienté objet

2.1 Définition et principales caractéristiques de l’audio orienté objet

On appelle "objet audio" l’association d’un fichier audio et d’un jeu de paramètres appelés *métadonnées* qui décrivent la façon dont ce son doit être reproduit lors de la diffusion d’un programme. Les métadonnées les plus importantes en terme de spatialisation sont les paramètres de position 2D ou 3D de chaque son mais elles peuvent aussi contenir des informations concernant sa largeur ou son diagramme de directivité par exemple. Ces métadonnées sont interprétées par un *moteur de rendu* ("*renderer*") qui fourni en temps réel à chaque enceinte un flux audio adapté au système de diffusion qui lui a été spécifié. L’enjeu est de restituer l’oeuvre la plus conforme à ce qui a été décidé lors de la phase de *mixage objet*. On appelle audio orienté objet l’ensemble des pratiques qui mettent en jeu cette notion d’objet audio.

Parmi les avantages qu’apporte l’audio orienté objet, on peut citer :

- Le fait de pouvoir diffuser une oeuvre sur un grand nombre de systèmes de reproduction (5.1, Atmos, binaural etc) sans avoir à réaliser un mixage différent pour chacun de ces systèmes.
- La possibilité d'offrir plus d'adaptabilité à l'utilisateur. Comme le rendu se fait en temps réel par le moteur de rendu, l'utilisateur peut intervenir sur certains paramètres. On peut penser à augmenter le niveau des dialogues d'un film afin de les rendre plus compréhensibles dans un environnement bruyant ou pour des personnes malentendantes par exemple.
- Faciliter le travail de création et de mixage pour des oeuvres dans des formats "immersifs" tels que l'ambisonie ou la WFS pour lesquels la configuration de restitution est très fortement variable (disposition et nombre d'enceintes par exemple).

Ces avantages présentent un intérêt en terme de méthode de travail et de coût de production dans des domaines très variés de l'audiovisuel : cinéma, *broadcasting*, spectacle vivant "immersif", jeu-vidéo, réalité virtuelle etc.

L'audio orienté objet apporte également ses contraintes :

- Il faut avoir tous les fichiers audios qui composent l'oeuvre, ce qui peut nécessiter beaucoup d'espace de stockage. Pour limiter ce problème on peut envisager de travailler avec des *stems* ou de compresser les fichiers audios.
- Le mixage objet est en général réalisé sur un seul dispositif de reproduction. Il faut donc "faire confiance" au moteur de rendu concernant le rendu sonore sur d'autres dispositifs. On peut vérifier son travail sur d'autres systèmes de diffusion en les simulant via des empreintes HRTF avec les limites que cela comporte.

2.2 Exemple d'utilisation de l'audio orienté objet

Le *Dolby Atmos* est le format orienté objet le plus populaire dans le domaine du cinéma. Examinons le logiciel *Dolby Atmos Renderer* qui permet le mixage orienté objet de films pour y identifier les points clés d'un outil de travail orienté objet.

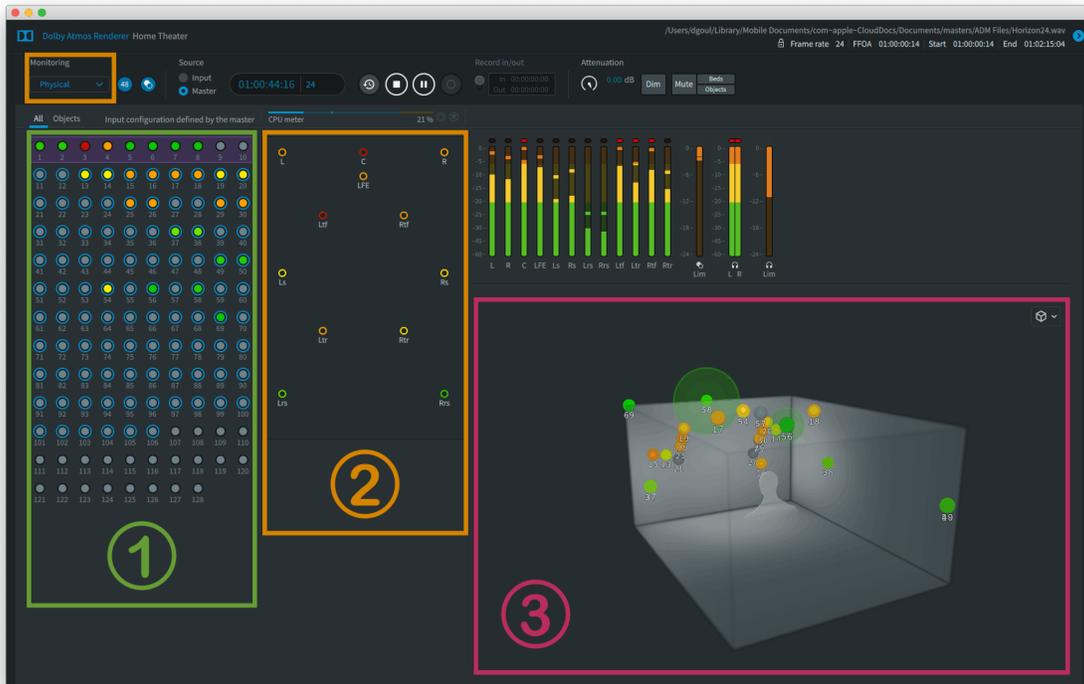


FIGURE 8 – Vue d’ensemble du logiciel *Dolby Atmos Renderer*, issu de [8].

On peut découper cette image en trois zones qui remplissent chacune un rôle essentiel pour travailler en audio orienté objet :

- Zone 1 : configuration des objets. Chaque objet est relié à une piste dans la station de travail audio numérique qui sert à lire les sons ou directement à une piste de console.
- Zone 2 : configuration du système de reproduction. On y définit et on y place les enceintes dans l’espace pour décrire le système de diffusion de la salle d’écoute. On peut changer ce système très rapidement pour vérifier le mixage sur un système plus petit comme un 5.1 ou un 7.1 par exemple (*down-mixing*).
- Zone 3 : Visualisation générale du placement des sources dans l’espace 3D.

Pour chaque objet, on a accès à un outil de panoramique (*Panner*) qui permet de renseigner plusieurs métadonnées dont sa position.

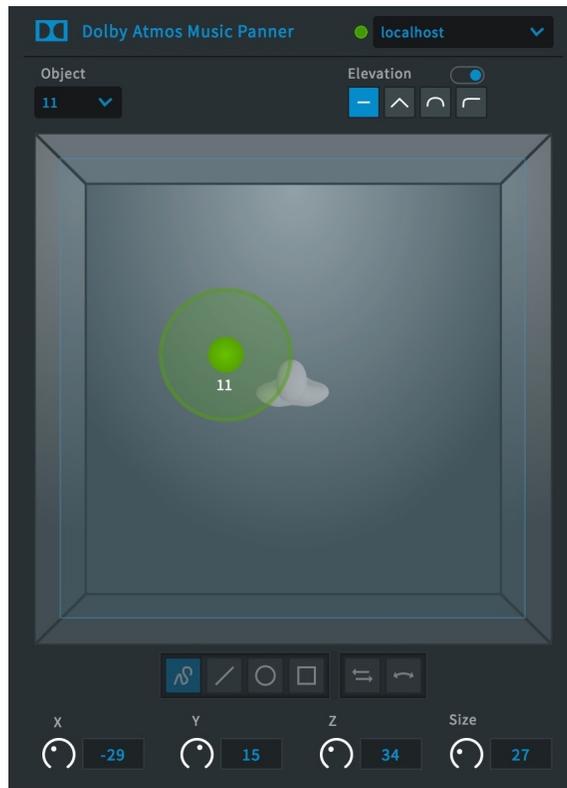


FIGURE 9 – Outil de panoramique du *Dolby Atmos*

L’outil de panoramique permet de placer précisément les sources dans l’espace via une interface graphique ou des coordonnées de position. On peut également choisir d’autres attributs pour l’objet comme sa taille pour le faire apparaître plus ou moins ponctuel dans l’espace. Tous ces paramètres peuvent être modifiés en temps réel ou automatisés dans la station de travail.

3 La synthèse de champ réverbéré

Pour élaborer notre prototype de réverbération objet, il est indispensable de passer en revue les principales méthodes de synthèse de champ réverbéré et d’estimer laquelle ou lesquelles seraient adaptées pour notre travail.

3.1 Les méthodes acoustiques et mécaniques

La méthode la plus ancienne utilisée pour ajouter de la réverbération en post-production consiste à enregistrer la champ réverbéré d'une pièce existante en y plaçant un microphone qui capte le son diffusé par une enceinte.

Les méthodes mécaniques incluent les réverbérations à plaque et à ressort (plus connues sous le nom de *Plate Reverb* et de *Spring Reverb*). Dans ces systèmes, c'est le son qui vient mettre en vibration un objet mécanique (la plaque ou les ressorts) via un transducteur. Cette vibration est ensuite captée par un autre transducteur. Différents phénomènes physiques ondulatoires comme des effets de bords sur la plaque et dans les ressorts contribuent à rendre le son en sortie de ces dispositifs plus diffus. Ces réverbérations produisent un son très caractéristique et sont très appréciées dans certains styles musicaux. Pour cette raison, on retrouve souvent des émulations de réverbération à plaque ou à ressort dans les réverbérateurs numériques.

Ces différentes méthodes ne sont pas assez modulables et nécessitent des dispositifs trop complexes pour être intéressantes dans la réalisation d'une réverbération orientée objet. On se tourne donc vers les méthodes de synthèse de champ réverbéré numériques, qui sont beaucoup plus paramétrables et simples à mettre en oeuvre.

3.2 La réverbération par convolution

La réverbération à convolution est basée sur l'opération mathématique du même nom : en convoluant un signal sonore avec la réponse impulsionnelle d'une salle dont on souhaite reproduire l'acoustique, on arrive à le replacer "dans la salle" désirée. Cette méthode produit des résultats très réalistes et est particulièrement appréciée des mixeurs cinémas qui trouvent une grande variété de pièces et d'espaces en explorant les banques de réponses impulsionnelles proposées par les constructeurs.

Cette méthode suppose d'avoir des échantillons audios qui serviront "d'empreintes sonores" pour chaque pièce que l'on désire simuler, les ré-

ponses impulsionnelles. On doit donc enregistrer la réponse impulsionnelle de chaque salle, à l'aide d'un ou plusieurs micros et d'une source sonore qui peut être soit une impulsion via un ballon qui éclate ou un pistolet à blanc, soit un balayage sinusoïdal logarithmique ("*sweep*") via une enceinte omnidirectionnelle. On parle très souvent par abus de langage de "la réponse impulsionnelle d'une salle", en réalité, il existe une infinité de réponses impulsionnelles pour une même salle car elle dépend de 3 paramètres :

- le format de captation : une réponse impulsionnelle peut être mono-phonique, stéréophonique, binaurale, 5.1 etc⁵.
- la position du(es) microphone(s) dans la pièce.
- la position de la source sonore dans la pièce.

L'opération de convolution mobilise une puissance de calcul conséquente et introduit de la latence. De nombreuses optimisations algorithmiques ont été réalisées pour limiter ces deux problèmes [9].

3.3 L'acoustique informatique

Comme nous l'évoquions dans le premier chapitre, les phénomènes physiques à l'origine la réverbération sont bien connus : réflexion, absorption, diffraction etc. Il semble alors naturel de chercher à émuler informatiquement ces phénomènes pour synthétiser une réverbération, c'est l'enjeu de *l'acoustique informatique*. Ces algorithmes ont besoin de connaître la géométrie de la pièce dont on cherche à synthétiser la réverbération, géométrie que l'on peut renseigner via un plan $3D$ de l'espace par exemple.

Ces méthodes permettent d'obtenir des résultats réalistes car elles modélisent vraiment le comportement du son dans une pièce. Cependant elles tendent à présenter plusieurs problèmes.

Tout d'abord, la qualité sonore de la réverbération dépend de la finesse de la description de la pièce considérée, il faut en premier lieu en décrire la géométrie pour que les trajets des multiples réflexions soient corrects ainsi que les différents types de matériaux présents sur les surfaces pour modéli-

5. Par abus de langage, on parle souvent de "la réponse impulsionnelle d'une salle" même quand elle est multicanale alors que l'on devrait plutôt parler d'un jeu de réponses impulsionnelles spécifique à la configuration microphonique choisie

ser les phénomènes d’absorption par bande de fréquence. Cette description peut s’avérer fastidieuse.

Le calcul des trajectoires et des phénomènes physiques ondulatoires mène à une charge de calcul très importante, à tel point qu’un rendu en temps réel commence tout juste à être envisageable [9]. On risque donc de mobiliser beaucoup de puissance de calcul pour obtenir un champ diffus très juste physiquement sans que notre cerveau soit en mesure de percevoir ces détails, ce qui revient à gâcher ces ressources.

Il est à noter que ces méthodes d’acoustique informatique peuvent être utilisées pour générer des réponses impulsionnelles de salle que l’on utilisera par la suite dans un réverbérateur à convolution standard. La génération de cette réponse n’est alors plus tributaire du temps réel et on peut alors pousser la précision de la modélisation au maximum. Cela permet par exemple d’extraire la réponse impulsionnelle d’un bâtiment modélisé en 3D, qu’il soit réel ou non.

3.4 Les méthodes utilisant des réseaux de retards numériques

3.4.1 Première initiative de Schroeder

Les premiers algorithmes basés sur des *réseaux de retards* ont été proposés par Schroeder dès le début des années 60. Dans sa publication [10], il propose des critères théoriques et une implémentation pratique qui auront une grande importance dans le développement des réverbérations numériques. Il considère la réverbération dans sa plus simple définition : la superposition d’un grand nombre d’échos. Schroeder choisit 1000 échos par secondes (*densité d’écho*) comme valeur minimale pour reproduire le comportement d’une pièce. L’idée est donc de simuler la réflexion et la diffusion du son dans un espace clos via un agencement de 2 modules élémentaires de traitement du signal : les filtres passe-tout (*All-Pass Filters* ou *APF*) et les filtres en peignes récursifs (*Feedback Comb Filters* ou *FBCF*).

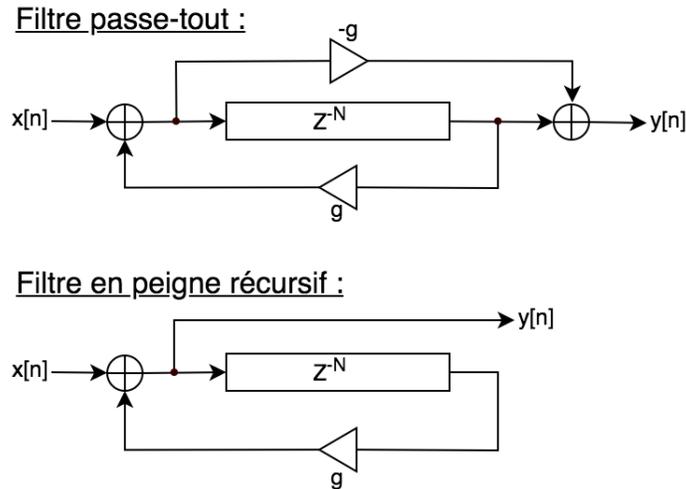


FIGURE 10 – Exemple de filtre passe-tout et de filtre en peigne récursif avec un retard de N échantillons et un gain g tel que $|g| < 1$.

Schroeder propose un agencement de trois filtres passe-tout en série qui rentre dans quatre filtres en peigne récursifs en parallèle comme l'indique la figure suivante :

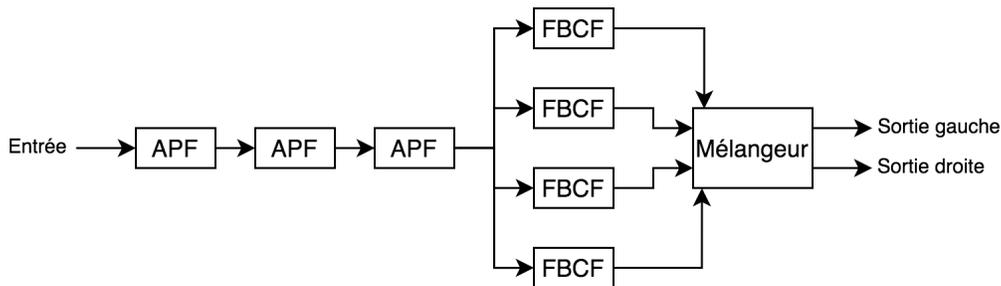


FIGURE 11 – Réverbérateur de Schroeder stéréophonique, inspiré de [11].

Les filtres en peigne récursifs simulent les réflexions du son entre deux parois opposées d'une pièce. C'est l'association en parallèle de ces filtres qui définissent la "pièce" que l'on souhaite modéliser. En particulier les différents retards sur ces filtres définissent la taille ressentie de la pièce et les gains son temps de réverbération.

Les filtres passe-tout modifient la réponse en phase d'un son et le rendent plus perceptivement plus "diffus" afin de simuler les multiples diffractions et diffusions que subit le son sur les murs et les objets dans une pièce. L'association en série permet de multiplier la densité d'écho de chaque filtre par celle des autres pour se rapprocher du seuil des 1000 échos par

seconde. Les filtres passe-tout peuvent cependant introduire des résonances et des effets de *flutter* donnant une couleur métallique indésirable au son. L'agencement et le réglage des temps et des gains des filtres passe-tout doivent donc être bien maîtrisés par le constructeur de la réverbération [12].

Enfin Schroeder suggère une matrice de mélange qui à partir des quatre sorties des filtres en peigne permet de construire plusieurs sorties les plus décorréélées possibles pour alimenter plusieurs enceintes.

La puissance de calcul nécessaire pour un réverbérateur de Schroeder est quasiment négligeable sur du matériel actuel. Si la version proposée par Schroeder lui même sonne relativement artificielle et métallique, des agencements plus complexes basées sur son principe offre une qualité sonore honorable. Cette première proposition aura un impact considérable sur l'architecture des futurs réverbérateurs numériques.

3.4.2 Les Feedback Delay Network

Dans le sillage de Schroeder et avec le développement de la puissance de calcul des ordinateurs, des ingénieurs ont cherché à améliorer les réverbérateurs algorithmiques.

Une des limitations des réverbérateurs de Schroeder est le manque d'interconnexion des éléments entre eux. Pour régler ce problème, l'idée d'utiliser une matrice de mélange au sein d'une boucle de rétroaction est apparue. Ces réverbérateurs ont été généralisés sous l'appellation *Feedback Delay Network*, abrégé en **FDN**. Ils sont à la base de nombreux processeurs de réverbération matériels et logiciels, c'est par exemple le cas de l'algorithme de réverbération du Spat de l'IRCAM [13][14] ou de réverbérateurs numériques bien connus comme le *Lexicon L224XL*.

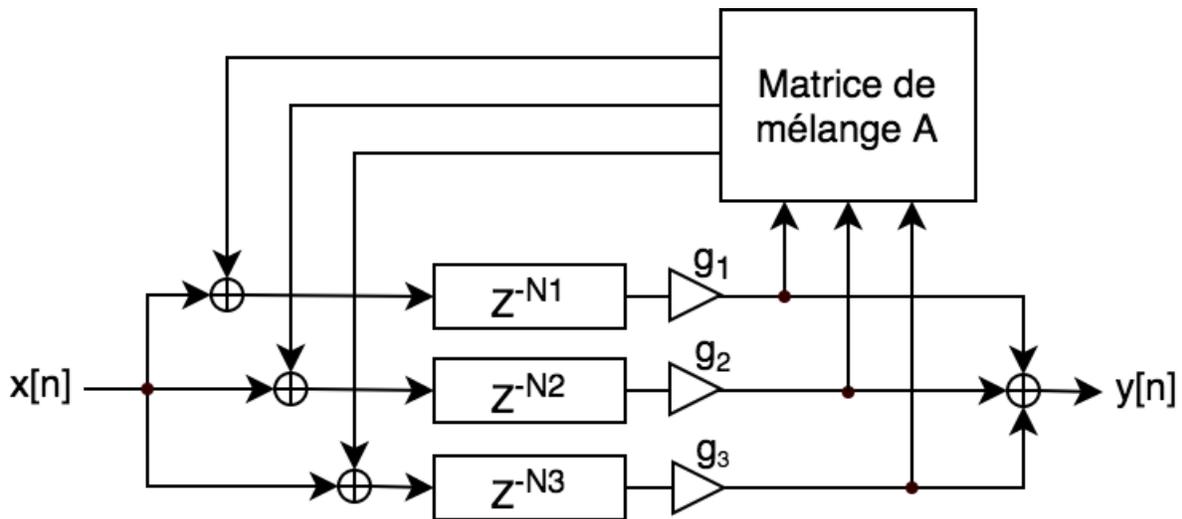


FIGURE 12 – Topologie simplifiée d'un FDN d'ordre 3.

Le signal d'entrée passe par n retards N_i mis en parallèle, la sortie de chaque retard est envoyée dans une matrice de mélange de taille $n \times n$ notée A qui réinjecte une combinaison des sorties dans chaque entrée des retards. Le nombre de retard n est appelé *l'ordre* du réseau, plus l'ordre est élevé, plus la densité d'écho augmente. Des réseaux d'ordre 8 jusqu'à 16 sont assez répandus et offre un bon compromis entre qualité sonore et puissance de calcul mobilisée.

La matrice de mélange A joue un grand rôle dans le comportement de la réverbération car c'est elle qui assure la bonne distribution de l'énergie dans le réseau. Des critères mathématiques concernant cette matrice ont été trouvés pour garantir la stabilité du réverbérateur [15].

L'un des enjeux dans la conception de ces réverbérateurs est leur *paramétrisation*. Il faut en effet proposer des paramètres de haut-niveau pour l'utilisateur, afin que ce dernier puisse rapidement ajuster la réverbération à ses besoins. Il s'agit donc de relier des paramètres de haut niveau, comme le temps de réverbération aux divers paramètres de bas niveau du réseau comme les gains et les temps de retard.

Supposons que nous voulions proposer un paramètre "temps de réverbération T_{60} " à l'utilisateur du FDN d'ordre 3 précédent. On peut pour cela agir sur trois paramètres : les temps de retard, les gains et la matrice de

mélange. Modifier les temps de retard N_i modifierait également la sensation d'espace de la réverbération. De la même manière, modifier la matrice A modifierait la distribution d'énergie dans le réseau et donc sa couleur globale. On choisit donc d'agir uniquement sur les gains g_i .

Une première approche serait d'avoir un gain unique $g = g_1 = g_2 = g_3$, ainsi valeur de g proche de 1 augmenterait le phénomène de rétroaction et donc la durée de la réverbération (et inversement pour une valeur de g proche de 0). Cependant un problème apparaît avec cette méthode : le signal qui va passer par le délai le plus court va passer par le gain g plus souvent et donc s'atténuer plus vite. En conséquence la queue de réverbération n'aura pas une décroissance homogène en fréquence et présentera des résonances liées au filtrage en peigne des délais les plus longs.

Pour régler ce problème, il faut relier le gain au temps du retard qui lui est associé, de manière à ce que les signaux qui passent par des petites valeurs de retards soient moins atténués que les autres. Pour assurer une décroissance exponentielle de la réverbération, on propose la relation suivante entre chaque g_i et un paramètre γ (plus de détails dans [16]) :

$$g_i = \gamma^{N_i}$$

Il nous reste à relier le paramètre γ au temps de réverbération T_{60} :

$$\gamma = 10^{\frac{-60T_{60}}{20T_s}} = 10^{\frac{-3T_{60}}{T_s}}$$

Où T_s représente la période d'échantillonnage.

Avec cette expression des gains g_i , la queue de réverbération suit une décroissance exponentielle homogène. De plus, on vient de relier les paramètres de bas niveau très peu parlants que sont les gains à un paramètre beaucoup plus explicite qu'est le temps de réverbération.

La qualité sonore de ces algorithmes dépend de l'ordre du FDN et de la qualité de l'implémentation du constructeur, notamment de sa paramétrisation. Une grande partie des paramètres, notamment la valeur des retards N_i , ainsi que l'ajout de filtres passe-tout supplémentaires restent choisis de façon arbitraire par le constructeur et la prise en compte de la position des sources (au delà du pré-délai) n'est pas intuitive.

3.5 Architecture générale d'un réverbérateur numérique et paramétrisation

La plupart des algorithmes de réverbération numériques sont des agencements de blocs fonctionnels plus élémentaires. Cela est particulièrement vrai pour les algorithmes à réseaux de retards bouclés, pour lesquels l'agencement et le réglage de ces différents éléments contribuent grandement à la qualité générale de la réverbération synthétisée.

On peut établir une architecture générale d'un réverbérateur numérique en le découpant en trois blocs : le module de premières réflexions, le module de champ diffus et le module de sortie. Le module de premières réflexions est le plus souvent constitué d'une ligne de retard de laquelle on extrait les réflexions. Le module de champ diffus peut être à convolution, ou un agencement plus ou moins complexe de filtres élémentaires et de FDN. Enfin le module de sortie s'occupe de spatialiser les différents signaux que l'on extrait des deux premiers modules.

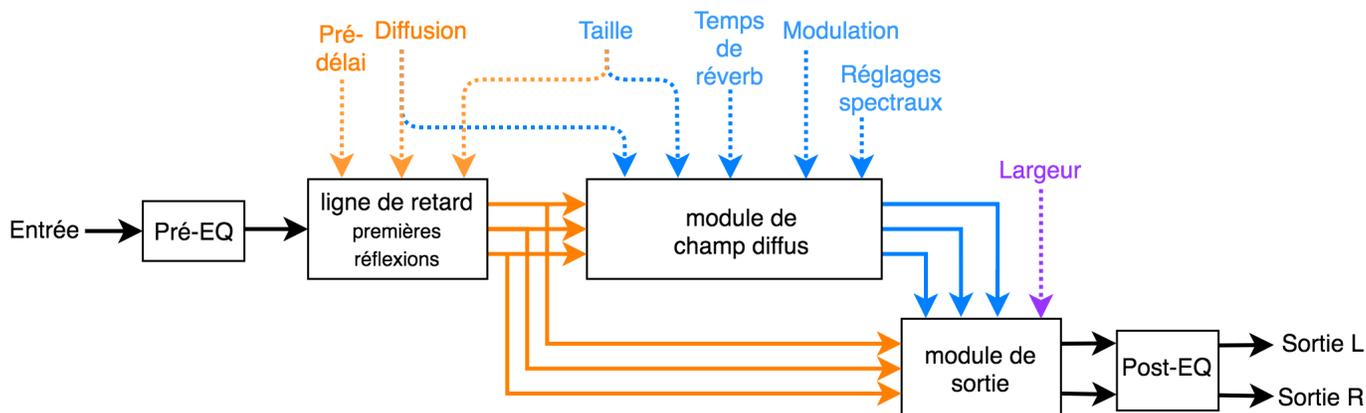


FIGURE 13 – Architecture simplifiée d'un réverbérateur algorithmique stéréophonique.

La plupart des réverbérateurs numériques, quelle que soit leur technologie offrent une grande variété de réglages pour affiner le rendu sonore de la réverbération. Le choix des réglages offerts à l'utilisateur est propre à chaque constructeur qui doit proposer un compromis entre simplicité d'utilisation et exhaustivité des paramètres. Cependant ils existent quelques réglages qui sont communs à la quasi-totalité des réverbérateurs numériques [17] :

- Le temps de réverbération (*Reverberation Time* ou *Decay Time*), parfois exprimé via un temps en secondes, un pourcentage ou une

valeur sans unité. L'utilisation d'un pourcentage est courante pour des réverbérations à convolution pour lesquelles on peut "étirer" la réponse impulsionnelle. L'utilisation d'une valeur de T_{60} est standard pour effectuer des mesures acoustiques mais est basée sur une dynamique trop grande pour le travail usuel d'un ingénieur du son, la queue de réverbération se retrouvera masquée par d'autres éléments si bien que la réverbération paraîtra trop courte, des valeurs de dynamique plus réduite comme le T_{30} ou le T_{20} sont plus adaptées pour indiquer un temps de réverbération.

- Un réglage de pré-délai qui va souvent de 0 à quelques centaines de millisecondes. Comme on l'a vu dans la partie sur la perception psychoacoustique de la réverbération, ce pré-délai joue un grand rôle sur la distance à la source et la taille de la salle perçues ainsi que sur le timbre général du son. Ce réglage est également fortement dépendant du type de son considéré.
- Un paramètre de taille (*Room Size* ou *Dimension*) qui contrôle la densité d'écho et la répartition temporelle des premières réflexions. Un petit paramètre de *Size* va permettre d'atteindre la densité d'écho finale rapidement, comme c'est le cas dans les petites pièces.
- Un paramètre de densité ou de diffusion. Ce paramètre est souvent lié au gain d'un réseau de filtres passe-tout placé en amont de la réverbération ou à l'intérieur de celle-ci dans le but d'émuler les aspérités de la pièce et de créer perceptivement de la diffraction.
- Un paramètre de modulation (*Modulation* ou *Chorus*) qui contrôle une légère oscillation de valeur des délais. Pour des réglages de modulation légers, cela contribue à éliminer certaines résonances et rendre la réverbération moins "métallique"⁶. Pour des réglages plus extrêmes, un effet de chorus se fait clairement entendre, ce qui peut être désirable pour certaines utilisations esthétiques [18]. Il est à noter que ce réglage ne renvoie à aucune réalité physique car il correspondrait à des murs qui "oscille".
- un paramètre spatial (*Width*) pour contrôler la sensation d'enveloppement et d'espace de la réverbération. Ce réglage agit sur le module de sortie en décollant plus ou moins les signaux gauche et droite par

6. Des valeurs de modulation de l'ordre de la dizaine d'échantillons à une fréquence de l'ordre du Hz sont souvent citées

exemple.

En plus de ces paramètres, la plupart des réverbérateurs intègrent des paramètres spectraux qui peuvent être de plusieurs natures : en amont ou en aval du réverbérateur ou à l'intérieur de celui-ci. Un égaliseur en amont ou en aval permet d'effectuer de grosses corrections pour intégrer la réverbération dans le mix global. Par exemple il est assez usuel pour un ingénieur du son de couper la partie grave de la réverbération pour qu'elle ne vienne pas polluer cette zone sensible du spectre (basse, grosse caisse etc).

Un égalisateur intégré au sein même du module de champ diffus permet d'agir plus subtilement sur le timbre de la réverbération car il contrôle alors directement le temps de réverbération par bande de fréquence.

La plupart de ces paramètres renvoient à des phénomènes physiques qui ont été décrits plus haut. L'intégration de ces paramètres est plus simple pour des réverbérateurs à réseaux de retards comme les FDN car on peut agir directement sur les valeurs des lignes de retards et sur les gains. Concernant les réverbérateurs à convolution, la modification de la réponse impulsionnelle nécessaire pour agir sur un paramètre n'est pas forcément aisée, il s'agira plutôt de trouver la réponse impulsionnelle adaptée à nos besoins plutôt que de chercher à trop la modifier.

4 Synthétiser le champ réverbéré en audio orienté objet

4.1 Une première approche de la réverbération en audio orienté objet : l'approche "multicanalement"

Une première façon d'aborder la réverbération en audio orienté objet est de chercher à utiliser les réverbérations les plus répandues chez les ingénieurs du son : les réverbérations multicanales (stéréo ou 5.1 par exemple, sous forme hardware ou software) en les adaptant à un environnement objet. Si on choisit une réverbération stéréophonique, on se retrouve avec les deux sorties à placer dans l'espace. On traite donc les deux sorties de la réverbération comme des objets comme les autres : chaque canal est déclaré

comme un objet et est placé dans l'espace acoustique, au même titre que les sources de son direct, comme illustré sur la figure ci-dessous.

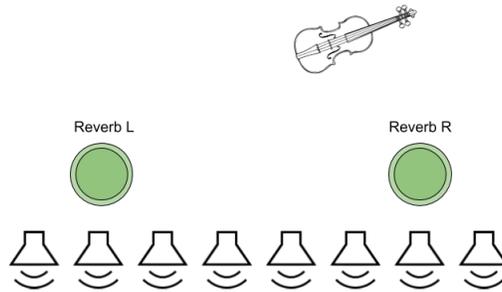


FIGURE 14 – Exemple d'intégration simple de réverbération en WFS

Dans cet exemple où l'on considère une ligne de haut parleur en WFS, la scène est composée de 3 objets : le violon, source de son direct et les deux canaux de réverbération. L'alimentation des canaux de réverbération se fait en amont, dans la console ou la station de travail par exemple.

Cette approche n'est pas mauvaise en soi et peut tout à fait être adaptée au contexte de travail, on peut penser par exemple à des réverbérations très colorées comme une réverbération à ressort pour lesquelles cette approche sera tout à fait adaptée. Néanmoins préciser les avantages et inconvénients d'une approche "naïve" de la réverbération nous permet d'identifier les enjeux de la réverbération orientée objet.

Tout d'abord voici les 2 avantages les plus évidents qu'offre cette méthode :

- Elle est facile à mettre en place.
- Elle permet de travailler avec les réverbérations dont on a l'habitude et dont on connaît à la fois le comportement et le rendu sonore.

On peut cependant relever un certain nombre de potentiels inconvénients voir de vrais problèmes avec cette méthode :

- On travaille avec des réverbérations qui ont été conçues pour un format de restitution bien précis (la stéréophonie, le 5.1 etc), différent a priori de notre configuration de sonorisation. On se retrouve donc souvent à placer nos points de réverbérations de façon plus ou moins heureuse dans l'espace pour essayer de recréer artificiellement le format d'origine de la réverbération.

- On ne tire pas vraiment parti de l'aspect objet puisqu'on ne tient pas en compte des métadonnées des objets comme leur position relative ou de leur directivité.
- On peut être sujet à des distorsions d'espace. Dans un exemple extrême, si notre violon est placé très loin derrière notre réverbération, en l'absence de gestion du pré-delay, il se peut que le champ réverbéré arrive avant le son direct.
- il n'y a pas de gestion automatique du rapport champ direct sur champ diffus : la source alimente la réverbération quelle que soit sa position. On doit gérer ce rapport manuellement via un niveau d'envoi si on désire faire un effet d'éloignement ou de rapprochement par exemple, ce qui peut être souhaitable ou non.

4.2 L'approche "paramétrique"

4.2.1 Méthode basée sur l'analyse d'une réponse impulsionnelle

La réponse impulsionnelle d'une salle est intéressante pour la qualité sonore et le réalisme qu'elle apporte une fois convoluée avec un son que l'on désire réverbérer. Le principal problème qui se pose dans un cadre orienté objet c'est qu'il n'est pas envisageable d'avoir un jeu de réponse impulsionnelle adapté à chaque configuration du système de reproduction. Le déploiement de réponses impulsionnelles pour des configurations génériques (5.1, quadriphoniques etc) pour des configurations différentes mène aux mêmes problèmes que nous évoquions dans la partie précédente.

Pour pallier à ce problème, l'idée est d'extraire des informations utiles concernant la réverbération d'un jeu de réponses impulsionnelles pour synthétiser un nouveau jeu de réponses impulsionnelles adapté au système de diffusion. On parle alors d'une méthode par analyse-synthèse, on analyse un jeu de réponses impulsionnelles pour en créer un autre.

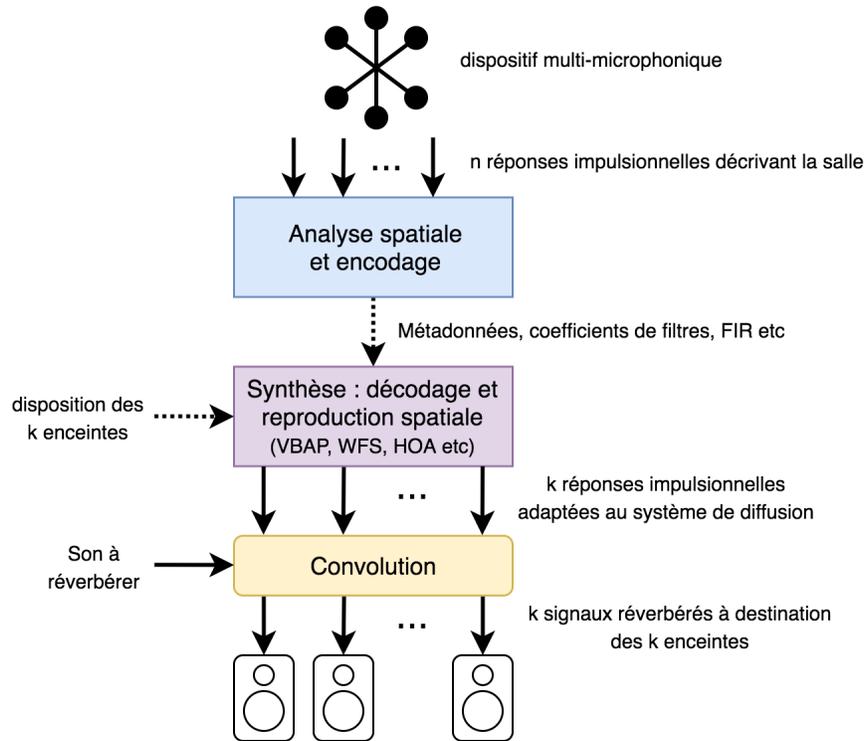


FIGURE 15 – Principe général de réverbération par analyse-synthèse pour une réponse impulsionnelle à n microphones et un système de diffusion à k enceintes.

Il faut donc disposer d'un jeu de réponses impulsionnelles qui décrivent précisément l'espace. Les dispositifs les plus utilisés sont les microphones format-B Soundfield ou des configurations circulaires de microphones (*circular microphone array*) comme présentés sur la figure ci-dessous. Plus le nombre de microphones est élevé plus la précision des informations extraites lors de la phase d'analyse est grande.

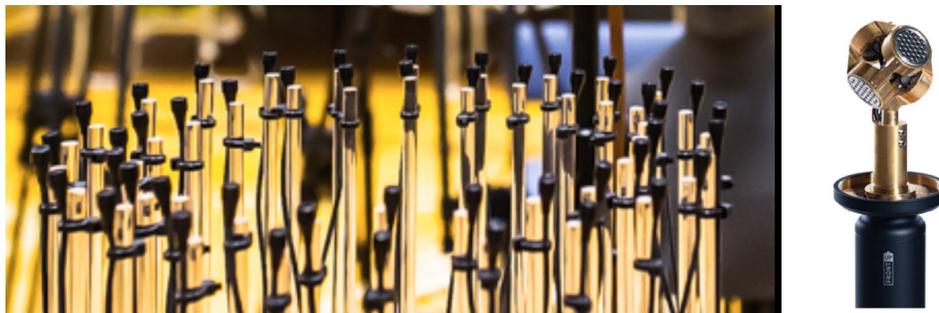


FIGURE 16 – Un dispositif multi-microphonique circulaire et un microphone ambisonique tétraédrique *AMBEO VR Mic* de chez *Sennheiser*.

Il existe plusieurs méthodes de réverbération objet par analyse-synthèse comme le SIRR [19] (*Spatial Impulse Room Rendering*), le RSAO [20] (*Reverberant Spatial Audio Object*), le SDM [21] (*Spatial Decomposition Method*) ou le R-WFS [22]. Ces techniques ont toutes en commun le fait de séparer la réverbération en deux parties distinctes : les premières réflexions et le champ diffus.

En connaissant la configuration géométrique du dispositif microphonique, on peut extraire d'un jeu de réponses impulsionnelles un grand nombre d'informations spatiales sur les premières réflexions et le champ diffus. Parmi ces informations peut retrouver :

- Pour les premières réflexions, leur direction d'arrivée (*DOA* pour *Direction of Arrival*), leur temps d'arrivée (*TOA* pour *Time of Arrival*) ainsi que leur amplitude et leur coloration spectrale.
- Pour le champ diffus, sa diffusivité, c'est à dire sa cohérence interaural ou encore son enveloppe de décroissance par bande de fréquence.

Lors de l'étape de synthèse, il s'agit de recréer un jeu de réponses impulsionnelles adaptés au système de diffusion en connaissant la configuration de ce dernier. Du fait de leur importance dans notre perception de la réverbération, les premières réflexions sont spatialisées précisément (par VBAP, ou comme des sources ponctuelles en WFS par exemple). Pour le champ diffus, pour lequel notre oreille est peu précise, plusieurs méthodes sont envisageables pour le reconstruire : on peut moduler un bruit gaussien par bande de fréquence ou utiliser divers algorithmes de décorrélation qui altèrent la phase ou de la convolution de diffusion.

Ces méthodes permettent donc d'utiliser des réverbérations à convolution pour n'importe quel système de diffusion spatialisé et permettent de placer une source dans un espace acoustique de façon réaliste. Elles permettent de modifier certains paramètres de la réverbération (amplitude des premières réflexions, diffusivité etc) en modifiant les informations extraites lors de la phase d'analyse.

En revanche, la prise en compte d'éventuels mouvements de la source n'est pas aisée et on se retrouve avec une scène sonore relativement fixe. En effet la position de la source réverbérée est identique à celle de la source sonore par rapport aux microphones lors de l'enregistrement de la réponse

impulsionnelle.

4.2.2 Interpolation de réponses impulsionnelles

Afin de permettre le rendu de mouvement, l'idée de "quadriller" une salle en prenant des réponses impulsionnelles à plusieurs endroits est apparue. L'enregistrement d'un très grand nombre de réponses impulsionnelles (une infinité en théorie) correspondant à chaque position de la source et des microphones est irréalisable en pratique. On choisit en général de faire varier soit la position de la source, soit celle du point d'écoute et d'enregistrer un nombre suffisant de réponses impulsionnelles pour couvrir l'espace.

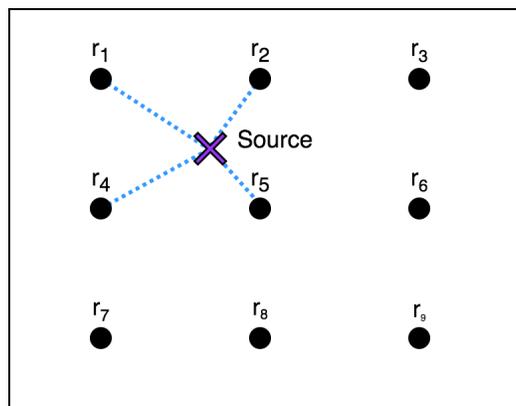


FIGURE 17 – Exemple d'une pièce quadrillée par 9 réponses impulsionnelles.

Il s'agit alors de réaliser une interpolation des réponses impulsionnelles proches de la source pour approcher la réponse appropriée. On voit bien comment une simple interpolation linéaire, c'est à dire une moyenne des réponses impulsionnelles proches pose problème : cela conduirait à l'addition de toutes les premières réflexions. Pour estimer les premières réflexions correspondant à la position exacte de la source, d'autres algorithmes d'interpolations plus complexes sont utilisés [23].

5 Réalisation de notre réverbérateur orienté objet

5.1 Cahiers des charges

Afin de mener à bien la réalisation de notre réverbérateur orienté objet, on en établit le cahier des charges. Notre réverbérateur devra proposer :

- Une gestion dynamique de la position des sources, pour mieux les intégrer à l'espace acoustique.
- Une gestion des premières réflexions indépendante du champ diffus et une spatialisation précise de celles-ci du fait de leur importance sur la perception de l'espace (estimation de la distance, de la taille du lieu etc).
- Une puissance de calcul nécessaire raisonnable car l'algorithme devra cohabiter avec un algorithme de spatialisation déjà gourmand en ressource.
- Un nombre réduit de paramètres pour l'utilisateur mais capables de simuler en quelques manipulations des acoustiques diverses (pour retrouver les réverbérations usuelles type "Room", "Hall", "Plate" etc). On veillera à bien paramétrer notre réverbérateur, c'est à dire offrir des réglages efficaces et avec de bonnes plages de valeurs pour qu'il soit confortable à utiliser.

En cohérence avec les deux premiers points du cahier des charges, on propose l'architecture suivante pour notre réverbérateur :

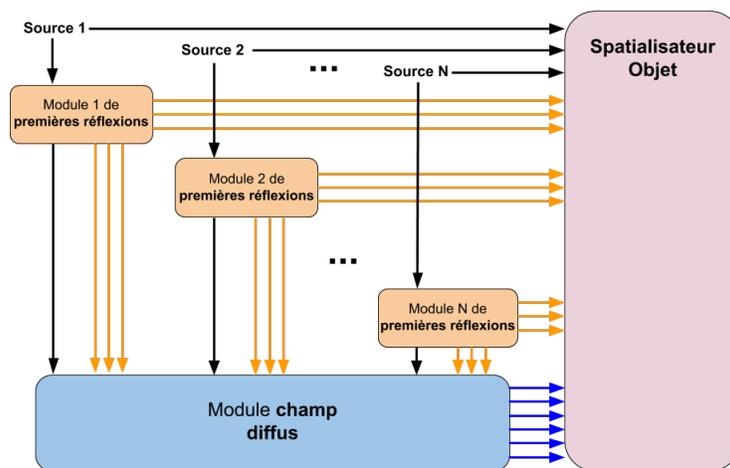


FIGURE 18 – Architecture générale du réverbérateur, chaque flèche représente un signal sonore accompagné de ses métadonnées de position.

L'utilisation d'un module de premières réflexions pour chaque source permet d'en avoir un contrôle précis. Il est en effet souhaitable de pouvoir gérer le niveau de premières réflexions pour chaque source en fonction de sa nature. De plus, cela offre la possibilité d'appliquer une spatialisation particulière pour les premières réflexions au sein du spatialisateur objet.

5.2 Choix d'un logiciel de travail

Pour développer notre réverbérateur orienté objet, on choisit le logiciel *Max* pour les raisons suivantes :

- C'est un logiciel de programmation visuelle qui permet de coder en connectant entre eux des blocs fonctionnels au sein d'un *patch*, il est donc assez simple à prendre en main pour un novice en programmation. Il permet également de modifier son code en temps réel sans avoir à compiler le patch à chaque changement.
- C'est un logiciel bien connu et documenté pour lequel on trouve énormément de tutoriels et d'exemples de patches en ligne.
- Il permet de réaliser des interfaces graphiques très correctes en intégrant facilement toute sorte d'éléments d'interaction pour l'utilisateur comme des boutons, des rotatifs etc. L'utilisation de *JavaScript* au sein de Max permet de réaliser des interfaces graphiques encore plus sophistiquées.
- Comme c'est un logiciel de programmation haut niveau, il est facile à utiliser mais ne permet pas de coder des applications très optimisées en terme de puissance de calcul. Une solution à ce problème est l'utilisation de l'extension *Gen*. Cette extension permet de coder des applications dans un environnement de programmation visuelle similaire à *Max* mais à un plus bas niveau car elles sont compilées directement en *C*. *Gen* a également l'avantage de travailler à l'échantillon près, là où *Max* travaille sur des vecteurs de sons ce qui offre une plus grande précision temporelle dans l'utilisation de retard par exemple.

5.3 Présentation du réverbérateur

On présente ici une version stéréophonique de notre réverbérateur. La spatialisation des objets sonores se fait par un simple calcul de différence de

temps et d'intensité entre les deux oreilles de l'auditeur (la distance entre les deux oreilles est réglable via le rotatif appelé "Largeur tête").

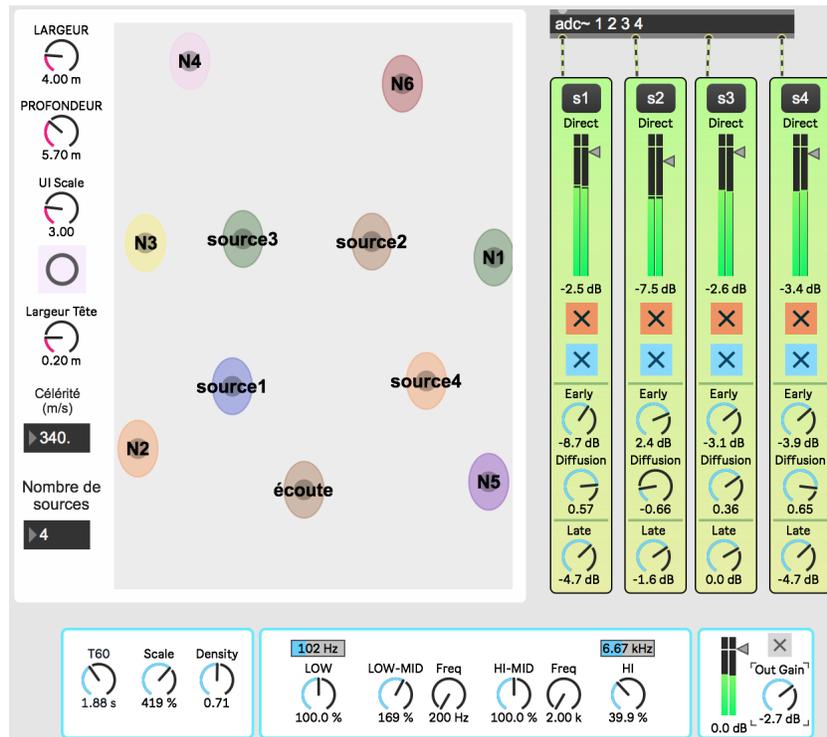


FIGURE 19 – Vue d'ensemble de l'environnement de travail avec 4 sources sonores

On retrouve trois éléments caractéristiques d'un environnement de mixage objet que l'on a décrits dans la section 2.2, à savoir :

- Une interface graphique pour positionner les sources, le point d'écoute et les noeuds de réverbération dans un espace, ici en 2D.
- La configuration du système d'écoute est ici très simple puisqu'elle se limite au choix de la largeur de la tête de l'auditeur qui intervient dans les calculs de différence de temps et d'intensité.
- La configuration des objets qui se fait via les 4 sorties du bloc *adc~* qui permet d'interagir avec des périphériques audios ou une carte son virtuelle pour "amener" le son dans le patch.

A chaque source est associée une tranche de mixeur pour régler le niveau du son direct, le niveau et la diffusion des premières réflexions ainsi que le niveau d'envoi dans le module de champ diffus.

L'interface graphique propose de positionner aussi bien les sources que l'auditeur dans l'espace. On peut également placer des objets nommés $N1$ jusqu'à $N6$ qui correspondent à nos points de réverbération. Ces points de réverbérations sont indispensables au fonctionnement du module de champ diffus, dont on retrouve le panneau de contrôle en bas.

5.4 Choix technologiques

Afin de réaliser notre prototype de réverbérateur orienté objet, nous devons choisir la technologie principale qui nous servira à le coder. On a le choix entre les trois grandes familles de méthodes de synthèse de champ réverbéré que sont : la convolution avec une réponse impulsionnelle de salle, l'acoustique informatique et les algorithmes à réseaux de retards bouclés. Ces trois méthodes sont répandues dans les réverbérateurs disponibles sur le marché et sont donc capables de produire des résultats sonores de grande qualité si elles sont assez bien implémentées. Afin de faire notre choix, nous examinerons ces méthodes selon trois critères :

- La *modularité* : les métadonnées de position des objets sont-elles facilement intégrables dans un algorithme utilisant cette technologie ?
- La *complexité algorithmique* : est-ce que cette technologie mobilise trop de puissance de calcul pour cohabiter avec un algorithme de spatialisation lui-même gourmand en ressource ?
- La *faisabilité* : y-a-t'il assez de ressources et de documentations sur cette technologie pour me permettre de l'implémenter dans mon prototype ? Si oui en ai-je les compétences techniques ?

La réverbération par convolution a l'avantage d'offrir une très grande qualité sonore en reposant sur une opération mathématique simple (l'opération de convolution est très facile à réaliser dans un logiciel comme *Max*). Cette simplicité fait également sa faiblesse dans le cadre de l'audio orienté objet : il est très difficile de modifier la réponse impulsionnelle en cohérence les positions des sources ou pire, de leurs déplacements. Il faut, comme nous l'avons vu au chapitre 4 envisager des méthodes par analyse-synthèse assez complexes ou des maillages de l'espace par un grand nombre de réponses impulsionnelles pour adapter la réverbération par convolution à l'audio orienté objet. De plus l'opération de convolution mobilise beaucoup de

puissance de calcul.

L’acoustique informatique, c’est à dire la simulation de la propagation du son dans une pièce est la méthode la plus modulaire et la plus facilement adaptable à l’audio orienté objet. Son principe même suppose la connaissance de la géométrie de la salle ainsi que des positions de la source et du point d’écoute. Cette méthode pose cependant deux problèmes dans le cadre de notre travail : elle demande beaucoup de puissance de calcul et elle me paraît compliquée à mettre en place techniquement (description de la salle, modélisation fine des phénomènes physiques etc).

Les algorithmes à réseaux de retards bouclés sont très intéressants pour leur très faible complexité algorithmique. Leur qualité sonore est assez variable car ils dépendent fortement de l’agencement et du réglage des différents blocs élémentaires qui les constituent. Ces algorithmes offrent une certaine modularité car on peut faire correspondre des valeurs de délais ou de gains à des paramètres d’espace (positions des sources, taille de la salle etc) mais cela reste assez grossier car ils ne se basent pas sur une modélisation physique fine de la réverbération dans une salle. Ils sont très bien documentés techniquement et relativement facilement implémentables dans un logiciel comme *Max* à l’aide de l’extension *Gen*.

	Modularité	Complexité algorithmique	Faisabilité
Réverbérateurs à convolution	faible	forte	forte
Acoustique informatique	forte	très forte	faible
Réseaux de retards bouclés	moyenne	faible	forte

FIGURE 20 – Les méthodes de synthèse de champ réverbéré évaluées selon nos critères de modularité, de complexité algorithmique et de faisabilité

La faible modularité des méthodes utilisant la convolution nous pousse à les écarter dans le cadre de ce travail. Dans la sous-partie suivante, on décrit en détail la méthode retenue pour notre réverbérateur, à savoir la technologie de *Scattering Delay Network*, à mi-chemin entre acoustique informatique et réseaux de retards bouclés. Le choix de cette technologie permet de concilier modularité et complexité algorithmique.

5.5 Module de champ diffus

5.5.1 Scattering Delay Network

Un *Scattering Delay Network*, abrégé en SDN, est un algorithme de réverbération qui mélange une approche réseaux de retards bouclés (FDN) et une approche d'acoustique informatique [24]. Un SDN est composé de plusieurs noeuds que l'on peut se représenter comme étant chacun une surface réfléchissante d'une salle (un mur par exemple). Chaque noeud reçoit les signaux venant de la source sonore et des autres noeuds, avec un retard et un gain qui dépend de leur distance respective. Chaque noeud réalise ensuite un mélange de ces différents signaux (via une matrice de mélange) et les renvoie vers les autres noeuds et vers l'auditeur. Les sources sonores, l'auditeur et les noeuds qui échangent entre-eux constituent donc un réseau qui assure la dispersion et le mélange du son pour créer un champ réverbéré.

On appelle ordre d'un SDN le nombre de noeuds qui le composent.

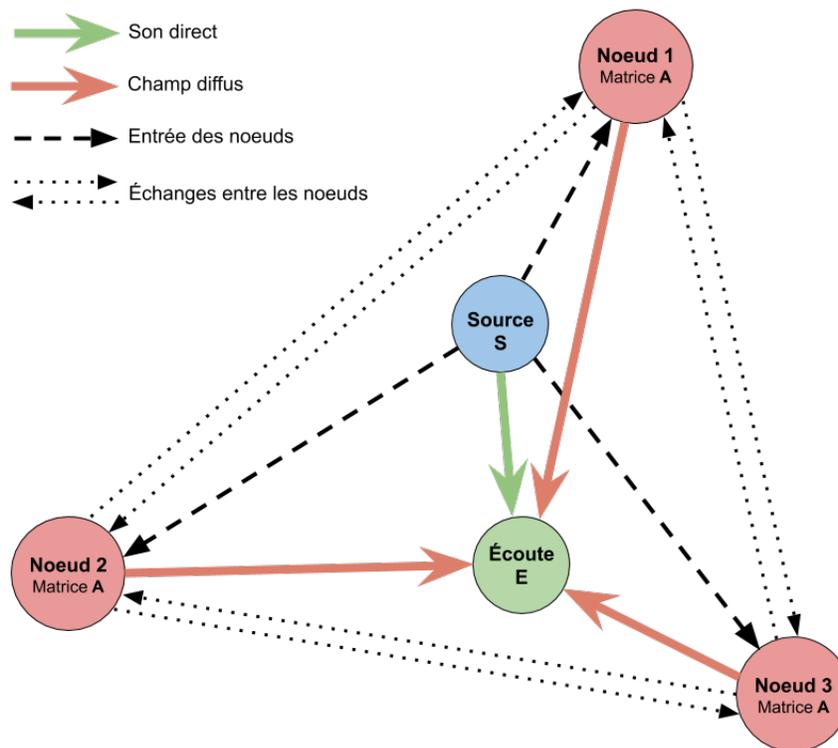


FIGURE 21 – Principe d'un SDN d'ordre 3

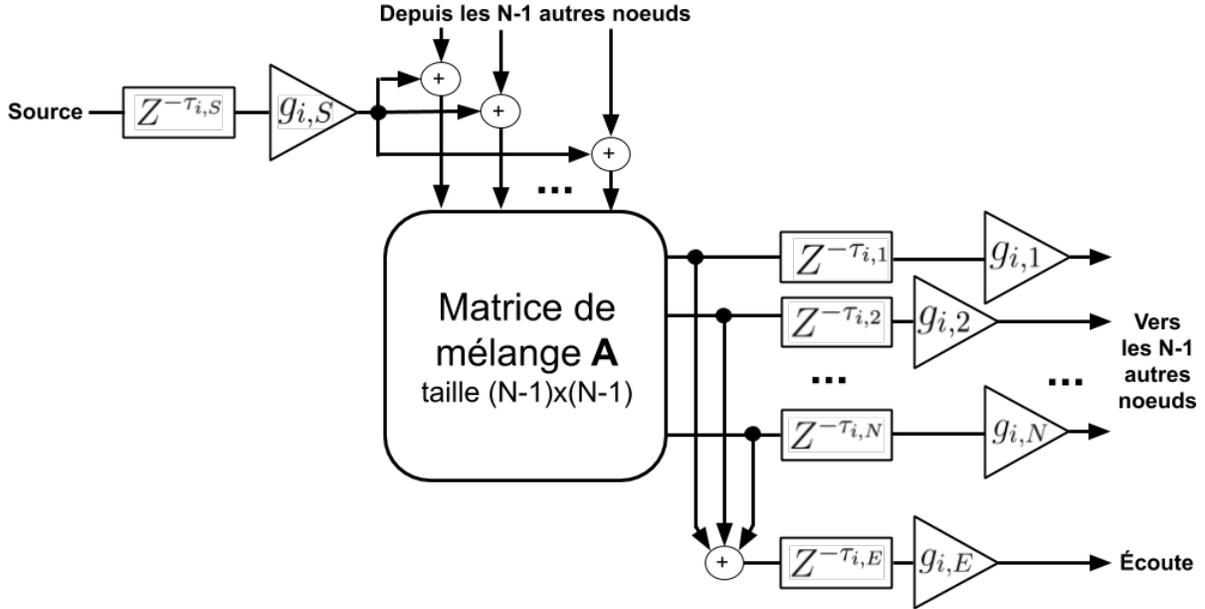


FIGURE 22 – Détail du i -ème noeud d'un réseau de N noeuds, g et τ représentent respectivement l'atténuation et le retard liés à la distance entre deux éléments du réseau

On retrouve donc bien l'esprit d'un FDN avec un réseau de retards interconnectés (on peut montrer qu'un FDN est un cas particulier de SDN) mais il y a deux différences majeures entre les deux.

Premièrement, toutes les valeurs de gains et de retards d'un SDN sont liées à la modélisation d'une salle via la position des noeuds. Dans le cas d'un FDN, ces choix de gains et de retards sont arbitraires, ils sont le plus souvent choisis par le constructeur du réverbérateur. Cette description de la salle par un faible nombre de noeuds qui représenteraient les "murs" est certes sommaire mais suffisante à rattacher tous les gains et retards du système à des paramètres de position, c'est ce que l'on a appelé dans la sous-partie précédente la *modularité*.

Secondement, là où dans un FDN, l'opération de matricage n'a lieu qu'une seule fois, elle a lieu à chaque noeud au sein d'un SDN. Il est résulte que l'énergie "circule" davantage et est mieux répartie dans un SDN, autrement dit, un algorithme à SDN aura une densité d'écho supérieure à un FDN de même ordre [24]. Ce sont ces multiples opérations de matricage qui donnent l'adjectif *Scattering* à cette méthode, car ce sont ces opérations qui participent à "dispenser" l'énergie dans le réseau, afin d'obtenir une réverbération diffuse.

Un SDN est adapté à l'audio orienté objet car sa structure même est définie par le placement dans l'espace des noeuds, des sources et du point d'écoute. C'est pour cette raison que l'on peut déplacer les noeuds $N1$, $N2$... $N6$ au sein de notre interface graphique (*c.f.* Figure 19).

Si un SDN s'inspire dans sa structure d'un modèle "simulation d'acoustique" où chaque noeud peut être vu comme un mur et le réseau comme une salle, il est important de noter que ce n'est pas en cherchant à reproduire une pièce existante que l'on obtient les résultats les plus intéressants. En particulier, c'est en disposant les noeuds en cercle ou en rectangle que l'on viendra reproduire certains défauts indésirables de pièces existantes comme des résonances très marquées. Il apparaît intéressant au contraire de casser au plus possible les symétries dans la répartition des noeuds et d'éviter de les placer à équidistance les uns des autres.

On notera que la puissance de calcul requise est plus élevée pour un SDN que pour un FDN du même ordre. En particulier un SDN d'ordre N mobilise $N(N - 1)$ lignes de retard là où un FDN en mobilise seulement N . Cela reste raisonnable sur des machines actuelles et ce coût supplémentaire en ressource se justifie par la meilleure densité d'échos des SDN que nous évoquions juste avant.

5.5.2 Le choix de la matrice de mélange

Le choix de la matrice de mélange A est important car il détermine la façon dont l'énergie est répartie au sein du réseau. On peut dire que la matrice A fixe la façon dont le son "rebondit" sur les murs que représentent les noeuds.

Comme pour les FDN, des critères mathématiques pour garantir la stabilité du système. Le document [24] propose la matrice de mélange isotrope suivante (*isotropic scattering matrix*, cas particulier d'une classe de matrice plus générale appelée *matrices de Householder*) :

$$A = \frac{2}{N - 1}J - I$$

où J est la matrice dont tout les coefficients sont égaux à 1 et I la matrice identité, toutes deux de dimension $N - 1$. Cette matrice sert à

repartir l'énergie de façon isotrope, c'est à dire en ne privilégiant pas de "direction" particulière.

Pour $N = 4$, c'est à dire pour le cas d'un SDN à 4 noeuds, on a :

$$A = \frac{2}{3} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \end{pmatrix}$$

Si on nomme x_2 , x_3 et x_4 les signaux provenant des noeuds 2, 3 et 4, alors le noeud 1 va envoyer vers le reste du réseau les signaux x'_2 , x'_3 et x'_4 le signal y_1 vers la sortie :

$$\begin{aligned} x'_2 &= -\frac{1}{3}x_2 + \frac{2}{3}x_3 + \frac{2}{3}x_4 \\ x'_3 &= \frac{2}{3}x_2 - \frac{1}{3}x_3 + \frac{2}{3}x_4 \\ x'_4 &= \frac{2}{3}x_2 + \frac{2}{3}x_3 - \frac{1}{3}x_4 \\ y_1 &= x'_2 + x'_3 + x'_4 \end{aligned}$$

Cette matrice possède une caractéristique qui la rend très efficace en terme de puissance de calcul. En notant $X = \frac{2}{N-1}(x_1 + x_2 + \dots + x_N)$ on peut écrire chaque signal de sortie x'_i sous la forme :

$$x'_i = \frac{2}{N-1}(x_1 + x_2 + \dots + x_N) - x_i = X - x_i$$

Ainsi il est nécessaire de calculer le terme X qu'une seule fois, il suffit ensuite d'une seule soustraction pour obtenir x'_i . Il en résulte que l'opération totale de matricage peut s'effectuer en 1 multiplication et $2N - 1$ additions à la place de N^2 multiplications et $N(N - 1)$ additions pour une matrice quelconque.

5.5.3 Implémentation pratique du module de champ diffus

Notre réverbérateur intègre un SDN d'ordre 6 pour synthétiser le champ diffus.



FIGURE 23 – Interface utilisateur du module de champ diffus

Les réglages sur lesquels l'utilisateur peut agir pour ajuster la réverbération sont :

- un contrôle du temps de réverbération T_{60} .
- un réglage de "Scale" qui vient simplement multiplier la valeur des retards du SDN d'un même facteur. Une faible valeur de Scale va générer une grande densité d'échos mais présenter des résonances très marquées donnant un effet "boîte de conserve". Une grande valeur de Scale permet d'obtenir des résultats plus homogènes fréquentiellement, même pour simuler de petits espaces (le facteur "Scale" est indépendant du T_{60}).
- un réglage de "Density" qui vient contrôler l'intensité d'un filtre passe-tout en série avec chaque délai et qui contribue à augmenter la sensation de "diffusivité" du champ réverbéré en augmentant encore la densité d'écho et en altérant la réponse en phase. C'est un paramètre qui fonctionne plus ou moins bien pour les signaux transitoires, le passe-tout peut générer des flutters pour les sons percussifs qui leur donnent une sonorité métallique indésirable.
- un contrôle du temps de réverbération par bande de fréquence pour modifier la couleur spectrale de la réverbération. On peut passer de 10% à 900% du T_{60} sur 4 bandes de fréquence (filtre shelf pour le grave et l'aigu et en cloche pour les médiums).

Examinons le code *Gen* qui implémente notre SDN d'ordre 6.

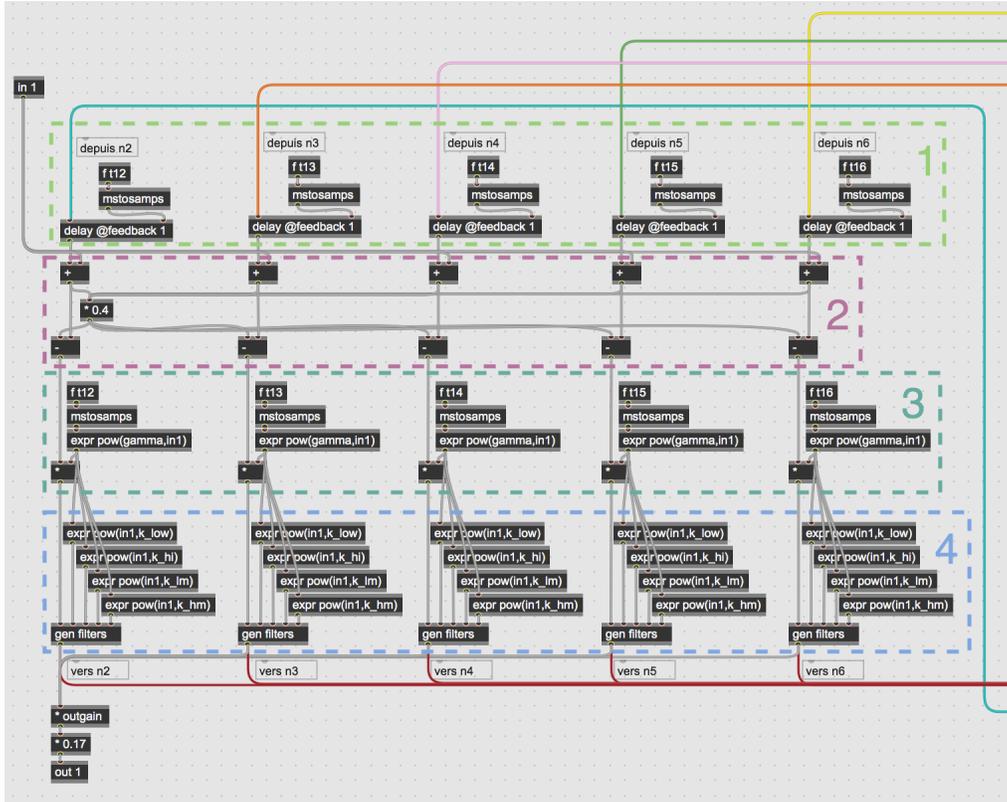


FIGURE 24 – Détail du premier noeud d’un SDN d’ordre 6 codé dans *Gen*.

- Zone 1 : application d’un retard correspondant à la distance entre le premier noeud et les autres noeuds du réseau via les blocs *delay*. Le bloc *mstosamps* converti un temps en millisecondes en un nombre d’échantillons en fonction de la fréquence d’échantillonnage choisie.
- Zone 2 : opération de matricage, grâce à l’utilisation de la matrice A décrite plus haut, cette opération peut être implémentée très élégamment et efficacement.
- Zone 3 : application d’un gain dépendant la distance entre le premier noeud et les autres noeuds du réseau et du réglage de temps de réverbération (appelé *gamma* dans le patch). Cela permet d’assurer une décroissance homogène de la queue de réverbération et d’éviter les résonances de certaines fréquences.
- Zone 4 : réglage du temps de réverbération par bande de fréquence via un égalisateur 4 bandes : filtre *shelf* pour le grave et l’aigu (gains k_{low} et k_{hi}), filtre en cloche pour les deux fréquences médiums (gains k_{lm} et k_{hm} pour *low-mid* et *hi-mid*). Ces réglages dépendent également de la distance entre le premier noeud et les autres

noeuds du réseau et du réglage de temps de réverbération.

5.6 Module de premières réflexions

Si en théorie un SDN recrée lui même des premières réflexions appropriées à l'espace que l'on souhaite modéliser, il s'avère intéressant de lui associer un module de premières réflexions séparé pour en avoir un contrôle plus précis. On veut pouvoir gérer leur niveau indépendamment du champ diffus pour chaque source. On veut également pouvoir contrôler la couleur de ces réflexions via un réglage de diffusion. Ce module, propre à chaque source permet de spatialiser ces réflexions : en WFS on les spatialisera comme des sources à part entière pour les localiser le plus précisément possible.

Le module de premières réflexions calcule 3 réflexions selon un modèle très simple de "boîte à chaussures" : les réflexions de 1er ordre sur les murs latéraux et sur le mur du fond. Les dimensions la pièce sont précisées dans l'interface utilisateur via le réglage de Profondeur et de Largeur (cf. Figure 19).

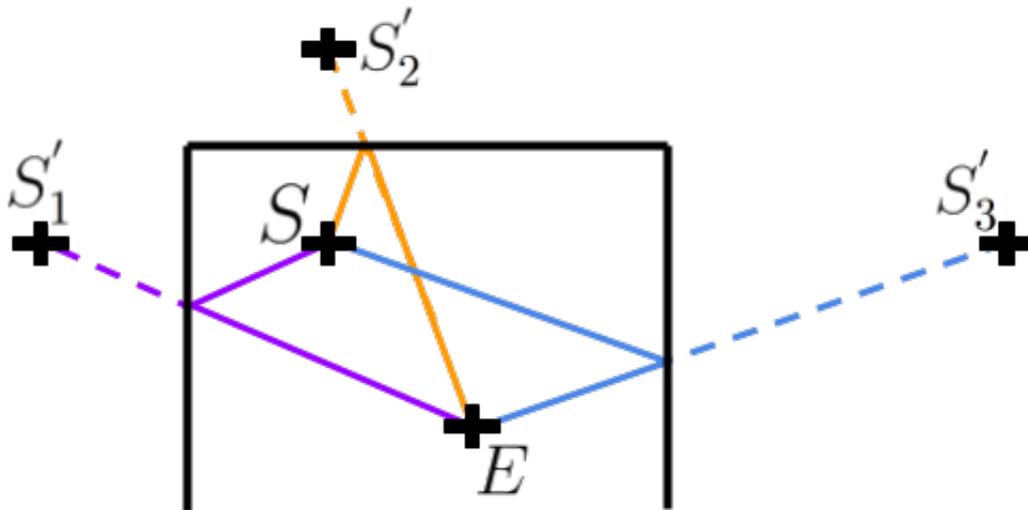


FIGURE 25 – Calcul des premières réflexions

6 Attentes relatives à la partie pratique

Du fait du confinement, j'ai eu l'occasion de développer une version stéréophonique de mon réverbérateur. Cela m'a permis de réaliser une bonne partie du codage puisque j'ai réussi à implémenter le SDN en lui même, un module (imparfait) de premières réflexions ainsi qu'un temps de réverbération réglable par bande de fréquence. Voici la liste de choses que je peux encore améliorer selon moi dans le temps imparti de cette partie pratique :

- Revoir le module de première réflexions afin de mieux faire le lien entre elles et le champ diffus (SDN). Cela améliorera la cohérence de l'espace acoustique en évitant un effet de superposition d'un module premières réflexions et d'un module de champs diffus.
- Améliorer la qualité sonore perçue de réverbération. On peut penser à la modulation de certaines lignes de retard ou à l'intégration de modules de diffusion via des filtres passe-tout. La difficulté est que la plupart de ces modifications sont arbitraires et sont difficiles à régler, nécessitant beaucoup d'expérimentations à tâtons, sans vraiment d'autre justification que "ça sonne mieux avec que sans".
- Ajouter une interface graphique en JSUI (*JavaScript User Interface*) pour améliorer le confort d'utilisation de la réverbération. Potentiellement très chronophage.

Le contrecoup de la situation épidémique et que j'ai n'ai pas vraiment eu l'occasion d'expérimenter avec mon réverbérateur en WFS. Pour cette raison, j'ai choisi de très peu évoquer le WFS dans cette première version du mémoire, qui sera par contre l'enjeu le plus important de la partie pratique. On étudiera et expérimentera les possibilités de spatialisation offertes par la WFS, en s'attardant sur les points suivants :

- La spatialisation des premières réflexions comme des point-sources à part entière. On se posera la question du calcul des premières réflexions qui dépendent du choix d'un "auditeur type" placé dans la salle et de l'impact que cela peut avoir pour un auditeur situé ailleurs.
- La spatialisation du champ diffus, c'est à dire de la sortie des différents noeuds du SDN. On exploitera la capacité de la WFS à reproduire des ondes planes puisqu'il a été montré qu'un champ réverbéré peut être reconstitué sous la forme d'une superposition de plusieurs

ondes planes [7].

- L'utilisation ou non d'enceintes latérales ou arrières pour estimer leur intérêt dans la sensation d'immersion et dans la précision de localisation des sources.

Conclusion

En étudiant le phénomène physique et psychoacoustique de la réverbération, on a identifié les opportunités que l'audio orienté objet pouvait offrir dans la synthèse de champ réverbéré. Parmi ces enjeux on a particulièrement insisté sur l'utilisation des métadonnées de position pour prendre en compte sur la double dépendance de position source-auditeur au sein d'une salle ainsi que les possibilités de spatialisation qu'offrent des techniques de sonorisation orientées objet comme la WFS.

Différentes approches ont déjà été proposées pour intégrer la réverbération en audio orienté objet, la plupart de ces méthodes font appel à des réponses impulsionnelles qui permettent difficilement de prendre en compte la position des sources ou leur déplacement. De ces études a découlé l'élaboration d'un réverbérateur utilisant une technologie de *Scattering Delay Network* à mi-chemin entre acoustique informatique et réseaux de retards bouclés pour concilier deux des enjeux primordiaux de la réverbération orientée objet : la facilité d'intégration des métadonnées de position des sources et une faible demande en puissance de calcul.

Notre réverbérateur produit des résultats sonores globalement satisfaisants. Son point fort pour moi réside dans la mise en espace des sources, le son direct semble bien immergé dans un espace cohérent et la réverbération n'est pas perçue comme un ajout artificiel. On se rend particulièrement compte de l'importance de la prise en compte des paramètres de positions des sources lors de leur déplacement. Les réglages de salle de petit volume et au faible temps de réverbération sont particulièrement efficaces pour "décoller" une source des enceintes. Son point faible réside pour moi dans la faible palette de réverbérations que l'on peut obtenir. Même si notre réverbérateur couvre un large de taille de salle, on retombe souvent sur la même "couleur" de réverbération.

L'aspect objet de notre réverbérateur peut encore être poussé d'avantage car sa structure est très modulable. On peut par exemple imaginer de prendre en compte le diagramme de directivité d'une source. On pourrait aussi créer des espaces plus complexes en attribuant des caractéristiques

fréquentielles et d'absorption pour chaque noeud et ainsi simuler divers matériaux pour les "murs" du réseau SDN.

Références

- [1] Octávio Inácio. Fundamentals of room acoustics. *Lda. R. Dr. Carlos Pires Felgueiras*, 2002.
- [2] Trevor J. Cox and Peter d’Antonio. *Acoustic Absorbers and Diffusers : Theory, Design and Application*. 2016.
- [3] C. Faller and J. Merimaa. Source localization in complex listening situations : Selection of binaural cues based on interaural coherence. *The Journal of the Acoustical Society of America*, vol. 116, novembre 2004.
- [4] Ruth Litovsky, H. Steven Colburn, William Yost, and Sandra Guzman. The precedence effect. *The Journal of the Acoustical Society of America*, 106, 1999.
- [5] David Griesinger. The physics and psycho-acoustics of surround recording. 2005.
- [6] J.J. Sonke. *Variable acoustics by wave field synthesis*. PhD thesis, Université de Technologie de Delft, 2000.
- [7] Diemer Diemer de Vries, Bernd Fröhlich, Frank Melchior, and Christoph Sladeczek. User-dependent optimization of wave field synthesis reproduction for directive sound fields. *AES Convention Papers Forum*, Mai 2008.
- [8] David Gould. What dolby atmos software is right for you? <http://www.avidblogs.com/what-dolby-atmos-software-is-right-for-you/>, 2020.
- [9] V. Välimäki, J. D. Parker, L. Savioja, J. O. Smith, and J. S. Abel. More than fifty years of artificial reverberation. *AES 60th International Conference*, février 2016.
- [10] M. R. Schroeder. Natural sounding artificial reverberation. *Journal de l’AES*, juillet 1962.
- [11] Julius O. Smith. Physical audio signal processing. https://ccrma.stanford.edu/~jos/pasp/Schroeder_Reverberators.html, 2010.
- [12] Sean Costello. Reverbs : Diffusion, allpass delays, and metallic artifacts. <https://valhalladsp.com/2011/01/21/reverbs-diffusion-allpass-delays-and-metallic-artifacts/>, 2011.

- [13] J.-M. Jot. Efficient models for reverberation and distance rendering in computer music and virtual audio reality. *ICMC : International Computer Music Conference*, 1997.
- [14] T. Carpentier, M. Noisternig, and O. Warusfel. Hybrid reverberation processor with perceptual control . *17th International Conference on Digital Audio Effects*, 2014.
- [15] Sebastian J. Schlechtand and Emanuël A.P. Habets. On lossless feedback delay networks. *IEEE Transactions on Signal Processing*, 2016.
- [16] J.-M. Jot. Digital delay networks for designing artificial reverberators. *90ème Convention de l’AES*, 1991.
- [17] Pawel Malecki, Katarzyna Sochaczewska, and Jerzy Wiciak. Settings of reverb processors from the perspective of room acoustics. *Journal of Audio Eng. Soc.*, Avril 2020.
- [18] T. Erbe. Building the erbe-verb : Extending the feedback delay network reverb for modular synthesizer use. *41st International Computer Music Conference*, Septembre 2015.
- [19] J. Merimaa and V. Pulkki. Spatial impulse response rendering 1 : Analysis and synthesis. *Journal de L’AES*, Décembre 2005.
- [20] L. Remaggi, P. Jackson, and P. Coleman. Estimation of room reflection parameters for a reverberant spatial audio object. *Journal de L’AES*, Mai 2015.
- [21] S. Tervo, J. Pätynen, A. Kuusinen, and T. Lokki. Spatial decomposition method for room impulse responses. *Journal de L’AES*, Janvier/Février 2013.
- [22] F. Melchior, C. Sladeczek, A. Partzsch, and S. Brix. Design and implementation of an interactive room simulation for wave field synthesis. *Journal de L’AES*, Octobre 2010.
- [23] V. Garcia-Gomez and J. J. Lopez. Binaural room impulse responses interpolation for multimedia real-time applications. *Convention de L’AES*, Mai 2018.
- [24] E. De Sena, H. Hacihabiboglu, Cvetkovic Z., and Smith J.O. Efficient synthesis of room acoustics via scattering delay networks. *IEEE/ACM Trans. Audio Speech Lang. Process.*, 2015.