



Conservatoire National des Arts et Métiers
292 rue de Saint Martin
75141 Paris Cedex 3



Trinnov Audio
2 avenue de l'Europe
94360 Bry sur Marne

Analyse des réflexions précoces

Auteur : Matthieu Berjon
Superviseur : Arnaud Laborie

Stage effectué du 2 Juillet au 24 Août 2007

Table des matières

1	Approche physique	1
1.1	Introduction	1
1.2	Approche Modale	1
1.2.1	Limitations	2
1.3	Approche géométrique	2
1.3.1	Description des méthodes	3
1.3.2	Limitations	3
1.4	Approche statistique	3
1.4.1	Limitations	4
1.5	Conclusion sur les approches physiques	4
1.5.1	Lien entre fréquence de Schroeder et temps direct/diffu	4
1.6	La réponse impulsionnelle	5
1.6.1	Le son direct	7
1.6.2	Les réflexions précoces	7
1.6.3	Le champ diffu	8
1.7	Les paramètres objectifs	8
1.7.1	Temps de réverbération	8
1.7.2	Clarté	9
1.7.3	Largeur apparente de la source	11
1.7.4	Indice de corrélation interaurale IACC	11
1.8	Prospectives	12
1.8.1	Clarté ED	12
1.8.2	Distribution spatiales des premières réflexions SDER	13
1.8.3	Conclusion	14
2	Approche perceptive	15
2.1	Introduction	15
2.2	Effet de précedence - Haas	15
2.3	Masquage temporel	16
2.4	Masquage fréquentiel	17
2.5	Masquage spatial	21
2.6	Hauteur	22
2.7	Localisation	22
2.8	Perception du timbre	22
3	Préconisation pratique	24
3.1	Réflexions latérales	24
3.2	L'effet de spatialisation	25
3.3	Bob Walker [Wal07]	25

3.4	Soren Bech [Bec94] [Bec95]	27
3.5	Les studios, lieux d'écoute, petits volumes	27
3.6	Conclusion	28

Résumé

L'acoustique des salles est un domaine qui existe depuis un grand nombre d'années. Bien qu'il y ait eu de grands progrès il y a quelques années, l'étude complète du champ acoustique en espace clos est relativement récente. Pour la construction de locaux, les acousticiens s'appuient sur un grand nombre de critères objectifs et subjectifs afin d'évaluer globalement le résultat. L'informatique a également apporté une nouvelle dimension dans l'acoustique prévisionnelle. Malgré tout, l'ensemble de ces méthodes apporte une connaissance limitée du domaine et montre leurs limitations dans la compréhension fréquentielle, temporelle et spatiale de plusieurs phénomènes comme la propagation des premières réflexions. Ce rapport vise à faire l'état de l'art concernant la propagation des premières réflexions afin de voir les critères les plus pertinents et leurs limitations.

Approche physique

1.1 Introduction

La propagation acoustique est définie par l'équation des ondes. Celle ci vise à exprimer de façon mathématique le comportement d'une onde dans son déplacement en fonction du temps, de la vitesse du son et de la pression. Ce modèle simple est valable pour des hypothèses précises tel que la propagation s'effectuant en champ libre ¹. Dans notre cas l'étude, de la propagation s'effectue avec des hypothèses plus strictes puisque nous nous intéressons à la propagation en espaces clos. Ceci amène des conditions telles, que l'onde va rencontrer des obstacles de différentes natures changeant ainsi le comportement de l'onde en terme de fréquence et d'espace. Les acousticiens ayant étudié ces changements de comportement ont pu dégager 3 types de phénomènes :

- Les réflexions spéculaires ²
- La transmission
- La diffraction ³

D'autres phénomènes peuvent influencer la propagation d'onde et notamment la température, les courants d'air et l'hygrométrie. Malgré tout, dans la plupart des locaux, on peut considérer que le milieu reste homogène et donc que ces paramètres sont constants.

D'autres paramètres peuvent également rentrer en jeu dans la description de l'onde comme la géométrie de la pièce, l'absorption de l'air et des murs, les types de sources utilisés ... On voit bien la difficulté d'un tel système à résoudre. Afin d'obtenir des résultats de façon efficace et relativement fiables, les acousticiens ont dégagé plusieurs approches permettant par le biais de simplification de tendre à un résultat cohérent et valable dans des utilisations de salles simples. Il est intéressant de faire un point sur ces approches et de voir leur limitations car les indices utilisés de nos jours pour la conception de locaux découlent directement de ces recherches.

1.2 Approche Modale

L'approche modale dans les locaux consiste à résoudre l'équation des ondes (1.1) dans le volume de la salle :

$$\Delta p - \frac{1}{C^2} \frac{\partial^2 p}{\partial t^2} = 0 \quad (1.1)$$

¹c'est à dire que l'onde se déplace dans les 3 axes sans jamais rencontrer d'obstacles

²si la longueur d'onde est petite devant l'obstacle, l'onde incidente de la réflexion possède le même angle que l'onde réfléchie (voir paragraphe 1.3)

³se présente lorsque l'onde rencontre un obstacle et le contourne

La superposition des différentes solutions de l'équation des ondes permet d'obtenir le champ sonore de la salle dans le domaine fréquentiel. Ces solutions correspondent aux modes propres ¹. Cela implique, pour être correct, d'avoir des informations sur les parois et les angles d'incidence des ondes incidentes. Une telle représentation mathématique du comportement acoustique des parois est extrêmement difficile à modéliser. Il varie en fonction de la fréquence, de la taille des matériaux constituant la paroi, des conditions de montage, de la direction d'incidence de l'onde, etc ... Ceci ne prend pas non plus en compte le fait que l'onde puisse rencontrer des obstacles tout au long de son parcours de propagation, de ce fait, le problème est difficilement soluble de façon analytique, même dans le cas de salle à géométrie simple.

Partant de ce constat, la résolution numérique de l'équation des ondes semble être une évidence. Malheureusement, les méthodes numériques actuelles ne permettent pas d'avoir des résultats satisfaisant sur tout le domaine audible dans un délai acceptable. L'utilisation d'une méthode numérique reste gérable uniquement dans le domaine basse fréquence car les modes sont peu nombreux et identifiables (Equation 1.2).

$$f_{l,m,n} = \frac{C}{2} \sqrt{\left(\frac{l}{x}\right)^2 + \left(\frac{m}{y}\right)^2 + \left(\frac{n}{z}\right)^2} \quad (1.2)$$

Où C représente la célérité du son, x, y, z représentent les dimensions de la salle et l, m, n représentent des indices entiers de 0 à n^2 .

1.2.1 Limitations

Comme il a été dit plus haut, la limitation de puissance des processeurs implique de rester dans le bas du spectre audible. Toutefois, définir le domaine basse fréquence est difficile et peut varier selon les cas. Cruz-Barney [Bar99] décrit l'appellation du domaine basse fréquence. Il dit qu'il est possible de classer les salles en fonction de leur dimensions et de la fréquence. On peut considérer qu'une salle est petite si la longueur d'onde à une fréquence spécifique est supérieure ou égale aux dimensions de la salle, de ce fait la densité modale sera faible ³.

1.3 Approche géométrique

L'approche géométrique est une technique de plus en plus utilisée car elle suit l'évolution des puissances de calcul des unités informatiques. Elle permet en haute fréquence de connaître la distribution spatio-temporelle des réflexions ⁴ et de la réverbération tardive. La solution du problème consiste à déterminer des rayons issus d'une source en connaissant les propriétés du milieu et les obstacles. Malgré tout, l'approche géométrique n'est qu'une approximation du phénomène de propagation acoustique. Les hypothèses sont les suivantes :

- Les dimensions des obstacles sont beaucoup plus grandes que les longueurs d'ondes considérées.
- Les irrégularités des surfaces sont inférieures aux longueurs d'ondes.
- Un rayon n'est affecté que par l'absorption de la surface au point d'impact.
- La source est placée loin des surfaces par rapport aux longueurs d'ondes considérées.

Ces hypothèses posées, plusieurs techniques de calculs ont été mises en place pour calculer la l'évolution des rayons au cours du temps :

- Méthode par sources images / rayons
- Méthode des rayons

¹Les modes propres correspondent aux ventres de la salle (axiaux, tangentiels et oblique) de la salle, donc les fréquences favorisées par rapport aux autres

²Ces indices permettent de définir les modes axiaux, tangentiels et obliques

³Ce qui implique d'être de ce fait dans le domaine basse fréquence

⁴Développé au chapitre 2

– Méthode des cônes

1.3.1 Description des méthodes

Sources images : pour une source sonore placée devant une paroi réfléchissante va correspondre une source virtuelle symétrique par rapport à la surface de réflexion (Fig 1.1). Dans le cas d'une paroi réfléchissante dont les dimensions sont beaucoup plus grande que la longueur d'onde de la source considérée, alors la totalité de l'énergie est renvoyée. Dans un cas concret, les parois possèdent souvent un coefficient d'absorption ¹ que l'on applique au signal considéré.

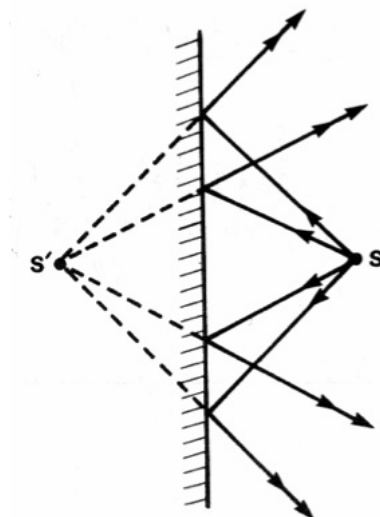


FIG. 1.1 – Principe d'une source image d'ordre 1

Cônes : Pas de référence à ce jour. Méthode non utilisée dans la "gros" logiciel d'acoustique prévisionnelle tel que Catt Acoustic.

1.3.2 Limitations

Plusieurs limitations viennent mettre en doute des résultats valables sur tout le spectre audible. Des soucis peuvent intervenir lorsque la pièce fait intervenir des parois non planes. Ceci a tendance à fausser les résultats. Afin d'atténuer les biais, il faut discrétiser la géométrie de la paroi et d'appliquer un coefficient de diffusion afin de réduire le biais induit. De plus, les logiciels commencent juste à prendre les coefficients de diffusion des parois en compte. Si les aspérités d'un mur sont supérieures à la longueur d'onde considérée alors, les résultats peuvent s'avérer altérés. Il en va de même pour le problème de la diffraction qui est mal interprétée par les logiciels de prédiction. De ce fait, l'acoustique géométrique est réservée à un spectre restreint en milieu de gamme, pour des salles dont les géométries ne sont pas trop tortueuses et facilement dessinable à l'ordinateur.

1.4 Approche statistique

L'approche statistique est née du besoin de caractériser aisément une salle avec des outils simples. Pour ce faire, une des hypothèses importantes pour travailler en acoustique statistique est de considérer

¹Le coefficient varie en fonction de la fréquence mais aussi de l'angle d'incidence, pourtant ce dernier n'est pas souvent pris en compte dans les calculs

que l'on est en champ diffu. Ceci permet d'avoir accès au niveau réverbéré en régime stationnaire. Les autres hypothèses nécessaires au travail en acoustique statistique sont les suivantes :

- Toutes les directions de propagation sont équiprobables
- La densité d'énergie acoustique est la même dans toute la salle

1.4.1 Limitations

Il s'avère que ces conditions sont difficiles à obtenir [Pol88] mais fonctionnent quand même correctement lorsque la salle n'est pas trop absorbante [Bru98]. L'approche statistique n'est justifiée ¹ qu'à partir d'un recouvrement des modes suffisant. Ces travaux menés par Schroeder ont abouti à la fréquence de coupure de la salle ² et s'exprime :

$$f_c = 2000\sqrt{\frac{Tr}{V}} \quad (1.3)$$

Où Tr représente le temps de réverbération et V Le volume de la salle considérée.

Une formule des plus connues de l'acoustique statistique est celle du temps de réverbération donnée par Sabine qui s'exprime sous la forme :

$$Tr = 0.16\frac{V}{A}$$

Avec Tr , le temps de réverbération exprimé en seconde (s). V le volume de la salle considérée et A l'absorption totale de la salle qui s'exprime par $A = \alpha \times S$.

De ce fait on voit que l'approche statistique ne fonctionne pas en basse fréquences (continuum de modes) et pour des salles pas trop absorbantes (condition de champ diffu).

1.5 Conclusion sur les approches physiques

La description des approches ci dessus montre que le savoir faire actuel à beaucoup évolué ces dernières années mais montre également la difficulté à maîtriser les champs sonores en milieux clos. Il est extrêmement difficile de décrire le champ acoustique de façon correcte ³ par le biais d'une seule méthode. L'approche modale est bien trop compliquée à l'heure actuelle pour être utilisée en industrie (trop de paramètres, pas de simplification), l'approche géométrique donne des résultats faux en basse fréquences et l'approche statistique ne fonctionne qu'à partir de la fréquence de coupure et pour des salles pas trop absorbantes.

Ces méthodes d'acoustique prévisionnelle, bien que possédant bon nombre de lacunes, sont adaptées à la réalité du marché et viennent en complément de l'expérience de l'acousticien dans la réalisation de ses préconisations.

Dans le cadre de notre étude, nous ne nous intéressons pas directement à la prévision acoustique d'une salle, mais à constater les défauts d'une salle existante afin de les corriger par le biais de différents traitements numériques. Pour ce cela il faut effectuer des mesures afin de dégager des paramètres pertinents permettant d'identifier ces défauts.

1.5.1 Lien entre fréquence de Schroeder et temps direct/diffu

Comme le montre la Fig. 1.2, les approches physiques dépendent de la fréquence de Schroeder et du temps. Ceci permet de voir là où les différentes approches fonctionnent de manière optimum. Au vu de ce

¹dans le domaine fréquentiel

²aussi appelée fréquence de Schroeder

³de façon spatiale, fréquentielle et temporelle

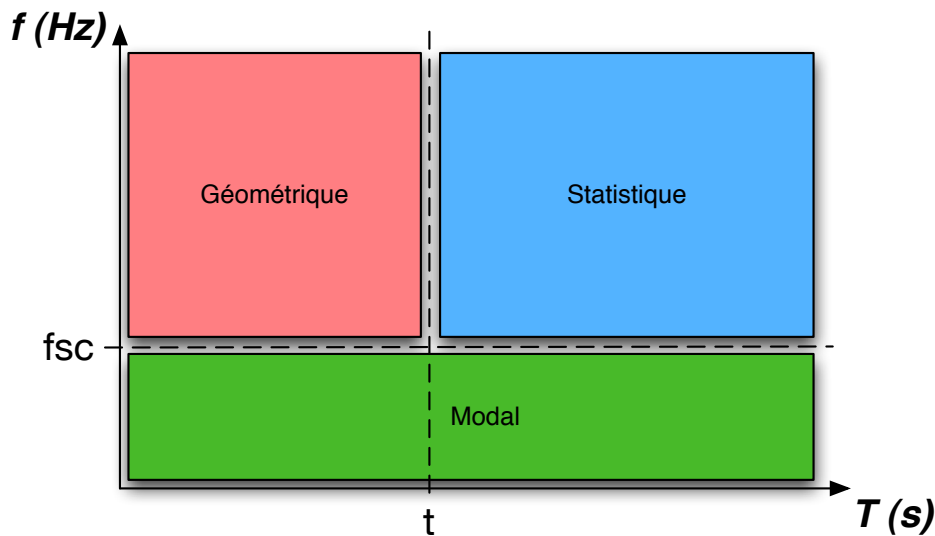


FIG. 1.2 – Zone d'utilisation optimum des différentes approches physiques. Vert : modal - Rose : géométrique - Bleu : statistique - f_{sc} : fréquence de Schroeder - t : temps où intervient la transition direct/diffu

graphique, on est en droit d'imaginer un lien entre le temps séparant réflexions précoces et champ diffu et la fréquence de Schroeder qui permet de savoir où se situe le continuum de modes. Au premier abord on pourrait penser à un lien (équation 1.4) de la forme :

$$t = \frac{1}{f_{sc}} \quad (1.4)$$

Ce lien bien qu'à priori pertinent est trop simpliste pour donner des résultats satisfaisants. En effet la fréquence de Schroeder dépend du Tr^1 et du volume. Si l'on observe le phénomène "avec les mains", pour une salle donnée :

- la fréquence de coupure va descendre si le volume augmente ou si le temps de réverbération augmente. De manière générale plus le volume est grand, moins l'absorption est grande (donc le Tr grand), plus le champ diffu va mettre du temps à s'établir (plus de trajet à parcourir). Donc ils sont plutôt opposés.
- la fréquence de coupure va augmenter si le volume est faible ou si le temps de réverbération est petit. De manière générale, plus le volume est faible (donc Tr souvent petit), plus le champ diffu va s'établir vite (sauf si Tr trop petit), donc la transition réflexions précoces / champ diffu va descendre, tandis que la fréquence de Schroeder va plutôt tendre à augmenter.

1.6 La réponse impulsionnelle

Comme expliqué dans le paragraphe 1.5, nécessiter de dégager un ensemble de paramètres permettant de juger les défauts (et les qualités) d'une salle. L'idéal est d'obtenir une mesure nous permettant d'avoir l'ensemble des informations accessible en lisant directement le résultat de la mesure ou en effectuant

¹Si l'on fait attention, le Tr dépend de la fréquence ce qui implique le f dépend de f . Cela n'empêche pas que le système fonctionne assez bien dans la pratique

un ou plusieurs post traitement sur celle ci. Ceci est partiellement possible ² en mesurant la réponse impulsionnelle.

Qu'est ce que la réponse impulsionnelle ? Elle correspond à la réponse de la salle pour un couple de position émetteur / récepteur. On appelle aussi la réponse impulsionnelle la signature acoustique de la salle. En théorie, la réponse impulsionnelle contient la totalité des informations permettant de caractériser une salle. Toutefois, la lecture brute d'une réponse impulsionnelle (Fig. 1.3) en temps est difficile et ne permet pas de lire directement l'ensemble des informations (temporelles, fréquentielles et spatiales) que l'on souhaite obtenir. Les premières informations qui sautent aux yeux (comme le montre la Fig. 1.3) sont la distribution temporelle des réflexions ainsi que l'amplitude de celle ci. Ceci reste quand même très léger et ne nous permet pas de tirer des conclusions fiable sur les défauts de la salle, ni sur ces qualités. Malgré tout, par le biais de post traitement informatique, nous sommes capables de tirer un ensemble d'indices objectifs nous amenant un certain nombre d'informations dont un des plus connus est le temps de réverbération.

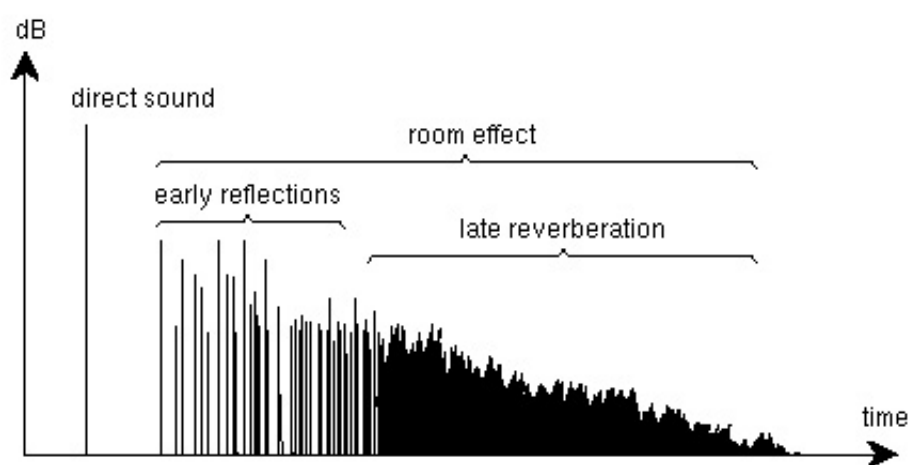


FIG. 1.3 – Echogramme d'une réponse impulsionnelle type

De quoi dépend la réponse impulsionnelle ? Elle dépend d'un grand nombre de paramètres couplés les uns aux autres :

- Du type de source
- De la directivité de la source
- De la géométrie de la salle
- De l'absorption effective de la salle
- De la position du couple source/récepteur
- Du type de récepteur

De ce fait, on peut déduire que la réponse impulsionnelle d'une salle peut facilement évoluer dans l'espace en fonction d'un nombre d'éléments.

La réponse impulsionnelle dans la plupart des salles peut "à l'oeil" se découper en 3 parties distinctes qui sont :

²La réponse impulsionnelle n'est valable que pour une position du couple émetteur/récepteur

- Le son direct
- Les premières réflexions
- Le champ diffu

Le but de ce rapport est de s'intéresser aux premières réflexions qui, comme expliqué par la suite, peuvent aussi se représenter par un découpage temporel. Cruz-Barney [Bar99] a montré que la réponse impulsionnelle d'une salle pouvait se découper plus finement au niveau des premières réflexions et du champ diffu, comme montré dans le Tab. 1.1 :

Plage Temporelle	Formule	nom
0 ms	$\int_{t=0}^{t=5\text{ms}} P^2 dt$	Onde directe
0 - 1.2 ms	$\int_{t=0}^{t=1.2\text{ms}} P^2 dt$	Onde de localisation
1.2 - 20 ms	$\int_{t=0}^{t=20\text{ms}} P^2 dt$	Onde directe étendue partie 2
20 - 40 ms	$\int_{t=20}^{t=40\text{ms}} P^2 dt$	Premières réflexions
40 - 80 ms	$\int_{t=40}^{t=80\text{ms}} P^2 dt$	réflexions tardives
80 - 160 ms	$\int_{t=80}^{t=160\text{ms}} P^2 dt$	Réverbération précoces
160 - ∞	$\int_{t=160}^{t=\infty\text{ms}} P^2 dt$	Réverbération tardive

TAB. 1.1 – Découpage temporel de la réponse impulsionnelle.

Il est maintenant nécessaire de s'intéresser à chaque partie de la réponse impulsionnelle afin d'avoir une vue d'ensemble des différentes parties qui compose le champ acoustique d'une salle pour une position donnée. Par la suite, nous nous intéresserons aux premières réflexions spécifiquement d'un point de vue physique et perceptif.

1.6.1 Le son direct

Le son direct représente le trajet le plus court de la source au récepteur sans aucune réflexion. On peut l'assimiler à la représentation d'une onde se déplaçant en champ libre. Le son direct ne dépend que des caractéristique de la source (directivité, puissance ...) et de la distance source-point de réception (décroissance en champ libre ¹).

1.6.2 Les réflexions précoces

Les réflexions précoces appellées aussi premières réflexions font parties de la première partie de l'effet de salle. Elles représentent un ensemble de contributions en provenance des parois et autres obstacles. Elles correspondent le plus souvent aux réflexions d'ordre 1 ². Ces réflexions sont souvent porteuses d'informations (timbre, taille de source ...) mais peuvent s'avérer préjudiciable dans la compréhension du message sonore [Bec94][Wal07] (echo, masquage, timbre trop prononcé, mouvement de l'image sonore ...). D'autres études psychoacoustique montre également l'importance du stimuli ([Bec94] [Bec95]). Les premières réflexions peuvent influencer le timbre d'un son mais dans des proportions différentes selon la nature de celui-ci (seuil de détection différent selon qu'il s'agit d'un bruit blanc, de la parole, d'un son faisant sens ou non). A partir de ces quelques constations, on comprend la complexité des réflexions précoces puisque celles-ci dépendent de la nature de la source et du récepteur (type de stimuli, directivité de la source ...) et de la géométrie de la salle (géométrie, absorption ...). Malgré tout, d'un point de vu purement physique, on peut estimer l'amplitude d'une réflexion par la formule suivante [AJ01] :

$$I_{\text{relexion}} = I_{\text{source}} - 20 \log_{10}(\text{longueurDuCheminParcouru}) + 10 \log_{10}(1 - \alpha) \quad (1.5)$$

¹Le signal décroît de 3dB par doublement de la distance

²L'ordre 1 correspond à la réflexion sur une seule paroi, et donc celle qui arrive le plus tôt au recepateur après le champ direct

Avec I_{source} , l'intensité de la source de bruit et α le coefficient d'absorption³ de la paroi considérée. La gestion des premières réflexions devient très épineux lorsque l'on a plus d'une source et d'un récepteur. En effet, une source engendre un certain nombre de réflexion sur les parois, une autre source va automatiquement multiplier ce nombre. Dans le même ordre d'idée, lorsque l'on a une unique récepteur fixe, il est possible d'effectuer des changements dans les réflexions puisque l'on cherche à faire une amélioration de façon locale. Par contre lorsque l'on augmente le nombre de récepteur ou du moins la zone où l'on souhaite obtenir une distribution optimum des premières réflexions, alors il devient beaucoup plus compliqué d'agir et d'obtenir des résultats satisfaisant. En effet chaque source est couplée entre elles et à la salle.

Les studios de la BBC ont essayer de mettre en oeuvre une technique de correction acoustique allant à l'encontre des cabines classique LEDE¹. Le but étant d'optimiser la distribution des premières réflexions. Cette technique apellée RFZ² consiste à rediriger toutes les réflexions d'ordre 1 afin qu'elles ne coupent pas la zone d'écoute.

1.6.3 Le champ diffu

Le champ diffu représente la dernière partie de la réponse impulsionnelle et la décroissance du signal. Le champ diffu correspond à une multitude de reflexion provenant de toutes les surfaces de la salle. D'un point de vue modal, le champ diffu représente un ensemble de modes très dense, d'un point de vue statistique, cela représente un ensemble de réflexions provenant de toutes les directions de façon équiprobable.

1.7 Les paramètres objectifs

1.7.1 Temps de réverbération

Le temps de réverbération selon Sabine est un des paramètres objectifs les plus connus en acoustique des salles. Il a été mis au point par Wallace Sabine et permet de mesurer le temps que met un son pour décroître dans une salle. D'un point de vue mesure, on a établi que le temps de réverbération correspondait au temps que met l'énergie sonore pour décroître de 60dB après extinction de la source (Fig. 1.4).

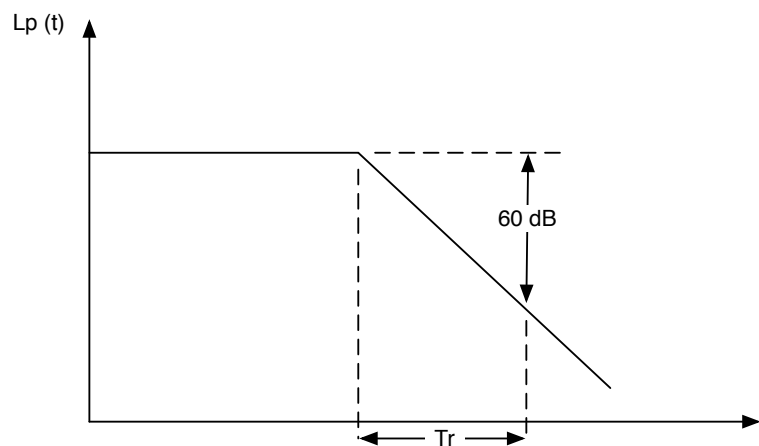


FIG. 1.4 – Mesure du T_r par décroissance du signal

³Varie avec l'angle d'incidence de la réflexion sur la paroi

¹Live End Dead End : absorbant devant l'auditeur, diffusant derrière

²Reflection Free Zone

Methode de mesure

La mesure du temps de réverbation se fait pour une position de couple émetteur / recepneur. L'emetteur se représente régulièrement par un bruit blanc ou un dirac et de façon omni directionnelle. Une fois la mesure effectuée, il s'agit de réaliser une opération de post processing afin d'obtenir le temps de décroissance pour obtenir 60dB. Le résultat final sera un couplage dépendant du système d'émission, de la directivité de la source et de la géométrie de la salle. Il est également possible de calculer le temps de réverbération en fonction de certain paramètre. Plusieurs formules ont été mise au point et la plus connue (car la plus facile à retenir et à calculer) est la formule de Sabine :

$$Tr = 0.16 \frac{V}{A} \quad (1.6)$$

Où V représente le volume de la salle en m^3 et A l'absorption des murs en m^2 , A s'écrivant :

$$A = \sum \alpha S \quad (1.7)$$

Avec S la surface des murs en m^2 et α le coefficient d'absorption de la paroi considérée.

Informations temporelle/spatiale

Le temps de réverbération étant un indice mesuré, le temps obtenu donne des informations sur le comportement général de la salle :

- Absorbant ou non
- Grand ou petit volume

Perception

Le temps de réverbération est quelque chose que l'on mesure relativement aisément à l'oreille. On distingue facilement lorsqu'une salle est mât ou plutôt réverbérante. De manière instinctive on adapte notre écoute en ayant quelques valeur clés en tête.

- Beaucoup de réverbération = mauvaise compréhension d'un locuteur
- Peu de réverbération = excellente compréhension du locuteur mais gêne à l'écoute

Robustesse spatiale et limites

La robustesse spatiale du temps de réverbération est théoriquement bonne, d'un point de vue pratique, le champ diffu peut varier dans d'assez grandes proportions dans certaines salles.

Le temps de réverbération tel qu'il a été expliqué plus haut est souvent difficile à obtenir. Une décroissance de 60dB dans une petite salle est souvent difficile à atteindre et, de plus, il existe des irrégularités importantes dans la décroissance d'un point à un autre de la salle, de ce fait on a tendance à utiliser la mesure d'EDT (Early Decay Time) qui correspond à une extinction de 15dB. Les salles de petits volumes sont plus sujette aux interférences de type diffraction. Il n'est pas aisé (voir impossible) de modifier un paramètre indépendamment des autres et de trouver des résultats simples et efficaces. Il est a noter que le Tr seul, ne constituait pas un indice assez pertinent [Ber62].

1.7.2 Clarté

La clarté dépend de la répartition de l'énergie de la réponse impulsionnelle. La clarté correspond à un rapport entre l'énergie précoce et l'énergie tardive. La clarté vise à évaluer la précision d'écoute d'un instrument.

$$C_x = 10 \log \left(\frac{\int_0^{x(ms)} P^2 dt}{\int_{x(ms)}^{\infty} P^2 dt} \right) (dB) \quad (1.8)$$

Les éléments pouvant venir perturber le calcul de la clarté sont la non linéarité de la distribution des réflexions (temporelle, spectrale et spatiale) et l'émergence de pics par rapport à d'autres (les échos par exemple).

Méthode de mesure

La mesure de la clarté se déduit d'une mesure de réponse impulsionnelle. De ce fait la clarté s'obtient en mesurant l'énergie contenu dans un interval de temps par rapport à l'énergie situé après cet interval¹ de la réponse impulsionnelle (le résultat étant mis sous la forme de décibel).

Informations temporelle/spatiale

La clarté étant l'intégration de l'énergie, on récupère uniquement des informations entre le taux de premières réflexions et le taux de champ diffu. Aucune donnée de la clarté permet de connaître la distribution temporelle, fréquentielle ou spatiale de la réponse impulsionnelle.

Perception

La clarté est très liée au temps de réverbération. Même si la clarté ne tiens pas compte de la géométrie de la salle ou du volume, il est facile de comprendre que certaines valeurs clés permette de le relié au Tr.

- $< 0dB$: Il y a plus d'énergie en champ diffu qu'en premières réflexions (temps de réverbération plutôt long).
- $= 0dB$: Il y a autant d'énergie en champ diffu qu'en premières réflexions. en général c'est un bon compromis pour le temps de réverbération (va dépendre du type de source présente dans la pièce)
- $> 0dB$ Premières réflexions beaucoup plus énergétique que le temps diffu (pièce plutôt mat et temps de réverbération plutôt court).

Robustesse spatiale et limites

A priori la robustesse spatiale de la clarté est homogène dans le cas d'une salle relativement simple et symétrique. De plus le fait qu'on l'intègre sur un temps t permet de pondérer un peu certaine irrégularité dans la distribution temporelle des premières réflexions. Malgré tout, comme il a été spécifié plus haut, une trop grande irrégularité dans la distribution temporelle, fréquentielle et spatiale des premières réflexions et l'émergence de pics peuvent venir perturber le calcul et donner des résultats qui ne seront pas en adéquation avec l'allure générale de la salle.

Prise en compte de la multidiffusion

La clarté se mesure pour une source et un point de réception, donc tiens compte de la géométrie de la salle. De plus la source d'émission rayonne omnidirectionnellement et donc ne tient pas compte de la directivité particulière d'une source en fonction de la fréquence.

Pour palier au problème de distribution il faut mettre en place des fonctions de pondération permettant de quantifier et de régulariser la distribution des réflexions. Pour le problème d'écho il faut évaluer la nuisance de l'écho produit (écart de 50ms minimum) et la nature de cet écho (coefficient d'écho direct de Niese, la tangente ajustée de Bolt et Doak, le rapport de diffusion temporelle de Kuttruff pour l'écho tonal et flottant). Il faut toutefois prendre cet indice avec des pincettes, deux salles peuvent avoir le même indice de clarté et pourtant donner une perception différentes de leur acoustique (du à la distribution des réflexions).

¹Rapport entre réflexions précoces et champ diffu.

1.7.3 Largeur apparente de la source

La largeur apparente d'une source dépend du niveau relatif des premières réflexions latérales. La largeur est donc caractérisée par un rapport d'énergie des premières milliseconde en fonction d'un angle (angle d'arrivée des premières réflexions) et l'énergie globale de la salle. La mesure de l'efficacité latérale en pratique s'effectue à l'air de deux microphones : l'un bidirectionnel et l'autre omnidirectionnel.

$$LF = \frac{\int_0^{80\text{ms}} p^2 \times \text{Micro_Bidirectionnel} dt}{\int_0^{80\text{ms}} p^2 dt} \quad (1.9)$$

Note : La largeur apparente de source n'est pas utilisé par Beranek dans ces études des opéras et salle de spectacles.

Méthode de mesure

La largeur apparente s'obtient en utilisant à l'émetteur une source omni. Au point de réception, le système s'apparente à un couple MS. La différence réside en l'utilisation d'un microphone omni en lieu et place du microphone à directivité cardioïde. Ce système permet par rapport à la clarté d'identifier les contributions latérales des premières réflexions (premières réflexions qui permette d'identifier la source plus facilement). Dans le cadre du post traitement, il est convenu de prendre le temps d'intégration 5ms après le son direct pour ne pas être perturbé par ce dernier (dans la pratique, la gêne générée par le son direct reste négligeable).

Perception

La largeur apparente de source tient compte du taux de premières réflexions latérale en fonction des caractéristique des micros sur l'énergie totale des premières réflexions.

Qu'est ce que l'on peut trouver ? Un rapport non mis en dB.

- $LF \simeq 0$: peu de contribution latérale, source plutôt large, moins bien définie.
- $LF \simeq 1$: égalité des contribution latérales par rapport au global, source correctement définie.
- $LF > 1$: abondance des contribution latérale, bonne définition de la taille de source, Tr sûrement peu élevé.

Robustesse et limite

Selon où se fait la mesure, les contributions latérales ne doivent pas réellement se compenser. Si près d'un mur latérale, contribution très forte de ce mur et disparition des contributions de l'autre murs. En fait il y a peut être compensation donc pas de souci particulier : Il faut trouver un moyen de vérifier.

Prise en compte de la multidiffusion

Non pas de prise en compte de la multidiffusion. Mesure spatial du champ sonore mais appliqué à une source mono. Le fait d'avoir un nombre plus grand de source doit sûrement perturber le champ (onde constructives et opposition de phase au point de réception).

1.7.4 Indice de corrélation interaurale IACC

L'IACC permet d'obtenir des informations en regardant les énergies qui arrivent à chaque oreille.

$$IACC = \frac{\int_{t_1}^{t_2} P_l(t) P_r(t + \tau) dt}{\sqrt{\int_{t_1}^{t_2} P_l^2(t) dt \int_{t_1}^{t_2} P_r^2(t) dt}} \quad (1.10)$$

P_l correspond à la pression à l'oreille gauche, P_r la pression à l'oreille droite et τ la différence de temps de parcours entre les deux oreilles. Dans le cadre du présent stage, il convient de calculer l' $IACC_e$

qui correspond à un temps d'intégration entre 0 et 80 ms. Il s'avère que seules les contributions latérales jusqu'à 1.6kHz contribuent à la sensation de largeur de la source [MISM94].

Méthode de mesure

Mesuré à l'aide d'une tête artificielle. Il s'agit d'utiliser une fonction de corrélation pour intégrer entre un temps t_1 et t_2 la pression reçue à l'oreille gauche et droite en fonction du retard interaural. Il s'agit donc de mesurer la pression obtenue à chaque oreille et d'effectuer un post traitement afin d'avoir la fonction d'intercorrélation normalisée.

Limites

Le souci de l'IACC est de faire appel à une composante binaurale, dont le biais n'est pas évident à mesurer et varie dans de fortes proportions entre individus.

1.8 Prospectives

Les critères cités ci-dessus sont ceux se rapprochant le plus de notre propos, à savoir les premières réflexions. Bien que trouvant leur entière légitimité en acoustique physique traditionnelle, ces critères sont malgré tout assez peu pertinent en traitement du signal. En effet, l'ensemble de ces critères prennent en compte l'énergie des premières réflexions par rapport à l'énergie du champ diffu. Ceci ne donne qu'une information partielle sur le champ acoustique et ne permet pas de quantifier correctement la direction de provenance, l'intensité, les fréquences des premières réflexions. Les meilleures conditions d'écoute se font en champ libre car seul le son direct est porteur d'information et celui-ci n'est perturbé par aucune autre réflexion. Les premières réflexions et le champ diffu nous apportent quant à eux des informations sur la dimensions de la salle ainsi que sur certains aspects de la source¹.

Ces constatations faites, il serait intéressant de modifier les critères existants pour les rendre pertinent par rapport à notre approche. Il est à l'heure actuelle improbable d'avoir une description parfaite du champ sonore, ceci dans un temps acceptable pour une utilisation industrielle. L'idée principale est, dans un premier temps de comparer l'énergie des premières réflexions par rapport au son direct² afin de constater l'éventualité de dégradation du front d'onde au point d'écoute. Dans un deuxième temps il est intéressant de regarder l'énergie des premières réflexions en fonction de la direction.

1.8.1 Clarté ED

La notion de clarté comme expliqué au paragraphe 1.7.2 vise à donner une information sur la "précision" du son. Pour ce faire, on regarde le rapport d'énergie entre les premières réflexions et le champ diffu. Ceci ne fonctionne évidemment que si la condition de champ diffu est respectée. Si elle est respectée cela implique que le champ diffu représente une homogénéité des réflexions dans leur direction de provenance et en intensité (équiprobabilité). Dans ce cas le champ diffu sera porteur d'information sur la taille de la salle uniquement et n'apportera pas de modification de timbre sur la source³. L'indice de clarté dans sa forme originale n'est pas donc pas des plus pertinents. Le chapitre sur la perception met en évidence l'importance que les premières réflexions sont susceptibles de modifier de façon importante le timbre, la taille de l'image, la stabilité de la source ... Pour pouvoir palier à ce genre de problème, des études ont été menées en comparant l'intensité des réflexions par rapport au son direct afin de déterminer les seuils d'intensité maximum que devait avoir les réflexions.

¹puisque ces réflexions perturbent la source, elles contribuent à lui donner de nouvelles dimensions et une nouvelle position.

²et non au champ diffu

³Éventuellement sur la taille de la source.

La démarche ci-dessus implique de faire des tests perceptifs sur chaque réflexion. Notre but consiste en une démarche différente qui par le biais de ce paramètre avoie une idée de la différence de niveau entre les premières réflexions et le son direct : la clarté ED (Early on Direct). Pour cela on peut suivre la démarche suivante :

$$C_{ED_x} = 10 \log \left(\frac{\int_{0.01(ms)}^{x(ms)} p^2 dt}{\int_{0(ms)}^{0.01(ms)} p^2 dt} \right) (dB) \quad (1.11)$$

La lecture en décibel permet de travailler avec chiffre plus parlant qu'un petit rapport. Le x représente donc le temps d'intégration de l'énergie et permet donc de travailler par bande de temps. Ces périodes si nous devons les définir par période pertinente, le mieux est de se référer au tableau 1.1. L'interprétation peut se faire en fonction de ces périodes qui vont correspondre à l'étude perceptive menée au chapitre 2. Il en ressort les idées suivantes :

- 0 - 1.2 ms : Détection des mouvements de l'image fantôme
- 1.2 - 20 ms : Modification du timbre
- 20 - 60 ms : Modification taille de la source (effet d'espace)
- 60 - 80 ms : Risque d'échos
- 80 ms et + : Champ diffu, critère inadapté

Ceci peut permettre de façon rapide, de connaître le comportement de l'image sonore et du timbre. Une autre méthode pourrait consister à emmagasiner une base de donnée de gabarits classés selon les types de salles et à les comparer à la réponse des premières réflexion en attribuant un écart acceptable.

1.8.2 Distribution spatiales des premières réflexions SDER

Le deuxième critère s'apparente au critère de la largeur apparente de source. Le principe est voir la contribution des premières réflexions dans l'espace au(x) point(s) d'écoute. Le SDER (Spatial Distribution of Early Reflections) consiste en un ou plusieurs points à mesuré à l'aide d'un micro la provenance des réflexions. Pour ce faire le doit être doté d'au moins 3 capsules omni (pour détecter l'arriver des réflexions sur le plan horizontal) et d'une référence sur l'orientation du micro¹. On sait que les premières réflexions seront situé au maximum dans les 80 premières millisecondes et que les temps de parcours les plus court correspondent aux fronts d'ondes directes de chaque enceintes (si on mesure chaque enceinte séparément). Dans notre cas on cherche à connaître la répartition des réflexions au point de mesure à un instant donné. Ceci pourrait permettre d'obtenir des informations sur la robustesse de l'image sonore.

Limites

Un des premiers problèmes est dans la mesure. La mesures s'effectue haut-parleur par haut-parleur, ce qui ne représente pas le champ sonore lorsqu'il y a plusieurs sources. La mesure doit à priori s'effectuer avec un même signal large bande sur l'ensemble des enceintes.

Si les haut-parleurs ne sont pas à la tous à la même distance du point d'écoute (Fig. 1.3) il devient compliqué de distinguer les réflexions des enceintes les plus proches avec le son direct des enceintes les plus éloignées. La solution la plus simple et la moins couteuse est l'utilisation de ligne à retard et supposer que le système d'écoute est bien aligné en temps. La distinction des réflexions avec les fronts directs semble nécessaire afin de ne pas prendre en compte ces derniers dans les calculs. Il est évident également que lorsque l'on considère le sweet spot dans une zone et non plus sur un point, la contribution des réflexions sera plus importante à un endroit qu'à un autre. Il pourrait être intéressant de quantifier le phénomène et de savoir jusqu'à quel point le sweet spot reste viable dans l'espace.

¹permettant de distinguer l'avant de l'arrière, la gauche de la droite, en résumé d'identifier le micro à un auditeur au point d'écoute.

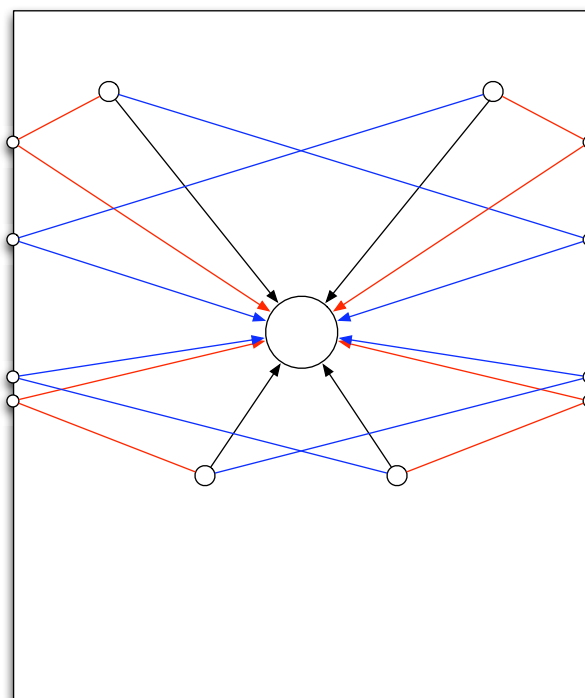


FIG. 1.5 – Représentation simple du champ sonore. Noir : front direct - Rouge et bleu : réflexions

1.8.3 Conclusion

Ces deux nouveaux critères semblent adaptés à la notre problématique, toutefois ils demandent à être validés par une batterie de tests. Si la validation des critères est satisfaisante, alors l'étude de valeurs clés par types de salles semble être une suite logique aux investigations.

Approche perceptive

2.1 Introduction

Comme vu plus haut, les différentes approches physiques impliquent de nombreuses simplifications par rapport à la réalité du champ acoustique. Malgré ces simplifications, l'idée d'effectuer des corrections par le biais de systèmes actifs dans un laps de temps acceptable pour une utilisation industrielle semble encore poser des soucis. Il faut garder à l'esprit que dans une salle, du point de vue de la mesure, tout un ensemble d'éléments peuvent venir perturber la réponse impulsionnelle (changement de trajet des réflexions, déplacement des points de mesure ...), qui sont difficiles à modéliser et à prendre en compte avec une approche physique du champ sonore. Toutefois, ces changements ne seront peu ou pas perçus par l'auditeur à cause des effets de masquages (temps d'intégration de l'oreille principalement). En adoptant une approche perceptive par dessus les modèles physique, il est possible de réduire les temps de calculs. Le but de ce chapitre est de voir les types de masquages et comment fonctionne l'oreille afin d'observer les éléments les plus pertinents. L'oreille amplifie naturellement les fréquences situées entre 4000 et 5000Hz (dimension du pavillon et du conduit auditif). Le système auditif n'a pas une écoute linéaire et les différents travaux menés par Fletcher et Munson en 1933 ont fait apparaître les courbes de sensibilité de l'oreille appelé courbe isosonique (Fig. 2.1).

L'audition variant avec la fréquence, on se doute donc que les effets de masquages temporels et fréquentiels qui dépendent de l'intensité n'auront pas le même "poids" les uns par rapport aux autres (Fig. 2.9).

2.2 Effet de précedence - Haas

L'effet de précedence¹. Cet effet vise à montrer les limites de fusion de l'oreille et donc la limite de perception de l'écho. Helmut a travaillé sur des signaux de parole. Il s'avère que la perception d'un echo va varier d'un stimulus à un autre. De ce fait ainsi que du fait que ce soit de l'ordre du perceptif, les valeurs trouvées dans la perception d'un echo va varier en fonction de plusieurs paramètres. Haas dit que l'on perçoit un écho entre 2 fronts d'ondes si ceux-ci sont espacés d'environ 50ms ou plus [Haa72]. Au cours de ses travaux, il a abouti à quelques lois empiriques sur les fronts d'onde en terme de temps et d'intensité :

- Temps entre son direct et 1ere réflexion : 32 ms \Rightarrow Le deuxième front d'onde ne doit pas dépasser 5dB par rapport au son direct pour ne pas devenir un echo
- Temps entre son direct et 1ere réflexion : 15 ms \Rightarrow Le deuxième front d'onde ne doit pas dépasser 10dB par rapport au son direct.

Si l'on prend le cas d'un même signal émis par plusieurs sources avec un léger décalage temporel, alors l'oreille fusionne les sources et ne perçoit qu'un seul signal dont la localisation sera au niveau de la source

¹Dit aussi effet Haas, venant de son Helmut Haas, scientifique allemand ayant identifié ce phénomène

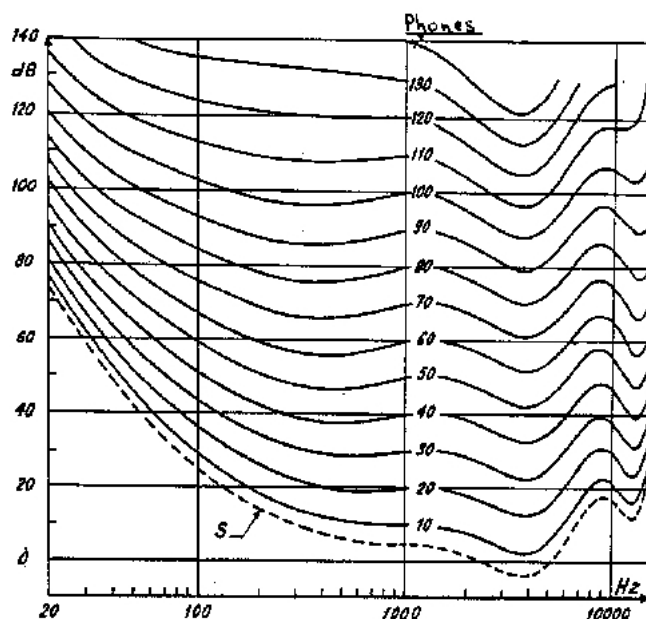


FIG. 2.1 – Courbes isoniques

émettant le premier front d'onde. Il faut quand même garder en tête que ces principes ne fonctionnent que pour les sons riches en transitoires et dépendant de surcroît du décalage temporel entre les deux sources. Si les stimuli sont des clics, alors le temps entre les deux sources avant de percevoir un écho est fortement réduit.

Selon Litovsky [LCYG99] l'effet de précedence varie avec l'individu. Un exemple est pris pour la détection d'un écho avec un clic pour stimulus. Le stimulus est diffusé en boucles (5 boucles). Il constate que la détection de l'écho augmente de façon asymptotique pendant un certain nombre d'essai avant de trouver un équilibre. Ceci suppose que l'auditeur adapte sa perception au fil des tests et démontre un apprentissage au niveau de la perception des échos.

2.3 Masquage temporel

Loi du premier front d'onde Cremer [Cre48] a étudié dans les premiers, dans le cadre d'une diffusion stéréo, la diffusion d'un signal cohérent non périodique. Tant que les deux HP diffuse le signal en même temps et au même niveau, l'événement auditif apparait au centre. Si un des signaux possède un délai qui augmente de façon continu à partir d'un délai de 0, la direction de de l'événement sonore bouge vers le haut parleur émettant en premier ([Bla01] et paragraphe Loi empirique du 1er front d'onde en temporel). On distingue malgré tout deux périodes "critiques", si le délai est compris entre $630 \mu s$ et $1ms$, le signal va vers le haut parleur émettant le front d'onde en premier. Si le délai augmente au delà de $1ms$, alors le signal reste à peu près stable.

Au delà de cette limite, différents phénomènes peuvent apparaitre. Dans le cadre d'un système stéréo et $1ms$ de décalage entre les deux sources, l'image fantôme passe du coté du haut parleur ayant émis le signal en premier. Cette perception de l'image fantome dure jusqu'à $5ms$ de décalage entre les deux haut-parleurs pour des sons très brefs (pour une durée de 30 à $40ms$ environ). Pour des décalages supérieurs à $5ms$, les signaux sont entendus séparément. Pour un décalage de $70ms$ et plus, la perception d'un écho

franc est totale. Si le système est redondant¹ alors l'écho est virtuellement augmenté par l'auditeur de quelques millisecondes (pour des stimuli de type clic). Blauert [Bla01] a mis en évidence la notion d'écho en fonction du temps et de la direction des signaux (Fig. 2.2).

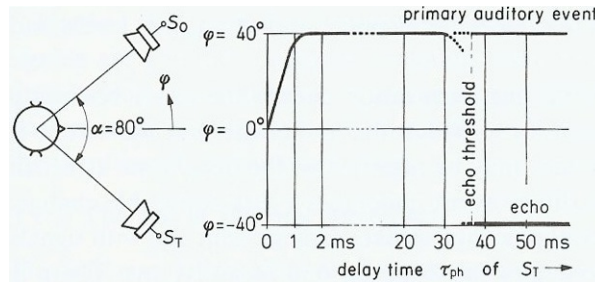


FIG. 2.2 – Chemin des événements audio en fonction des délais des réflexions

Post masquage De tous les effets temporels inclus dans le mécanisme d'écoute, le post masquage possède la plus grosse influence dans le codage des systèmes audio numérique [ZZ91]. Il résulte d'un arrêt progressif de l'effet de masque. Effectivement, l'effet de masquage ne disparaît pas immédiatement avec le son masquant. Tout d'abord à quoi correspond l'effet de post masquage ? Il se produit en présence de deux sons, le premier étant le masquant, le second le masqué. Lorsque qu'un son dure un certain temps, alors il produit un effet de masquage d'une durée relativement courte lors de son arrêt. Le post masquage dépend donc de la durée du son masquant (Fig. 2.3).

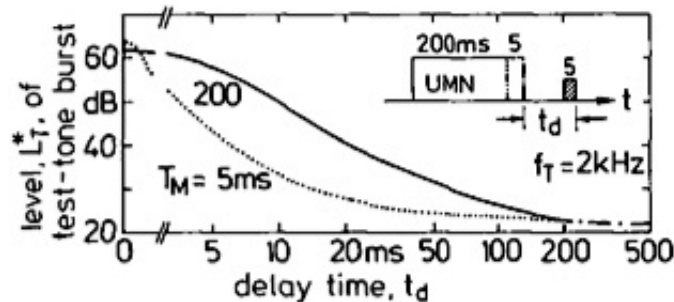


FIG. 2.3 – Effet du post masquage dans le cas d'une séquence de 5ms (pointillé) et 200ms à 2kHz. La fin du son masquant est présenté sur l'axe des abscisses par les traits parallèles.

La décroissance du masquage (par le son masquant) n'est pas linéaire et dépend de l'intensité du son masque (Fig. 2.4).

2.4 Masquage fréquentiel

Haas a démontré que le masquage dépendait du temps mais également des niveaux des signaux. Si une réflexion possède une amplitude faible en regard du son direct, alors la réflexion ne sera pas perçue par l'auditeur. Si l'on augmente le niveau de cette réflexion, alors elle devient audible et se manifeste par une modification du timbre, de l'intensité ou de la taille de la source ([BMB01]). Si la réflexion dépasse un certain niveau, alors on perçoit un écho. Il a été montré que le niveau nécessaire pour reproduire une

¹Le stimulus est diffusé en boucle avec même décalage temporel entre les deux haut-parleurs identique à chaque boucle

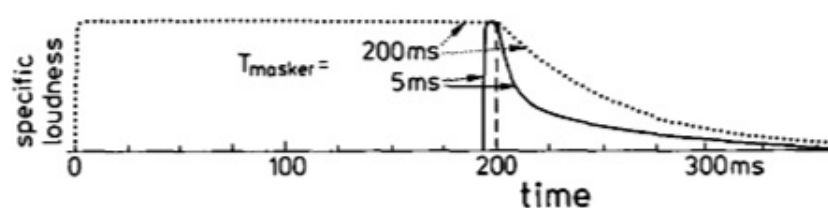


FIG. 2.4 – Graphique similaire pour montrer l'importance en intensité

réflexion d'une pièce, il fallait que celle-ci soit au moins 6dB supérieur au seuil de masquage des réflexions [Ser61]. Donc pour résumer, si le niveau de la réflexion est en dessous du seuil de détection en tant que réflexion unique (echo) alors, la ou les réflexions peuvent potentiellement affecter le timbre, le volume ou la taille de la source.

Buchholz définit le masquage simultané comme le post masquage et le masquage non simultané qui correspond au pré masquage. Par la suite il est dit que le pré masquage est finalement négligeable surtout lorsque le signal arrive de plusieurs directions différentes. Les post et pre masquages sont définis à l'aide des bandes critiques. Pour obtenir ces bandes critiques (qui donnent les niveaux pour lesquels un signal sera masqué dans une certaine bande de fréquences Fig. 2.5).

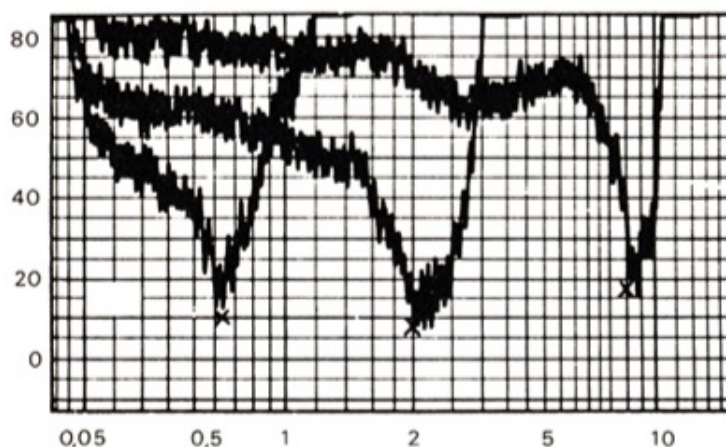


FIG. 2.5 – Seuil d'effet de masquage pour 3 sons purs "masquant" de fréquences 630, 2000, et 8000Hz. Abscisse : fréquence (kHz) - Ordonnée - Niveau de pression (dB). Seuil à partir duquel un son va être perçu. Se reporter à Fig. 2.9.

Méthode d'obtention On choisit un son pur d'une fréquence donnée que l'on cherche à masquer par un autre son pur. Le son masqué est diffusé à très faible niveau afin de n'exciter qu'un seul filtre. On relève le niveau minimum du son masquant permettant de masquer le signal. On obtient ainsi la forme du filtre pour toutes les fréquences centrales (celle du signal) en inversant les courbes de sélectivité de fréquence psychoacoustique. Une autre méthode consiste en l'utilisation d'un filtre de type "Gamma tone" (Fig. 2.6) qui se rapproche beaucoup de l'écoute naturelle [Sla93].

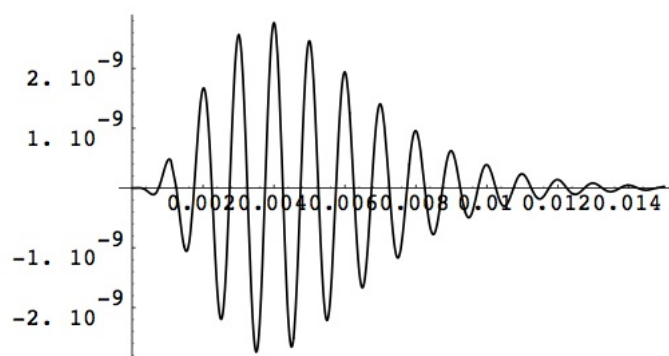


FIG. 2.6 – Réponse d'un filtre Gamma tone

Notion de masquage partiel Le masquage partiel est l'utilisation également de deux sons. Il s'agit en fait de réduire l'impact d'un stimuli à l'aide d'un autre à différentes fréquences. Ceci permet également de réduire l'effet de post masquage dans une bande de masquage relativement fine lorsque le deuxième stimuli y est présent. Malheureusement la fonction de ce processus n'est pas encore bien compris à ce jour.

Propriété de masquage [BMB01] montre que le seuil de masquage descend de façon linéaire lorsque le délai augmente (Fig. 2.6). Ainsi, Il est également possible d'observer le niveau global du son en fonction des angles d'incidence du son direct et d'une réflexion (Fig. 2.7). Deux configuration ont été testées :

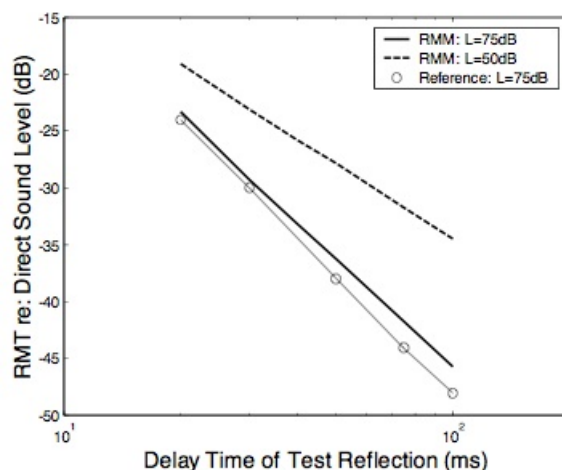


FIG. 2.7 – Seuil de masquage en fonction du temps

- Le son direct fixé pour un angle d'incidence 0° et l'angle d'incidence de la réflexion variant.
- La réflexion fixée pour un angle d'incidence 0° et l'angle d'incidence du son direct variant.

On s'aperçoit que le plus gros écart existe lorsque les angles d'incidence du son direct et de la réflexion sont équivalentes. Ceci peut malgré tout être altéré à cause des HRTFs.

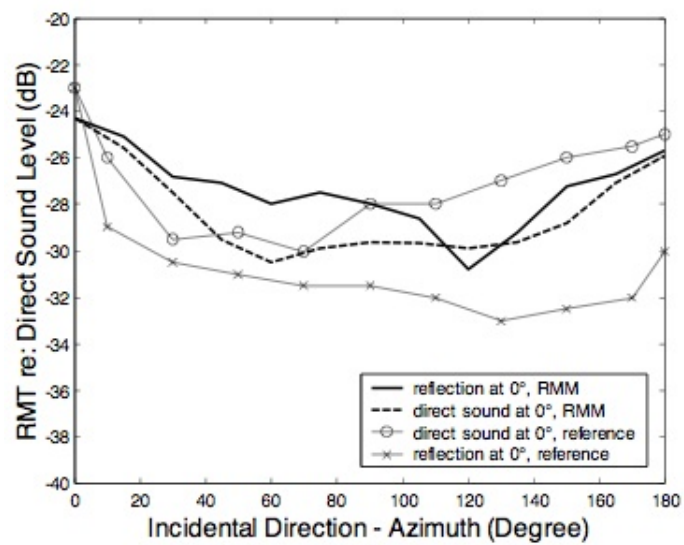


FIG. 2.8 – Dépendance du niveau global en fonction des angles d'incidences du son direct et d'une réflexion

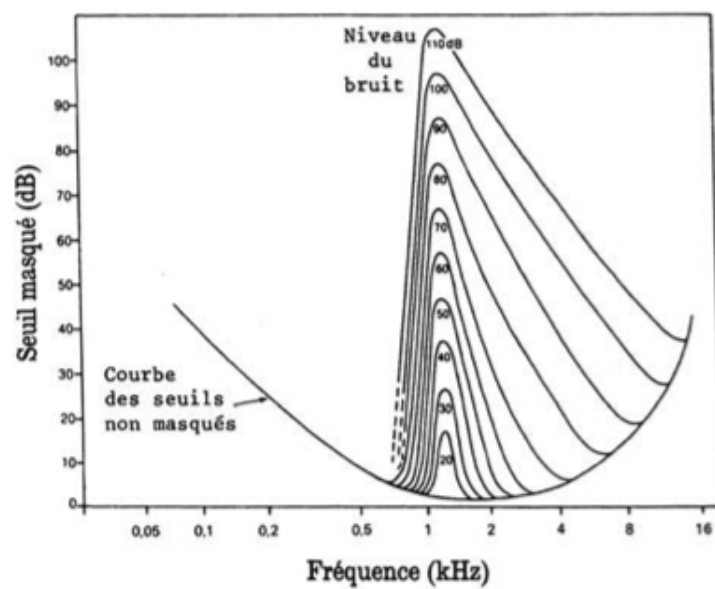


FIG. 2.9 – Seuil de masquage en fonction de l'intensité de la fréquence.

2.5 Masquage spatial

Toole et Olive([OT89]) s'accordent à dire que les réflexions agissent directement sur la perception du timbre et de l'espace. L'étude porte sur la détermination de seuil en fonction du niveau relatif par rapport au son direct. Toole dit que dans beaucoup de cas, on cherche à diminuer les réflexions mais que certaines permettent d'apporter un confort d'écoute plus grand (notamment dans l'intelligibilité de la parole). Leur système de mesure simule 4 réflexions (4 haut-parleur) :

- A 0° horizontal et 0° vertical (en face de l'auditeur)
- A 0° horizontal et 60° vertical
- A 65° horizontal et 0° vertical
- A 115° horizontal et 0° vertical

Utilisation de signaux plus pertinent pour montrer les effets (impulsions et bruits). Les premiers tests montrent pour un auditeur le seuil absolu de détection pour 2 réflexions arrivant respectivement du côté et du plafond. Les stimuli utilisés étaient des impulsions et des bruits roses et l'expérience a été réalisée en chambre anéchoïque.

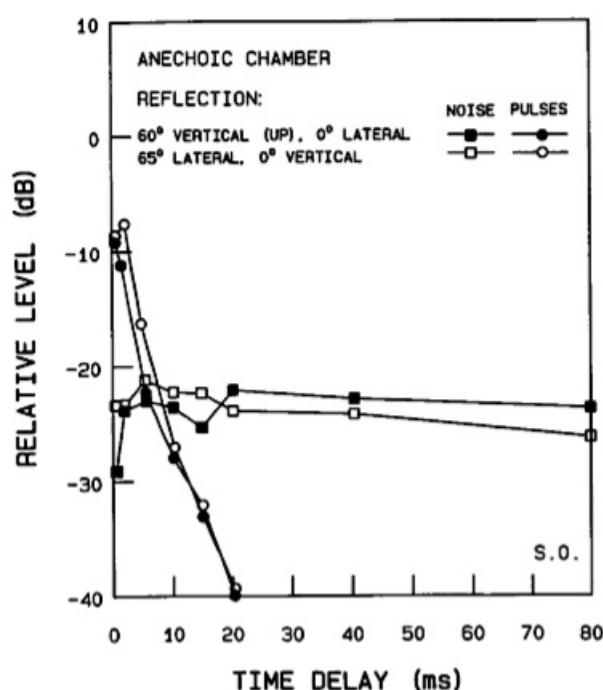


FIG. 2.10 – Détection de seuil absolu en fonction de la direction de provenance de 2 réflexions

Le graphique indique les différences induites par les directions de provenance sont faibles. Par contre la différence entre les deux stimuli est énorme. L'utilisation d'une impulsion montre que le niveau par rapport au son direct est très élevé mais que la sensation de la réflexion décroît très vite, tandis que dans le cadre d'un bruit rose le niveau par rapport au son direct plus faible mais reste constant dans le temps. Il trouve par contre que la direction de provenance a une importance sur le masquage. En effet une réflexion arrivant de la même direction que le son direct pourra être plus forte de 5 à 10 dB avant d'être détecté par l'auditeur comparé à une réflexion arrivant de côté. Les résultats montrent également que les réflexions latérales ont une influence plus grande sur la largeur de source alors que les réflexions venant du sol et du plafond ont apparemment une interaction sur le timbre (recoupement avec [Bec95]). Si l'on rajoute de la réverbération aux réflexions simulées alors, cela a pour effet de baisser le seuil de détection tant que le délai est au-dessus de 20ms

2.6 Hauteur

Définition La hauteur tonale d'un son pur est l'attribut perceptif pour lequel il est possible d'apparier, par ajustement de fréquence, un autre son pur différent par le niveau d'intensité (Demany).

Résolution fréquentielle du système auditif Le seuil différentiel de fréquence exprimé en valeur relative ($\frac{df}{f}$) varie avec la fréquence de référence selon une courbe en U avec un minimum autours de 1500Hz.

Stevens en 1935 a constaté plusieurs phénomènes lorsque l'on élève l'intensité d'un signal :

- élève la hauteur tonale des sons de fréquences élevées
- Laisse à peu près inchangée la hauteur tonale des sons de fréquences moyennes
- Diminue la hauteur tonale des sons de fréquences basses.

A ceci, il y a plusieurs facteurs à prendre également en considération comme la diplacousie binaurale qui correspond à une variation de la perception de la hauteur entre les deux oreilles. Cette différence peut aller jusqu'à 3%¹ La perception de la hauteur n'est possible que pour des sons d'une certaine durée et ceci varie avec la fréquence (Tab. 2.1).

300 Hz	1000 Hz	3000 Hz
27 ms	16 ms	11 ms

TAB. 2.1 – Temps de détection de la hauteur pour trois fréquences.

Hauteur et masquage Un bruit blanc sur un son pur élève sa hauteur tonale. Elle sera d'autant plus élevée que la fréquence du son masqué sera élevée et que le masquant sera fort. La différence avec ou sans masquage peut atteindre 1/4 de ton. Le même phénomène est ressenti pour le cas d'un bruit passe bas dont la fréquence de coupure est légèrement inférieure à celle du son masqué. Il est également possible de réaliser du masquage entre deux fréquences pures. Ceci est possible si la fréquence du son masquant est inférieure à celle du son masqué. Deux possibilités s'ouvrent alors :

- Quand les deux sons ont le même niveau le masquant élève d'autant plus le masqué qu'il est proche fréquentiellement
- Quand les deux sons sont dans un rapport entier de fréquence (1 à 2, 1 à 3) et que le masquant a un niveau bien plus fort, la relation de phase entre les deux sons joue un rôle important (plus de 1/2 ton) sur la hauteur tonale

Evidemment la probabilité d'une écoute musicale ou audiovisuelle avec des sons purs uniquement est assez faible. De ce fait ce dernier point n'est pas de la plus grande pertinence.

2.7 Localisation

Voir rapport de Gaëtan Parseihian sur les lois de panoramique.

2.8 Perception du timbre

La perception du timbre d'un point de vue perceptif est assez floue. Le timbre décrit les caractéristiques d'un son permettant de le différencier d'un autre son possédant la même note². D'un point de vue physique, le timbre se définit par le contenu spectral du son, sa distribution modale. D'une certaine

¹Ce qui correspond à une variation de 1/4 de ton d'un point de vue musical.

²dans le cas d'un son musical.

façon, le contenu spectral fait appel à la notion d'enveloppe du signal. Si il y a des fréquences propres émergentes, de distribution de pics (formants) ou des réjections (trous dans le spectre) alors le timbre va être "altéré". Le problème est que la source possède un timbre intrasèque mais va dépendre d'un grand nombre de paramètres. Tout d'abord, le signal peut passer dans une chaîne électroacoustique qui aura pour conséquence de modifier le timbre, puis si la source émet dans un espace clos, un effet de salle va s'ajouter au signal modifiant par conséquent sa réponse en fréquence et pour terminer le récepteur (que se soit l'oreille ou un microphone) va modifier le spectre car l'oreille ne possède pas une réponse linéaire en fréquence et les micros sont rarement linéaires également dans tous le spectre audible (surtout en prise de son). La définition du timbre est difficile à caractériser par manque de vocabulaire. D'un point de vue plus physique et en rapport avec les réflexions précoces, la modification du timbre va également dépendre du type de sources [Bec94] (bruit blanc, parole) et du nombre de réflexions (isolée ou groupées).

Préconisation pratique

3.1 Réflexions latérales

Travaux de barron : De manière perceptive, on peut constater les choses suivantes (ne dépendant pas nécessairement de la géométrie de la salle ou du volume). Mise en oeuvre de l'essai : 4 voix avec un atténuateur pour les réflexions, une réverb pour le champ diffu délayé de 100ms afin de ne pas chevaucher les 1ere réflexions. La réverb est diffusé sur 4 HP placés symétriquement par rapport à la tete de l'auditeur (arrangé pour produire des incohérences temporelles).

- Les réflexions latérales sont souvent bonne pour l'intelligibilité et la largeur de la source.
- la réflexion provenant du mur arrière est problématique dans la compréhension du message (du souvent au retard induit par cette onde)
- la réflexion provenant du sol peut s'avérer gênante si elle arrive très tôt (dans les 20 premières milliseconde) et créer de la distortion.

Au final [AJ01] à établi un gabarit utilisable pour les premières réflexions dans un petit volume (type studio d'enregistrement).

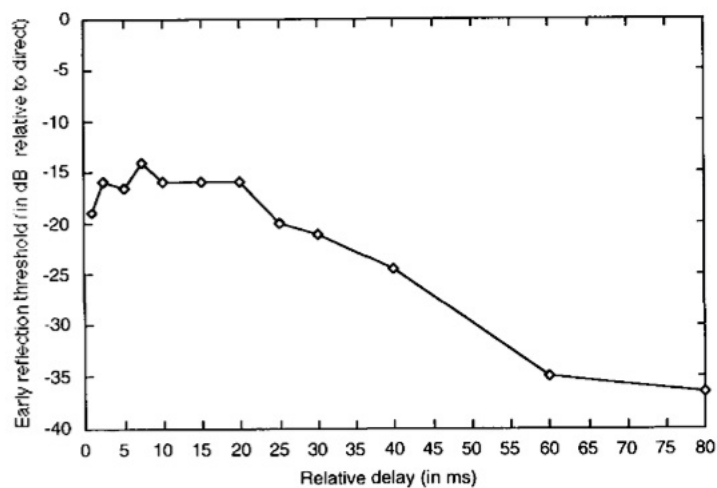


FIG. 3.1 – Gabarit des premières réflexions par rapport au son direct [AJ01]

3.2 L'effet de spatialisation

Les réflexions latérales sont à priori les plus importantes puisqu'elles contribuent activement à la localisation et à la sensation d'enveloppement. Pour avoir un complément d'information il faut se tourner sur les travaux de Barron (1971 et 1974) de Barron et Marshall (1981). La figure ci-dessous va montrer la gamme de niveau et de délai dans lequel Barron a trouvé l'effet d'espace en fonction du plaisir ressenti pour des signaux musicaux. Son expérience fut réalisée pour une seule réflexion latérale.

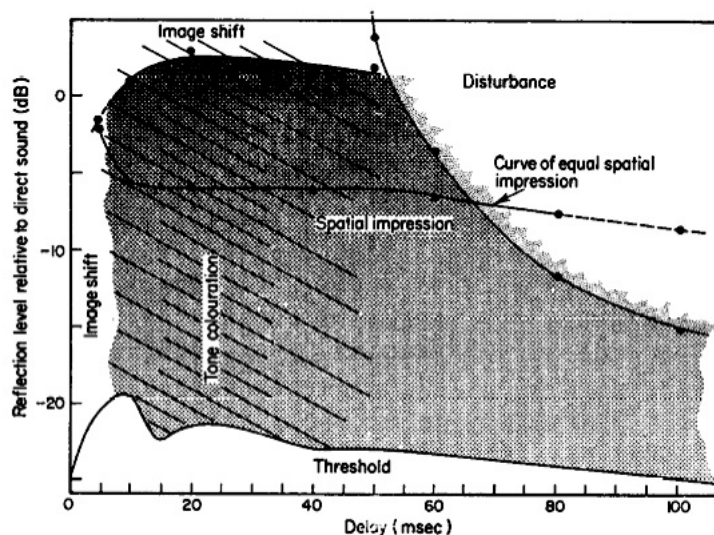


FIG. 3.2 – Zone acceptable pour un bon effet d'espace (d'après [Bar71])

Un effet d'espace est apparu immédiatement au seuil de masquage et devient plus fort lorsque la réflexion a augmenté. La sensation d'espace est limitée à fort niveau de réflexions à cause des mouvements d'image, la sommation de localisation apparaît. Cela peut être également lié au temps de délai si il est supérieur à 50ms (limite de l'effet Haas). La gamme limite pour un effet d'espace agréable est donnée par la gêne du à l'écho. Il a été également observé des changements dans le timbre pour des délais entre 10 et 50 ms. L'effet d'espace est indépendant du délai dans le cas d'une réflexion (pareil pour 2 réflexions latérales [BM81]). Les fréquences en dessous de 1.5kHz sont importantes pour l'effet d'espace.

Délais (ms)	Son direct	Reflexion latérale
0	Intensité	Taille mouvement de l'image sonore
5	Timbre	
10 - 20		Timbre
30-60		Effet d'espace
≥ 80	Echo	Echo

TAB. 3.1 – Détermination des seuils et effets d'une seule réflexion en présence du son direct [BM81]

Il a donc été montré que l'effet d'espace n'était pas dépendant du temps à laquelle la réflexion arrive mais par contre elle est dépendante de l'angle d'incidence (voir paragraphe 1.4.4 : Largeur apparente de source).

3.3 Bob Walker [Wal07]

Le temps de réverbération intéressant dans les studios d'enregistrements est de l'ordre de 0.3 - 0.4s. Les réflexions situées dans les 20 premières milliseconde après le son direct cause des distortions sur

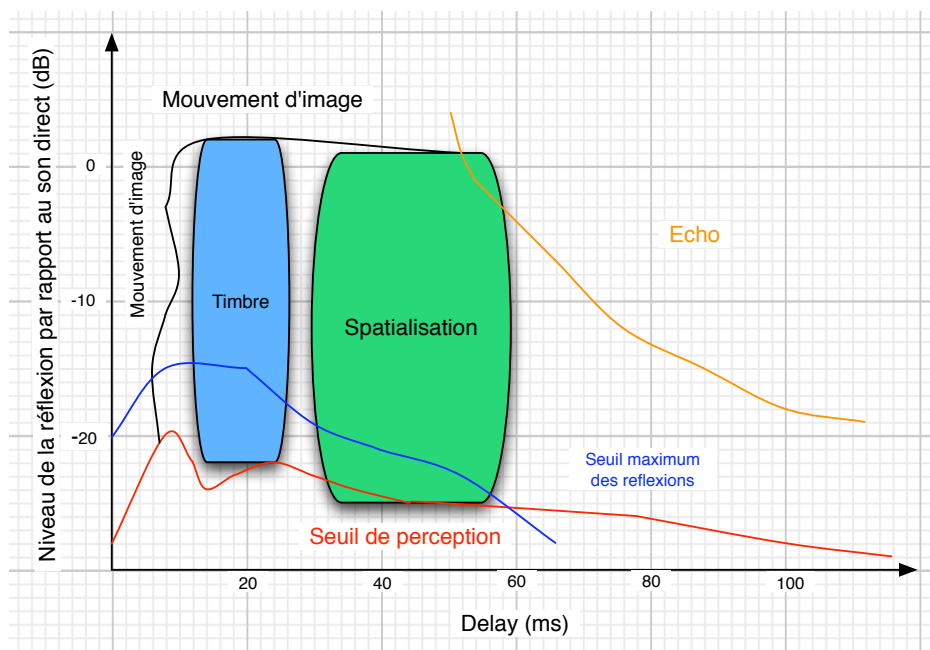


FIG. 3.3 – Représentation des phénomènes perceptifs pour les premières réflexions

le signal. Il faut les atténuer un maximum et Walker parle du critère nommé 20ms/-20dB qui dit que les réflexions comprise dans les 20 premières millisecondes suivant le son direct ne doivent pas dépasser -20dB par rapport au son direct. Une des réflexions qu'il est en général difficile d'atténuer dans les 20 premières millisecondes est celle provenant du mur arrière (cela dépend de la taille de la salle). Du à la distance, la réflexion parvient à l'auditeur avec une atténuation d'environ 10dB, ce qui facilite grandement la correction. Les premières réflexions arrivant à l'auditeur sont en général celles de la console et des équipements studio (≈ 2 ms), le plafond ($\approx 3/4$ ms), les murs latéraux ($\approx 4/6$ ms) et le mur arrière (≈ 12 ms). Les longueurs d'ondes courantes en audio sont de l'ordre de 10mm à 10m (34Hz à 22kHz). Au vu des longueurs d'ondes, les dimensions des studios d'enregistrement en général ne possèdent pas des grandes dimensions, donc il y a souvent des problèmes de modes dans les salles considérées. Les fréquences medium et aigues participent activement à la formation de l'image stéréo et aucunes informations en dessous de 300Hz contribue à cete dernière. A cause de l'ensemble des phénomènes de diffraction induit par les objets dans la pièce, le contrôle de la taille et du point d'écoute autour de l'auditeur est défini relativement arbitrairement (pour palier à ce problème et donner des solutions à des soucis d'esthétique et de propreté, Walker propose de placer les objets dans des armoires fermé par une vitre transparente). Le but étant d'élargir le cercle d'écoute autant que faire ce peut, mais pour une salle donnée, il est difficile de réaliser des cercles d'écoute optimum très large, surtout en élévation car souvent la hauteur sous plafond est plus petite par rapport à la largeur ou la longueur. De ce fait, il a estimé que dans des studios de tailles raisonnables, on pouvait espérer atteindre un sweet spot de 3m de diametre en horizontal pour 2.5m de diamètre en élévation.

Dans le cadre de système de multidiffusion, la multiplication des réflexions rend difficile leur contrôle. Les spécifications données ci-dessus sur la stéréo sont trop strictes pour être appliqué à un système de multidiffusion. Il a été calculer que pour un système de diffusion à 5 haut-parleurs placé dans une salle parallélépipédique, le nombre de réflexions d'ordre 1 est multiplié par 30 et les réflexions d'ordre 2 multipliées par 150. Il faut donc trouver des simplifications pour prendre en compte correctement le problèmes des réflexions précoces. Tout d'abord on peut considérer que les réflexions d'ordre 2 peut être exclu du modèle puisqu'elle vont faire parti du champ diffu et qu'elles auront atteintes une atténuation assez forte du à la distance, l'aborsption des matériaux et l'absorption de l'air. Au niveau des réflexions d'ordre 1, il semble plus adéquate de se fixer comme objectif qu'elles ne dépasse pas -10dB par rapport au son direct. Si l'on considère un distance approximative de 2m entre les HP et l'auditeur, alors toute réflexion ayant un

chemin de moins de 6.34m sera déjà en deçà des -10dB.

3.4 Soren Bech [Bec94] [Bec95]

Soren Bech a voulu quantifier l'influence des premières réflexions sur le timbre. Pour cela, il a mis en oeuvre un banc de mesure utilisant 8 sujets dans une chambre de petit volume en simulant un seul haut parleur (indépendant de la directivité). Un dispositif electroacoustique permet la simulation des 17 premières réflexions ainsi que le champ diffu. Le questionnaire est sous la forme "oui ou non" pour mesurer le seuil de détection et les JND pour évaluer l'augmentation du niveau des réflexions qui ont été mesurées. Son étude démontre que seules les réflexions d'ordre 1 du sol et du plafond contribue activement au timbre dans le cadre du bruit ou de la voix.

Dans son introduction, il commence par introduire une vérité toujours valable jusqu'à ce jour : l'interaction entre le haut parleur, la salle est l'auditeur n'est toujours pas correctement comprise et maîtrisée.

Des expériences ont montré que la décomposition de la décroissance d'un son dans une salle se scindait en 3 parties :

- Son direct
- avant 21-22ms ce sont les premières réflexions
- après les 21-22 ms il s'agit du champ diffu

Ceci correspond à ce qui a été trouvé dans [Jou03] [Too06] [Wal07] [Bar99]. Le nombre de réflexions choisies dépendent des caractéristiques de la salle et du haut-parleurs. C'est pour cela que l'article de Bech étudies les 17 premières réflexions, au delà, on tombe dans le champs diffu.

Au niveau des stimuli, mesuré au point d'écoute, le bruit faisait 66dB de pression et 50 pour la voix. Le panel n'était pas novices puisqu'ils avait participé à une campagne de mesure de 5000 essais sur le bruit et 1200 essais sur la voix avant ces expérimentations. A priori son analyse a été réalisé sur la méthode du seuil différentielle (méthode permettant d'obtenir le seuil de détection ou les JND à partir d'une comparaison + fort / - fort). Le niveau naturel correspond au niveau censé arrivé à l'oreille de l'auditeur après réflexion (calculé sur un modèle de source image).

Les résultats montrent qu'en général, pour le bruit rose, que le seuil de détection est plus haut que le niveau naturel (sauf pour les réflexions 1, 3, 8 et 12, il ne sait pas pourquoi ces réflexions sont en dessous du niveau naturel). Pour ce qui est de la voix, il y a une augmentation sensible du seuil de détection par rapport au bruit rose et donc au niveau naturel (l'augmentation dans le seuil de détection serait du à une incertitude sur le caractère non stationnaire de ce type de signal).

Le but des JND (Just Noticeable Differences) sont mesurés sur les niveaux sur chaque réflexion individuellement. Les réflexions ne sont pas audible individuellement si elles sont au niveau naturel. Si le JND est négatif, il correspond à un abaissement du niveau individuel de chaque réflexion. Si le JND est positif, il correspond alors à une augmentation du niveau qui est la différence entre la mesure du seuil de détection et le niveau naturel théorique (calculé sur un modèle de source image). Dans les tests, la plupart des mesures de JND sont positives du fait que les seuils de détection sont plus haut que les niveaux naturels. Les différences entre voix et bruit ne sont pas très différentes l'une de l'autre hormis pour les réflexions 13 et 17 (ce qui fût sensiblement la même chose pour les seuils de détection). Se serait des effets de masquage qui serait à l'origine des JND élevés. Au final, si on augmente l'intensité d'une réflexion sur le sol ou les murs latéraux, elle s'entendra, pour les autres réflexions : non.

3.5 Les studios, lieux d'écoute, petits volumes

Dans les studios qui correspondent à des petits volumes et donc des soucis de champs diffu, quelques pratique cité ci-dessous concordent souvent. Certaines valeur clés ressortent :

- Tr entre 0.3 et 0.4 s
- Pas de réflexions dans les 20ms suivant le son direct (pas plus de -20dB par rapport au son direct)

- Les autres réflexions selon les cas ne doivent pas excéder -15dB
- dans le cadre d'un système multidiffusion, les conditions ramenée à -10dB
- taille du sweet spot : $\phi 3\text{m}$ en horizontal et $\phi 2.5\text{m}$ en vertical
- les réflexions du sol et du plafond influencent le timbre individuellement (ainsi que les latérales)
- Entre 10 et 20ms c'est la réflexion latérale qui influence le timbre
- Entre 30 et 60ms cette réflexion va influencer sur l'impression d'espace

L'idée est d'atténuer au maximum les réflexions situées dans les 20ms suivant le son direct. Pour un système stéréo atténué les réflexions suivant les 20ms après le son direct, Walker a émis un principe empirique : ne pas dépasser les -20dB par rapport à l'amplitude du son direct. En multicanal, le fait d'avoir un réseau de transducteur modifie la loi et ramène les premières réflexions à un maximum de -10dB. En pratique il s'avère que cette technique donne des résultats satisfaisants dans les messages, mais on tendance à provoquer un souci dans la stabilité de l'image lorsque l'émetteur impose un mouvement. Petit plus sur les tables de mixage. Elles ont tendances à générer de la diffraction entre 1000 et 5000Hz et donne des réflexions précoces vers entre 0.8 et 1.2ms. Toutefois un des avantages des consoles qui sont souvent posées sur une table est le fait qu'elles masquent ou du moins atténuent la réflexion venant du mur d'en face qui a tendance à perturber l'image sonore.

Les réflexions précoces sont à regarder au dessus de la fréquence de Schroeder, en dessous on est sur du modal.

3.6 Conclusion

Ce chapitre nous permet de mettre en évidence un ensemble de "bonnes pratiques" concernant la gestion des premières réflexions en acoustique des salles. L'ensemble des informations situées dans ce chapitre concorde bien avec l'approche perceptive et permet d'envisager des simplifications dans les algorithmes pour ne travailler que sur les parties "utiles" du champ acoustique et réaliser des traitements individualisés en fonction de la fréquence, du temps et de la disposition spatiale. La majorité des préconisations sont quand même effectuées dans le cadre d'une diffusion mono ou stéréo. L'utilisation de diffusion multicanale type 5.1 et consorts complexifie énormément le champ acoustique et il devient difficile de prévoir le comportement des réflexions au point d'écoute et encore plus dans un espace plus large. De ce fait les préconisations sont revues à la baisse et les conditions donc moins strictes. Il serait intéressant de comparer ces données aux normes broadcast afin de faire des correspondances et éventuellement les compléter.

Bibliographie

- [AJ01] J Angus and A.S James. The effect of specular versus diffuse reflections on the frequency response at the listener. *JAES*, 49(3) :125–133, March 2001.
- [Bar71] M Barron. The subjective effects of first reflections in concert halls - the need for lateral reflections. *Journal of sound and vibration*, 15(4) :475–494, April 1971.
- [Bar99] Federico Cruz Barney. *Evaluation des performances d'un environnement informatique d'acoustique prévisionnelle*. PhD thesis, Université du Maine, December 1999.
- [Bec94] Soren Bech. Perception of reproduced sound : Audibility of individual reflections in a complete sound field. *AES*, (preprint 3849), February 1994.
- [Bec95] Soren Bech. Perception of reproduced sound : Audibility of individual reflections in a complete sound field 2. *AES*, (preprint 4093), October 1995.
- [Ber62] L Beranek. *Music, Acoustics and architecture*. John Wiley, 1962.
- [Bla01] Jens Blauert. *Spatial hearing, The psychophysics of human sound localization*. The MIT press, 2001.
- [BM81] M Barron and A.H Marshall. Spatial impression due to early lateral reflections in concert halls : the derivation of a physical measure. *Journal of sound and vibration*, 77(2) :211–232, July 1981.
- [BMB01] J M Buchholz, J Mourjopoulos, and J Blauert. Room masking : Understanding and modelling the masking of the room reflections. *J Audio Eng. Soc.*, May 2001.
- [Bru98] Michel Bruneau. *Manuel d'acoustique fondamentale*. Hermes, 1998.
- [Cre48] L. Cremer. Die wissenschaftlichen grundlagen der raumakustik. *S. Hirzel Verlag*, 1, 1948.
- [Haa72] Helmut Haas. The influence of a single echo on the audibility of speech. *JAES*, 20(2) :146–159, March 1972.
- [Jou03] Jacques Jouhaneau. *Acoustique des salles et sonorisation*. Tec et doc, 2003.
- [LCYG99] Ruth Y. Litovsky, H. Steven Colburn, William A. Yost, and Sandra J. Guzman. The precedence effect. *Journal Acoustical Society of America*, 106(4) :1633–1654, October 1999.
- [MISM94] M Morimoto, K Iida, K Sagami, and A.H Marshall. Physical measures for auditory source width (asw) : Part 1. discussion of the competing measures iacc and lf. in proceedings of the wallace clement sabine centennial symposium. *Acoustical Society of America*, pages 109–112, June 1994.
- [OT89] Sean E. Olive and Floyd E Toole. The detection of reflections in typical rooms. *J Audio Eng. Soc.*, 37(7/8) :539–553, July-August 1989.
- [Pol88] Jean Dominique Polack. *La transmission de l'énergie sonore dans les salles*. PhD thesis, Université du Maine, Le Mans, 1988.
- [Ser61] Seraphim. Über die wahrnehmbarkeit mehrerer rückwürfe von sprachschall. *Acustica*, 11 :80–91, 1961.

-
- [Sla93] Malcolm Slaney. An efficient implementation of the patterson-hildsworth auditory filter bank. *Apple computer technical report*, 35, 1993.
- [Too06] Floyd E Toole. Loudspeakers and rooms for sound reproduction. *JASA*, 54, June 2006.
- [Wal07] Bob Walker. Room acoustics for multichannel listening : Early reflection control. *AES*, AES 22ND, 2007.
- [ZZ91] E Zwicker and T Zwicker. Audio engineering and psychoacoustics. matching signals to the final receiver, the human auditory system. *J Audio Eng. Soc.*, 39(3) :115–126, March 1991.