



Historique et description des principaux indicateurs de sonie

Toolbox Sonie pour Matlab

GENESIS S.A.
www.genesis.fr

Sommaire

Table des matières

1	Objet du document.....	2
2	Introduction.....	3
2.1	Le système auditif	3
2.2	Métrologie	4
3	Modèles de sonie.....	7
3.1	Modèles de sonie pour les sons stationnaires	7
3.2	Modèles de sonie pour les sons non stationnaires	18
3.3	Modèle de sonie pour les sons impulsionnels	22
4	Indicateurs de sonie pour les sons non stationnaires	24
4.1	N4, N5 ou N7 selon Zwicker et Fastl (1999)	24
4.2	STLmax et LTLmax selon Glasberg et Moore (2002)	24
4.3	LMIS selon Boullet et coll. (2006)	24
	ANNEXE 1 : Références bibliographiques.....	I



1 Objet du document

Ce document détaille l'ensemble des modèles de sonie et des indicateurs psychoacoustiques qui permettent de calculer la sonie globale. Le cas des sons variant dans le temps et impulsionnels est aussi pris en compte. La Toolbox Sonie pour Matlab permet ainsi de connaître la sonie de n'importe quel type de son à partir de modèles normalisés ou publiés dans la littérature scientifique.



2 Introduction

Cette section est consacrée à un historique des recherches effectuées ces dernières décennies dans le domaine de la mesure et du calcul du niveau sonore perçu (sonie).

2.1 Le système auditif

La sonie est la grandeur subjective qui correspond au niveau sonore perçu. Elle dépend du niveau sonore mais aussi de la fréquence et de la durée des sons. Il y a donc un lien intime avec le fonctionnement de l'oreille humaine.

La Figure 1 est une représentation schématisée du système auditif périphérique humain.

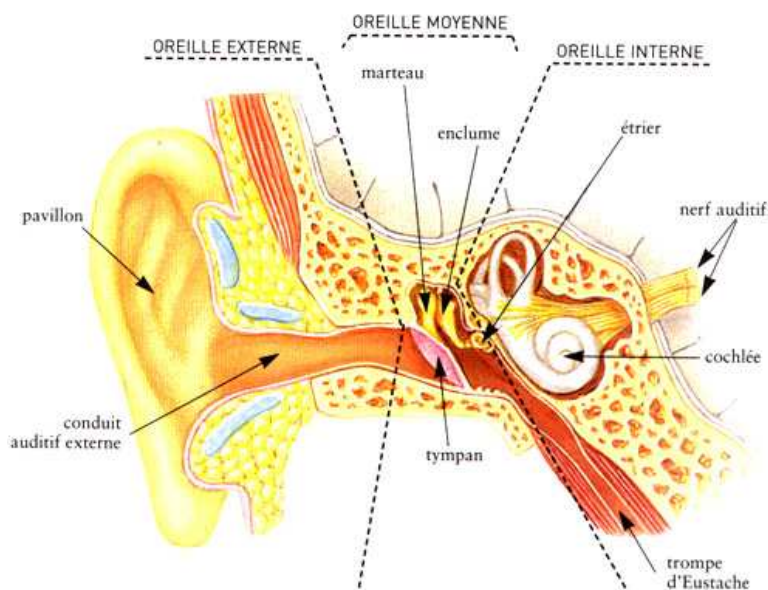


Figure 1: Schéma représentant le système auditif périphérique

Examinons le cheminement d'un son qui parvient à l'oreille humaine :

Un son parvient tout d'abord à l'oreille externe. Ainsi, le pavillon et le conduit auditif externe ont une influence sur l'amplitude de la vibration sonore (gain acoustique) qui arrive au tympan. Ces éléments de l'oreille externe agissent sur le son original à la manière d'un filtre, et ce son est transformé en fonction des caractéristiques de l'oreille : des fréquences sont plus ou moins amplifiées dans son spectre.

Arrivé à l'oreille moyenne (tympan, osselets, fenêtre ovale), le signal acoustique y est transformé en vibrations mécaniques. L'oreille moyenne joue le rôle d'adaptateur d'impédance, et permet ainsi d'éviter de trop grandes pertes d'énergie au passage vers l'oreille interne (cochlée).



Dans l'oreille interne, les vibrations sont ensuite transmises à l'intérieur de la cochlée jusqu'à la membrane basilaire. Cette membrane joue en quelque sorte le rôle d'un analyseur spectral. C'est ici que se passent des phénomènes responsables du masquage fréquentiel. La vibration mécanique est ensuite transformée en impulsions nerveuses (transduction électrique), à l'intérieur de l'organe de Corti, par les cellules ciliées internes et externes. Enfin, cet influx nerveux est transmis par le nerf auditif au cerveau et interprété par ce dernier.

2.2 Métrologie

La sonie est une grandeur subjective qui s'exprime en sones. Le sone est une unité basée sur une échelle sensorielle. Cette échelle a été établie à partir de méthodes de mesure dites directes (Stevens, 1956), consistant à demander à des auditeurs lors d'une expérience de donner des nombres proportionnels à la sonie des sons qui leurs sont présentés à différentes intensités.

La relation entre l'intensité acoustique et la sonie est appelée fonction de sonie (voir Figure 2). Elle est une simple fonction puissance pour des niveaux supraliminaires (**S (sones) = $k P^{0,6} = I I^{0,3}$**).

Par convention, la valeur de 1 sone est attribuée à la sonie d'un son pur de fréquence 1000 Hz qui a un niveau de 40 dB SPL. Un son de 2 sones est perçu deux fois plus fort qu'un son de 1 sone.

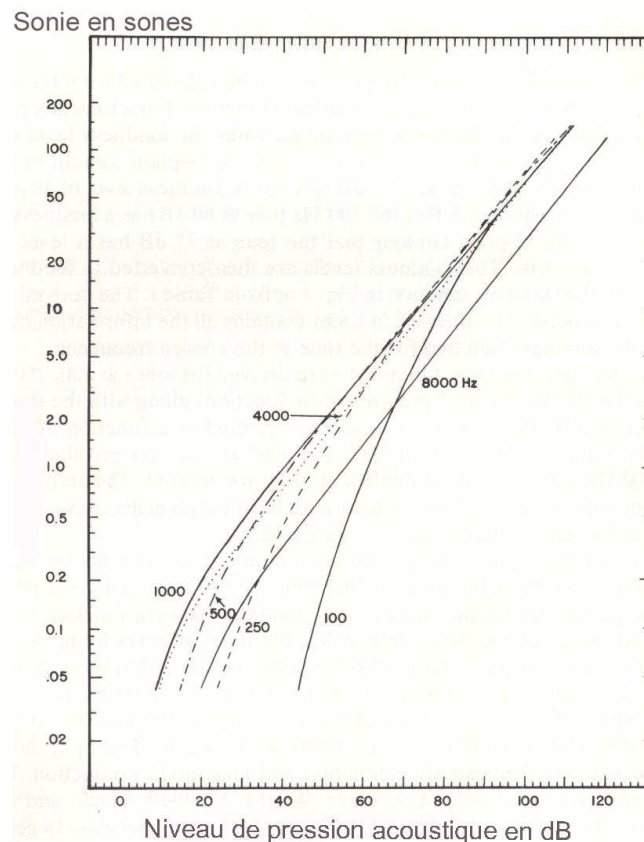


Figure 2 : Fonctions de sonie de sons purs pour plusieurs fréquences : 100, 250, 500, 1000, 4000 et 8000 Hz (D'après Scharf, 1978).



Une autre méthode de mesure de la sonie consiste à demander à des sujets de déterminer la sonie d'un son « test » par comparaison à un son pur « de référence » à une fréquence de 1000 Hz et un niveau de référence.

On obtient alors le niveau d'isophonie en phones, qui est le niveau en dB SPL que doit avoir le signal de référence pour que les deux sons (le son test et le son de référence) aient la même sonie. L'échelle des phones, établie à partir de méthodes dites de comparaison (Fletcher et Munson, 1933), coïncide avec l'échelle des dB SPL pour un son pur de 1000 Hz. Les lignes isophoniques ainsi obtenues (voir Figure 3) mettent en évidence les variations de la sensibilité de l'oreille humaine en fonction de la fréquence.

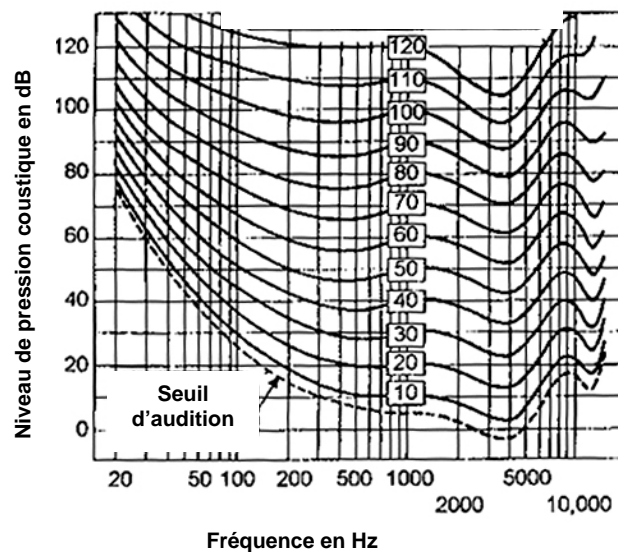


Figure 3 : Lignes isophoniques (Robinson et Dadson, 1956). L'échelle des phones coïncide avec l'échelle des dB SPL pour un son pur à une fréquence de 1000 Hz.

La sonie dépend non seulement du niveau sonore et de la fréquence mais aussi de la durée du signal. En effet, pour des sons d'amplitude constante, la sonie augmente régulièrement lorsque la durée du son croît de quelques millisecondes jusqu'à une certaine valeur. Cette valeur, appelée durée critique, varie de 50 à 400 ms selon les auteurs. Ce phénomène est appelé intégration temporelle. Il apparaît qu'au-delà de cette durée critique, et jusqu'à des durées de plusieurs secondes, la sonie reste constante.

Pour évaluer la sonie, des grandeurs physiques ont été développées. La plus connue et la plus utilisée est le dB(A), qui prend en compte les variations de la sensibilité de l'oreille en fonction de la fréquence (voir Figure 3). Le dB(A) est un calcul de niveau acoustique pondéré par les sensations d'égale sonie (isophonie) pour les sons de niveaux faibles (la valeur retenue est de 40 phones). En d'autres termes, le dB(A) vise à prendre en compte le comportement auditif en fonction de la fréquence, et est donc théoriquement valide pour des sons autour de 40 phones.

Il existe aussi les dB(B) et dB(C) pour la pondération respective des sons de niveaux moyens (70 phones) et forts (100 phones). Le dB(A) est cependant le plus



généralement employé en acoustique industrielle et environnementale. Il faut noter pourtant que les modèles de pondérations A, B ou C ne tiennent pas compte des phénomènes physiologiques comme le masquage fréquentiel ou le fonctionnement de l'oreille comme un banc de filtres, et restent donc d'une manière générale insuffisants pour estimer correctement la sonie.

Il est important de souligner aussi que la mesure du niveau en dB(A) a par définition tendance à sous-estimer le niveau perçu des sons complexes de niveau élevé. En effet, le dB(A) étant basé sur la ligne isosonique de sons stationnaires à 40 phones, il pondère fortement les basses fréquences (voir la Figure 3).

A l'heure actuelle, deux méthodes - celles de Zwicker et de Moore - permettent de calculer la sonie pour des sons stationnaires. Le modèle de Zwicker a d'abord été décrit dans une publication de 1958. En 1975, une méthode graphique permettant de calculer le niveau d'isonie a fait l'objet d'une norme internationale (ISO532B), et un programme en BASIC correspondant à l'implémentation de cette norme a été publié en 1984. Plus récemment, la méthode de Zwicker a fait l'objet d'une norme allemande DIN45631 (1991) qui contient un programme BASIC basé sur la méthode graphique de la norme ISO532 Méthode B. L'autre modèle est celui de Moore dont le principe a été publié en 1996. Le modèle a été révisé en 1997 pour introduire le calcul de la sonie masquée. Il a aussi fait l'objet en 2007 d'une norme américaine (ANSI S3.4-2007). La norme ANSI S3.4-2007 a remplacé la norme ANSI S3.4 1980, basée sur la méthode graphique de la norme ISO 532 Méthode A issue de travaux de Stevens (1961). La norme internationale ISO 532 est en cours de révision et devrait aboutir en 2010 à une nouvelle version, conservant l'approche de Zwicker et incluant celle de Moore.

En ce qui concerne les sons non stationnaires (c.-à-d. dont les caractéristiques temporelles et fréquentielles évoluent dans le temps), deux modèles ont été développés. Le premier par Zwicker et Fastl (1999) et le deuxième par Glasberg et Moore (2002). Un projet de normalisation internationale est en cours d'élaboration. Il existe toutefois une norme allemande DIN45631/A1(2008) basée sur les travaux de Zwicker et Fastl (1999) qui permet d'obtenir la sonie en fonction du temps.

Le cas des sons dits impulsifs - dont l'allure temporelle peut être caractérisée par une attaque rapide ne dépassant pas 5 ms, une absence de palier au maximum d'amplitude et un temps de descente plus ou moins long - a été étudié par Boullet (2005). Le modèle développé permet d'estimer la sonie globale de ce type de son.

Tous ces modèles permettent de calculer, à partir d'un signal acoustique en Pascal, la sonie de sons émis en incidence frontale en champ libre. Il est important de noter que la modélisation du système auditif à travers les différents modèles de sonie permet d'avoir une estimation de la sonie. Le seul moyen de connaître véritablement la sonie d'un son reste la mise en place de tests psychoacoustiques.



3 Modèles de sonie

Dans cette section, le détail des principaux modèles permettant d'estimer la sonie est donné. Trois sections différentes abordent respectivement le cas des sons stationnaires, des sons variant dans le temps et des sons impulsionnels.

3.1 Modèles de sonie pour les sons stationnaires

Les modèles présentés dans cette section visent à calculer la sonie des sons dits stationnaires, c.-à-d. présentant des propriétés statistiques temporelles et spectrales stables au cours du temps. Deux modèles sont détaillés ici : celui de Zwicker et celui de Moore.

Le modèle de Zwicker (1991) est fondé sur le calcul de l'excitation de la membrane basilaire par bandes critiques. Pour cela, il prend en compte:

- un modèle de transmission du signal acoustique au travers de l'oreille externe et de l'oreille moyenne,
- le phénomène de masquage dit fréquentiel (calcul de sonie spécifique),
- l'intégration de la sonie spécifique sur les vingt-quatre bandes critiques.

Le principe du modèle de Moore et coll. (1997) est le même que celui de Zwicker (1991) à trois différences près qui sont :

- le mode de calcul des filtres,
- les corrections de champ,
- le mode de calcul de l'excitation.

La Figure 4 présente les principales étapes dans le calcul de la sonie pour les deux modèles. Leur principe général étant le même, les deux modèles sont décrits en parallèle dans ce qui suit.

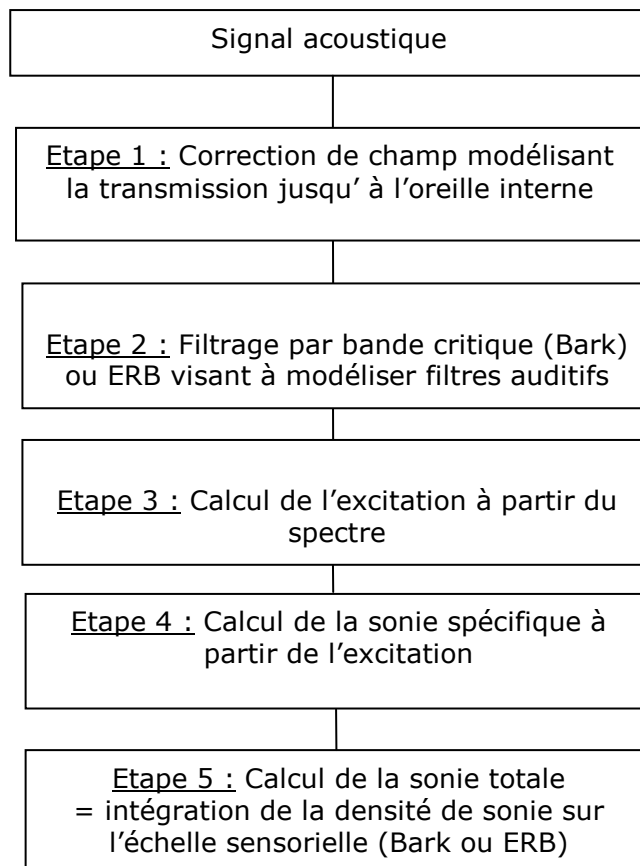


Figure 4: Etapes du calcul de la sonie pour les sons stationnaires

Etape 1 : Correction de champ (transmission du son au travers de l'oreille externe et moyenne)

Zwicker

Le facteur de transmission, appelé a_0 par l'auteur (voir Figure 5), prend en compte la transformation qui a lieu entre le champ libre et l'oreille interne. Il est appliqué au signal d'entrée pour modéliser cet effet (par soustraction de niveau au signal originel, avant de procéder aux calculs de sonie proprement dits).

Au-dessus de 1000 Hz, cette fonction de transfert a la forme de la courbe du seuil d'audition absolu, mais inversée. En dessous de 1000 Hz, Zwicker suppose que la transmission du champ libre à l'oreille interne est uniforme et de gain unité. Il suppose ainsi que l'augmentation du seuil d'audition pour les basses fréquences est uniquement due au bruit interne de l'oreille.

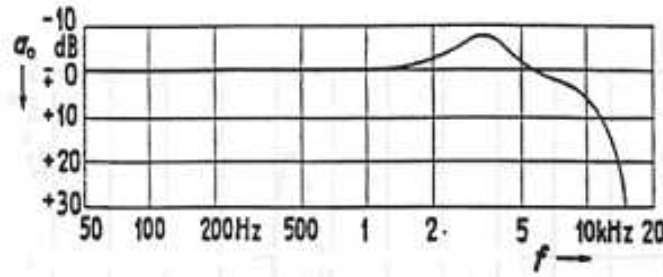


Figure 5: Variations avec la fréquence du facteur de transmission a_0 entre le champ libre et l'oreille interne (Zwicker et Feldtkeller, 1981, p. 111). En champ diffus, ce facteur est modifié et il faut introduire une autre correction.

Moore

Contrairement à Zwicker, Moore et coll. (1997) n'associent pas l'augmentation du seuil d'audition en basse fréquence au seul bruit interne. En effet, des données physiologiques (mentionnées dans Moore et coll., 1997), laissent penser que d'autres phénomènes, en plus du bruit interne de l'oreille, peuvent expliquer l'augmentation de ce seuil en basse fréquence. Ils utilisent donc dans leur modèle une correction dite ELC (Equal Level Contour). Cette correction est basée sur deux hypothèses :

- 1) Au-dessus de 1000 Hz, l'oreille interne a la même sensibilité à toutes les fréquences. Ainsi, la variation du seuil absolu avec la fréquence n'est due qu'à l'effet de filtrage de l'oreille externe et moyenne. Le facteur de transmission varie donc avec la fréquence de manière inverse au seuil absolu d'audition. La correction correspondante, appelée MAF (Minimum Audible Field), est donc la même que celle utilisée par Zwicker.
- 2) Au-dessous de 1000 Hz, on suppose que le bruit interne ne peut pas à lui seul rendre compte de la courbe de seuil. La fonction de transfert correspondante est alors modélisée par une fonction dont la forme est celle de la ligne isosonique à 100 phones mais inversée. Cette correction totale est appelée ELC (Figure 6).

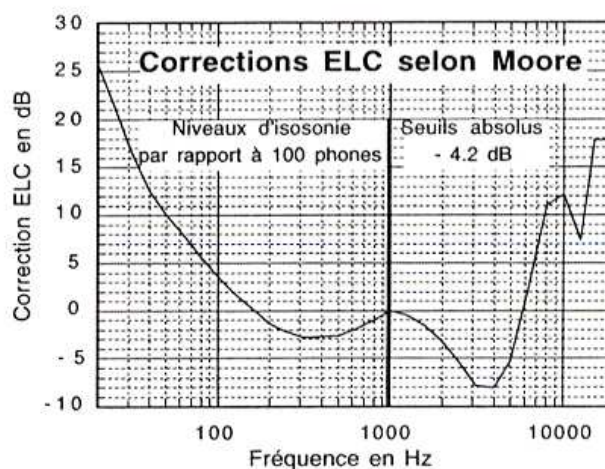


Figure 6: Variations avec la fréquence du facteur de transmission ELC entre le champ libre et l'oreille interne selon le modèle de Moore (1997)

Etape 2 : Modélisation et calcul des filtres auditifs

Zwicker

La sonie est liée à la répartition de l'excitation, provoquée par un stimulus, le long de la membrane basilaire (située dans la cochlée). Par conséquent, Zwicker représente le fonctionnement de l'oreille, ou plus précisément de la membrane basilaire, comme celui d'un banc de filtres.

Ces filtres, dits « physiologiques », sont modélisés par des bandes appelées bandes critiques. Ces bandes sont des bandes spectrales adjacentes, dont l'unité est le Bark, à l'intérieur desquelles la sonie reste constante. Au-delà, la sonie augmente avec la largeur de bande. En dessous de 500 Hz, la largeur des bandes critiques est constante (100Hz). Au-dessus de 500 Hz, cette largeur augmente avec la fréquence (voir Tableau 1 et Figure 7 pour les valeurs).

Numéro de la bande	Fréquence centrale en Hz	Largeur de la bande critique en Hz	Fréquence de coupure supérieure en Hz
1	50	80	100
2	150	100	200
3	250	100	300
4	350	100	400
5	450	110	510
6	570	120	630
7	700	140	770
8	840	150	920
9	1000	160	1080
10	1170	190	1270
11	1370	210	1480
12	1600	240	1720
13	1850	280	2000
14	2150	320	2320
15	2500	380	2700
16	2900	450	3150
17	3400	550	3700
18	4000	700	4400
19	4800	900	5300
20	5800	1100	6400
21	7000	1300	7700
22	8500	1800	9500
23	10500	2500	12000
24	13500	3500	15500

Tableau 1: Tableau définissant les 24 bandes critiques selon Zwicker (Zwicker et Feldtkeller (1981), p. 71). Pour chaque bande sont recensées la fréquence centrale (Hz), la largeur de la bande critique (Hz) et la fréquence de coupure supérieure (Hz).

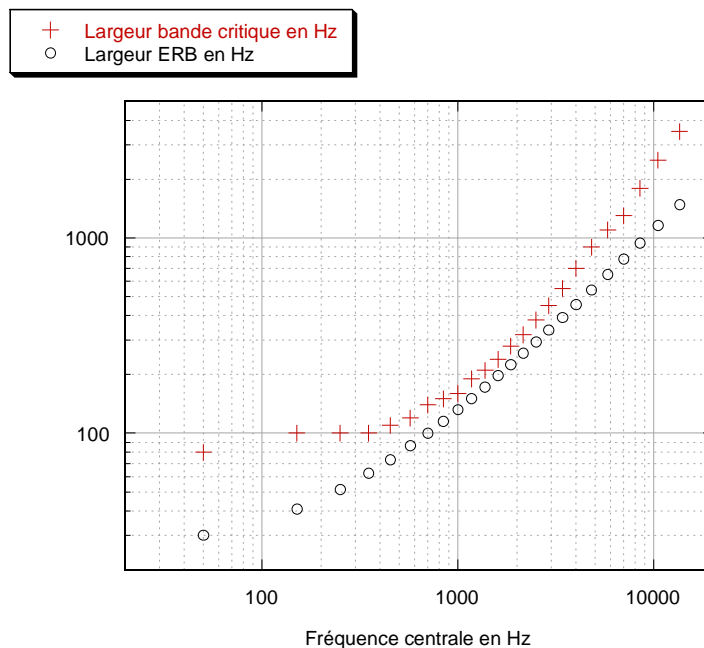


Figure 7: Largeur des bandes critiques (croix) et des ERB (cercles) en fonction de la fréquence centrale en Hz. Les valeurs des largeurs des bandes critiques sont tirées d'un tableau donné par Zwicker (1961).

La largeur des ERB en Hz est égale à $24,7(4,37f+1)$ où

Moore

Moore et coll. (1997) proposent une autre représentation de la capacité de filtrage du système auditif, qui reste proche des bandes critiques. La notion d'ERB (Equivalent Rectangular Bandwidth) remplace alors la notion de bande critique utilisée par Zwicker. Les largeurs de ces bandes sont indiquées sur la Figure 7.

Etape 3 : Calcul de l'excitation

Le diagramme d'excitation induit par un son est calculé à partir du niveau de ce son dans chaque bande de fréquence définie précédemment (bande critique ou ERB) après avoir tenu compte des facteurs de transmission à travers l'oreille externe et l'oreille moyenne (voir Etape 1). L'amplitude de sortie de chaque filtre en bandes est appelée diagramme d'excitation, et est calculé en fonction des fréquences centrales des filtres correspondants. Le calcul du diagramme d'excitation est une partie importante des modèles de sonie et diffère selon la méthode utilisée (Zwicker ou Moore).

La notion d'excitation est intimement liée au phénomène de masquage des sons. Le masquage d'un son par un autre se traduit par la baisse d'audibilité de ce son. Le masquage peut être partiel (le signal dit masqué est encore audible) ou total (le signal masqué est alors inaudible). Les expériences classiques mises en place pour étudier les phénomènes de masquage consistent à mesurer le seuil d'audition d'un signal en présence d'un autre signal dit masquant. A partir de ces expériences, des courbes de masquage (présentant une dissymétrie) ont été mises en évidence (voir Figure 8). Elles visent à modéliser comment un signal masque plus ou moins fortement un autre signal. Cela met notamment en exergue comment un signal donné masque plus fortement un autre signal de fréquence plus aigüe.



Zwicker

Selon Zwicker, le diagramme d'excitation reflète le diagramme de masquage d'un son pur masqué par un bruit à spectre étroit. Zwicker suppose en effet que les diagrammes d'excitation ont la même forme que les courbes de seuils de détection masquée. Les courbes de masquage traduisent la sélectivité de l'oreille et dépendent de la fréquence centrale et du niveau du signal d'entrée (Figure 8).

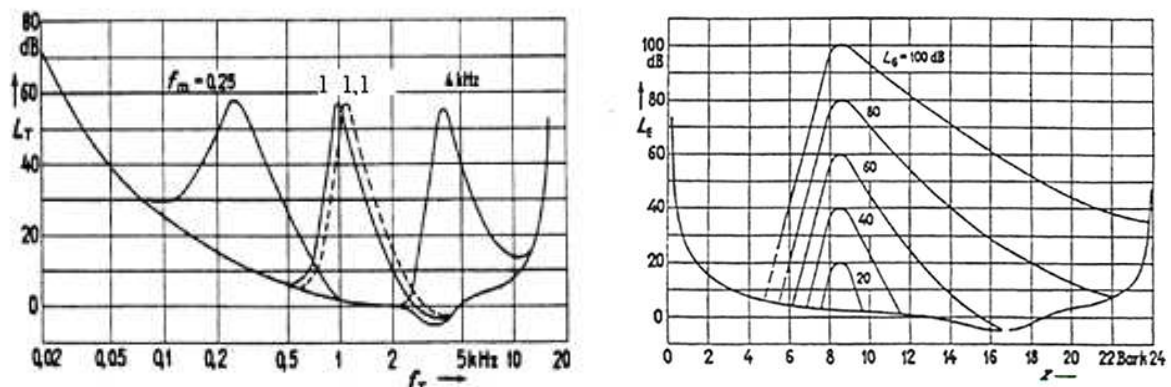


Figure 8: A gauche : courbes de masquage par des bandes de bruit à quatre fréquences centrales différentes (250 Hz, 1 kHz, 1.1 kHz et 4 kHz). Les courbes correspondent aux seuils de détection mesurés en présence de ces bandes de bruits (de largeur spectrale égale à une bande critique) émises à un niveau de pression de 60 dB SPL.

A droite : diagramme d'excitation produit par un bruit centré sur 1 kHz et de 160 Hz de largeur spectrale. Le niveau de pression de la bande de bruit est noté L_g et vaut successivement 20, 40, 60, 80 et 100 dB SPL (Zwicker et Feldtkeller, 1981).

Moore

Selon Moore, le diagramme d'excitation est calculé à partir de la sortie des filtres auditifs centrés sur les fréquences composant le son. Prenons par exemple un son pur (sinusoïde) de fréquence 1 kHz. Le diagramme supérieur de la Figure 9 montre les courbes de réponse des filtres centrés au voisinage de 1 kHz pour lesquels il y a une contribution à 1 kHz (c.-à-d. pour lesquels la courbe pleine croise la ligne verticale en pointillés). Le trait vertical en pointillés représente la fréquence du son pur d'excitation. Sur le diagramme inférieur, les abscisses correspondent aux fréquences centrales des filtres voisins (cf. diagramme supérieur), et les ordonnées représentent la réponse à 1 kHz de ces filtres. Le diagramme d'excitation est ainsi construit par la courbe qui joint les points, ramenés à chaque fréquence centrale des filtres voisins.

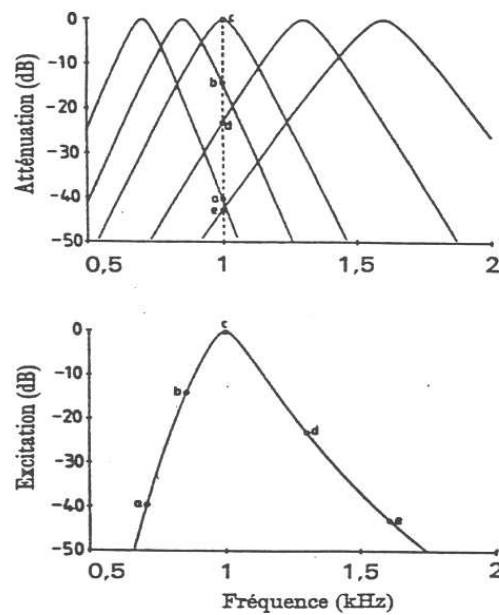


Figure 9: Principe de la détermination du diagramme d'excitation d'un son pur à 1 kHz selon le modèle de Moore.

Pour un son complexe, le diagramme d'excitation est obtenu en répétant l'opération décrite ci-dessus pour toutes les fréquences du signal. Ainsi, le diagramme d'excitation final est la somme des excitations partielles.

Etapes 4 et 5 : Calcul de la sonie spécifique et de la sonie totale

Le calcul de la sonie totale passe par la détermination de la quantité appelée densité de sonie (notée N'), aussi dénommée sonie spécifique, à partir du diagramme d'excitation (E) dans chaque bande de fréquence (bande critique ou ERB selon le modèle). La sonie spécifique correspond à un spectre sensoriel, c'est-à-dire le spectre d'un signal tel qu'il est perçu (en tenant compte de l'effet de l'oreille externe et moyenne, du fonctionnement de l'oreille comme un banc de filtres et du masquage fréquentiel).

Selon la loi de Stevens, la relation entre l'excitation et la densité de sonie est une fonction puissance (relation 1) :

$$N' = c \cdot E^\alpha \quad (1)$$

où c est une constante déterminée par ajustements successifs à partir de données expérimentales. Quant à l'exposant α , inférieur à 1, il traduit la relation non linéaire et compressive entre le niveau d'excitation et la densité de sonie.

Zwicker

Zwicker calcule la densité de sonie à partir du diagramme d'excitation. Pour chaque bande critique, un niveau d'excitation est calculé (voir Figure 8), appelé excitation de cœur. Pour cela, quelques ajustements sont faits : l'allure du flanc supérieur est globalement conservée (sous une certaine contrainte décrite plus loin) et dépend du niveau et de la bande critique concernée. L'allure du flanc inférieur est quant à elle transformée en une ligne verticale, en s'assurant de ce que l'aire sous les



courbes reste la même que celle avant ajustement. Le calcul de la densité de sonie est donc décomposé en un calcul de la sonie de cœur effectué à partir du niveau d'excitation et un calcul de la sonie de flanc (Figure 10).

La sonie de cœur est calculée à partir de l'équation ci-dessous, tirée de Zwicker et Fastl (1999):

$$N' = c. (E_{Thq})^{\alpha} . [(0,5 + 0,5 \cdot (E_{Stimulus})/(E_{Thq}))^{\alpha} - 1] \quad (2)$$

L'exposant α (égal à 0,23) est celui de la droite correspondant à la fonction de sonie d'un bruit uniformément excitant (courbe du haut en tirets-points de la Figure 11). $E_{Stimulus}$ est l'excitation produite par le stimulus. E_{Thq} , donnée par Zwicker, est l'excitation au seuil d'audition dans le silence, et correspond à l'excitation qui est produite par le bruit interne physiologique de l'oreille.

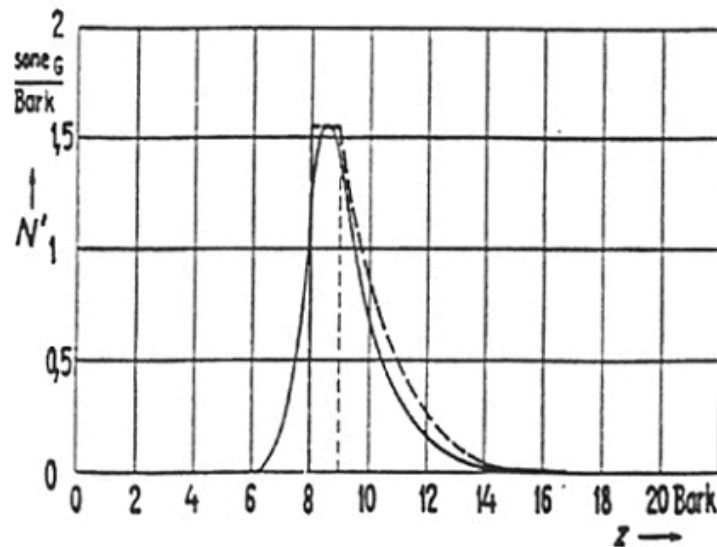


Figure 10: Courbe de densité de sonie, en fonction de la tonie en Bark, d'un son pur de 60 dB à 1 kHz. La courbe en traits-tirets représente l'approximation du modèle, décomposée en une partie rectangle (sonie de cœur) et une partie courbe décroissante (sonie de flanc) utilisée pour le calcul de la sonie globale (Zwicker et Feldtkeller, 1981).

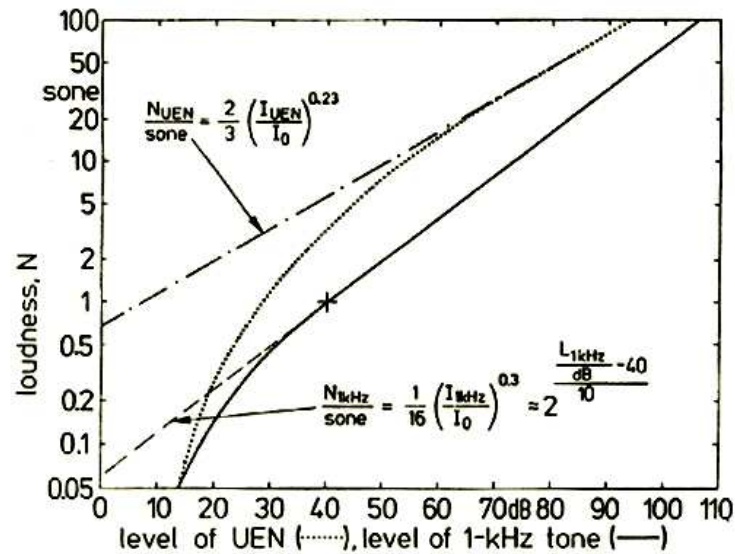


Figure 11: Fonctions de sonie d'un son pur à 1 kHz (courbe du bas, en trait plein) et d'un bruit uniformément excitant ou UEN (courbe du haut, en pointillés). Les droites approchent les fonctions de sonie, suivant une loi de puissance, d'un son pur à 1 kHz (en tirets) et d'un bruit uniformément excitant (tirets-points). La sonie est représentée en fonction du niveau de pression des signaux (Zwicker et Fastl, 1999).

Le calcul de la sonie de flanc par Zwicker est laborieux car il n'existe pas de relation mathématique explicite pour les courbes de seuils masqués. Les courbes des sonies de flanc des diagrammes patrons (Figure 12), utilisés autrefois, ont été approximées pour les programmes de calcul par des lignes brisées (Figure 13). Les segments de droite de la ligne brisée sont délimités par des lignes horizontales. Ces lignes horizontales correspondent à différents échelons de densité de sonie dépendant du niveau d'excitation. En effet, plus le niveau d'excitation est élevé, et donc la sonie de cœur aussi, plus la pente du premier segment de la ligne brisée est raide.

La densité de sonie ainsi obtenue est la succession dans chaque bande critique des sonies de cœur et/ou des sonies de flanc. Par principe, si la sonie de cœur d'une bande critique **n** se situe au-dessous de la sonie de flanc de la bande critique précédente (**n-1**), alors la sonie de cœur de la bande critique **n** est masquée. C'est alors la sonie de flanc de la bande critique **n-1** qui est prise en compte pour déterminer la densité de sonie dans la bande critique **n** (voir Figure 12).

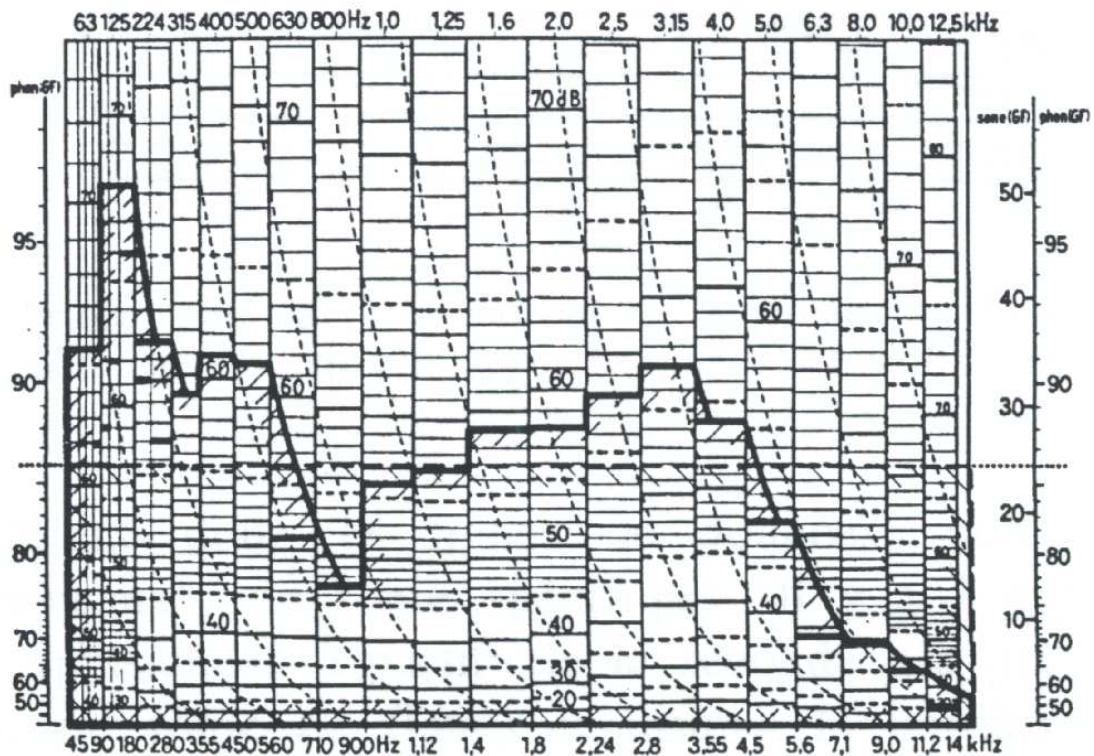


Figure 12: Exemple de distribution des niveaux par bandes critiques en fonction de la fréquence, pour un bruit de machine. La sonie de cœur calculée dans chaque bande critique est représentée par les traits épais horizontaux. La sonie de flanc, modélisant le masquage fréquentiel, est représentée sur la même figure par les courbes descendantes en trait épais déterminées d'après les diagrammes patrons (courbes en pointillés). La sonie globale est repérée par la ligne horizontale en traits-tirets. Le niveau d'isophonie correspondant est de 86 phones.

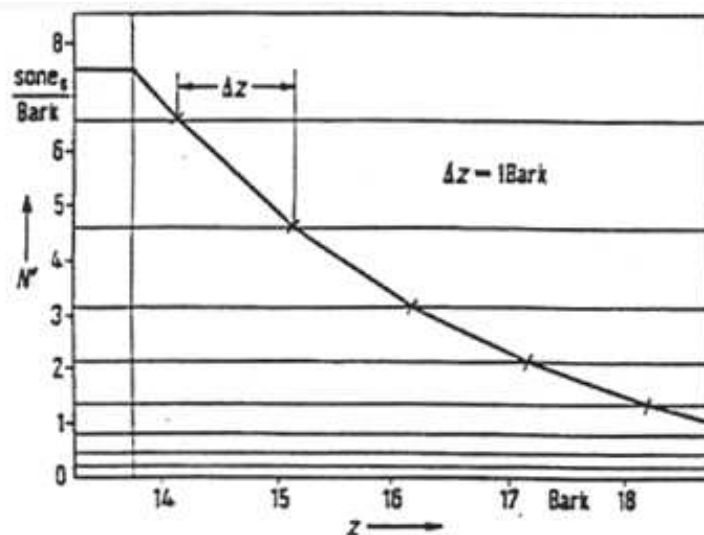


Figure 13: Exemple de ligne brisée utilisée pour approximer la sonie de flanc induite par une bande située à 2 kHz ayant une sonie de cœur de 7,5 sones/Bark.

Finalement, le calcul de la sonie totale est effectué par intégration de la sonie spécifique sur les 24 bandes critiques (procédure de sommation de sonie) : la sonie



totale correspond en fait à l'aire sous la courbe de sonie spécifique (par exemple représentée sous la courbe en trait plein épais de la Figure 12).

Moore

Le principe de calcul de la densité de sonie par Moore est légèrement différent de celui de Zwicker. En effet, Moore ne décompose pas la densité de sonie en sonie de cœur et sonie de flanc. La densité de sonie est calculée directement à partir du diagramme d'excitation (voir Figure 9).

Moore et coll. (1997) supposent aussi que le bruit de fond auditif est inaudible. Ils modélisent donc ce phénomène en soustrayant la densité de sonie produite par le bruit interne à la densité de sonie du stimulus :

$$N' = N'_{\text{Stimulus}} - N'_{\text{BruitInterne}} = c.(E_{\text{Stimulus}})^{\alpha} - c.(E_{\text{Thq}})^{\alpha}$$

$$N' = c. [(E_{\text{Stimulus}})^{\alpha} - (E_{\text{Thq}})^{\alpha}]$$

Dans le cas où l'excitation E_{Stimulus} produite par le stimulus serait plus faible que l'excitation au seuil, la densité de sonie, ne pouvant être négative, est supposée nulle.

Finalement, la sonie totale est calculée par intégration de la sonie spécifique selon les numéros des bandes d'ERB.



3.2 Modèles de sonie pour les sons non stationnaires

Le principe de base de ces modèles est le même que pour les sons stationnaires. Cependant, ils prennent en compte le masquage temporel¹, et la sonie est calculée en fonction du temps et non pas de manière globale.

La Figure 14 représente les principales étapes dans le calcul de la sonie pour les deux modèles.

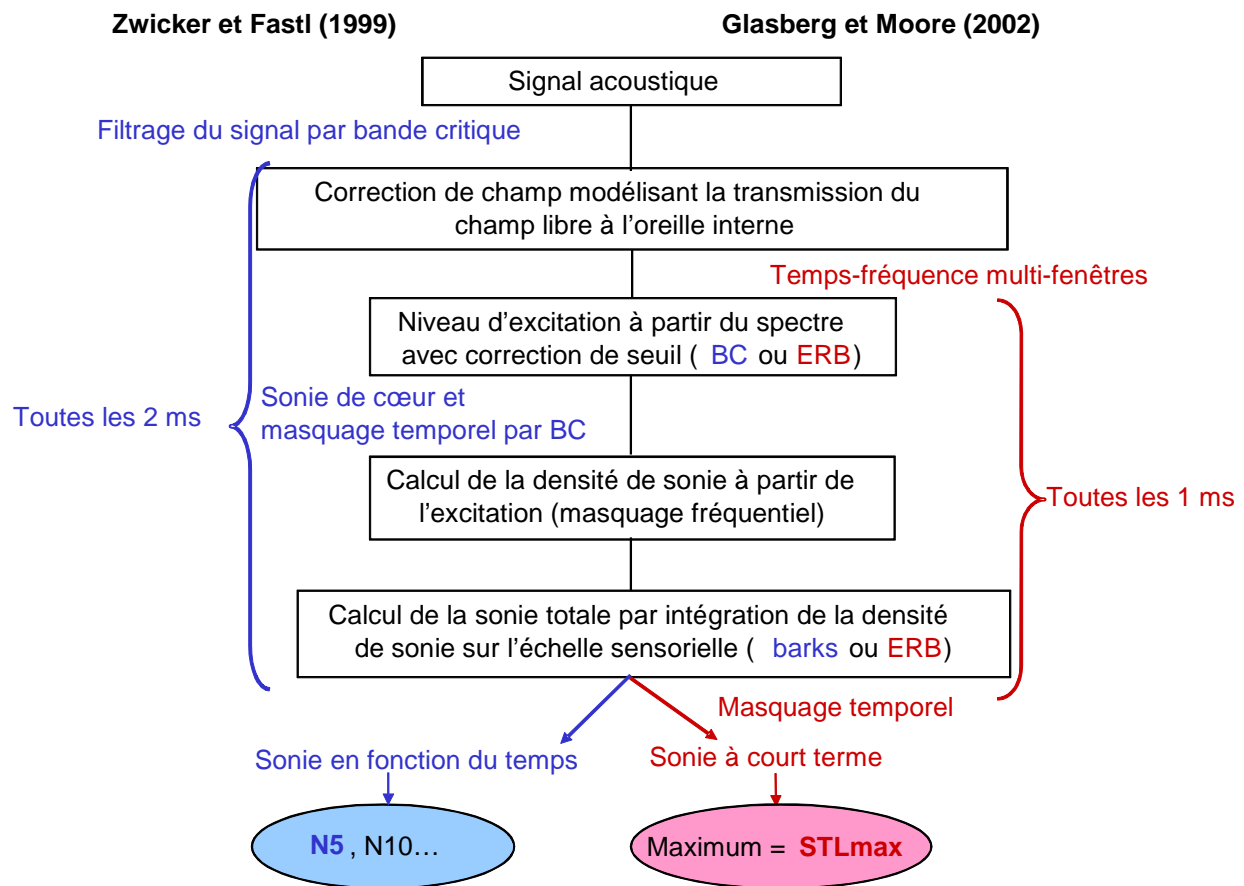


Figure 14: Etapes du calcul de la sonie pour les sons non stationnaires

Zwicker et Fastl (1999)

Zwicker et Fastl (1999) ont travaillé à l'élaboration d'un modèle de calcul de la sonie de sons non stationnaires. Pour cela, il est nécessaire de prendre en compte les phénomènes de masquage dits temporels. En effet, un signal peut être masqué (ou difficile à détecter) s'il est précédé de près dans le temps par un autre signal (masquage postérieur). Un signal peut aussi être masqué par un son qui le suit de

¹ Dans le cas d'un effet de masquage non simultané, l'antériorité d'un bruit masquant par rapport à un signal entraîne une élévation du seuil de détection de ce dernier. Il est généralement accepté que le masquage temporel peut avoir lieu sur une durée de 200 ms après l'extinction du bruit masquant.



près dans le temps (masquage antérieur, ou rétroactif). Toutefois, Zwicker ne tient pas compte dans son modèle de ce deuxième effet.

Zwicker utilise un schéma équivalent de quadripôle (Figure 15) pour modéliser le masquage temporel postérieur en fonction de l'intensité du signal et de sa durée. En d'autres termes, selon la longueur du signal, l'effet de masquage s'étend sur une durée plus ou moins longue après la fin du signal. Le masquage temporel est modélisé par une décharge de condensateurs.

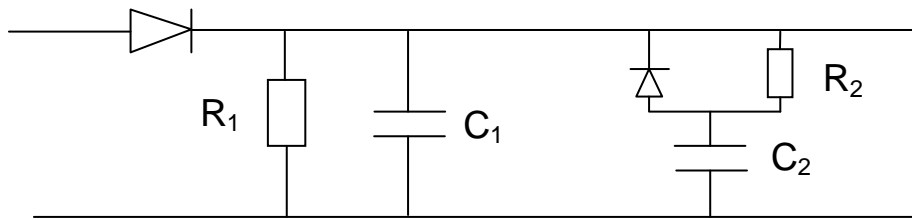


Figure 15: Circuit RC du quadripôle servant à modéliser le masquage temporel postérieur. $R_1=35\text{k}\Omega$; $C_1=0,7\text{ }\mu\text{F}$; $R_2=20\text{ k}\Omega$; $C_2=1\text{ }\mu\text{F}$ (Zwicker, 1984).

Les charges et décharges des condensateurs (C_1 et C_2) dépendent des tensions d'entrée et de sortie. Supposons que la tension du signal d'entrée soit un créneau. C_1 est chargé instantanément et C_2 se charge avec une constante de temps $T_2=R_2 \cdot C_2$ égale à 20 ms. C_2 est considéré comme chargé après une durée comprise entre $3 \cdot T_2=60\text{ ms}$ (condensateur chargé à 95%) et $5 \cdot T_2=100\text{ ms}$ (condensateur chargé à 99%).

Lorsque le signal est interrompu, les condensateurs se déchargent selon deux cas :

1. cas où le signal d'entrée dure moins de 100 ms :

Le condensateur C_2 n'a pas eu le temps de se charger complètement. C_1 se décharge alors dans R_1 et charge C_2 par R_2 .

2. cas où le signal d'entrée dure plus de 100 ms :

Les condensateurs C_1 et C_2 sont complètement chargés, le système est en équilibre. C_1 et C_2 se déchargent alors dans R_1 . La décharge, se faisant avec une constante de temps égale à $R_1 \cdot (C_1 + C_2)$, donc plus lentement que dans le cas numéro 1.

La Figure 16 présente un exemple de masquage temporel occasionné par des créneaux de sons purs de fréquence 5 kHz et de durée respective de 10 ms et de 100 ms. Le modèle de Zwicker décrit ci-dessus permet de calculer la sonie en fonction du temps (Figure 16 (c)), mais ne donne aucune indication sur la façon dont un auditeur jugerait la sonie globale d'un tel son variant dans le temps.

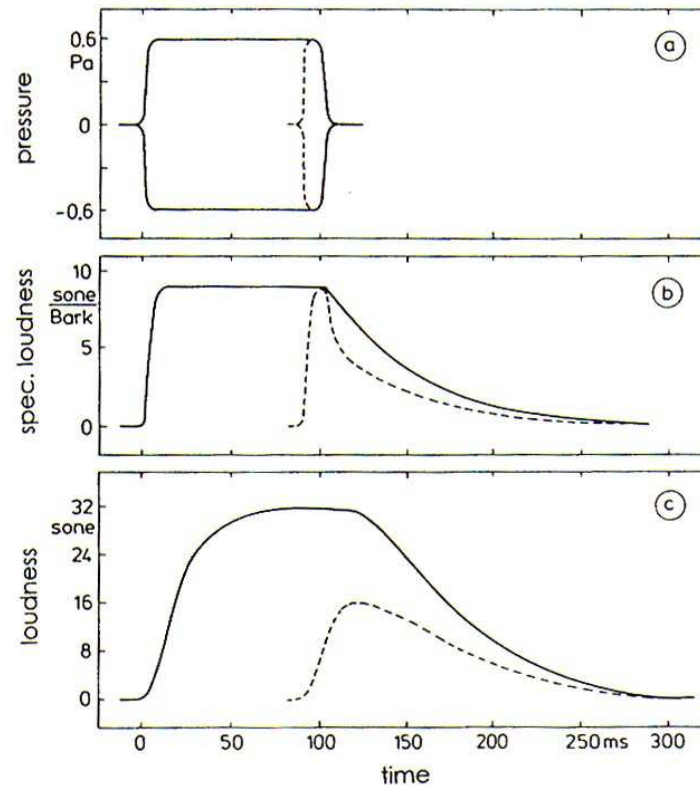


Figure 16: Etablissement de la sonie de créneaux à 5 kHz de 100 ms (trait plein) et 10 ms (pointillés). (a) Enveloppe temporelle des deux créneaux, (b) Densité de sonie correspondant aux deux créneaux dans la bande critique numéro 19, (c) Sonie en fonction du temps (Zwicker et Fastl 1999).

Glasberg et Moore (2002)

Glasberg et Moore (2002) ont aussi développé un modèle de sonie pour des sons variant dans le temps. De la même manière que les modèles précédents, ce modèle prend en compte la transmission à travers l'oreille externe et l'oreille moyenne, mais cette fois en utilisant un filtre à réponse impulsionnelle finie (FIR). La fonction de transfert du filtre (FIR) simulant la transmission du champ libre à l'oreille interne est donnée sur la Figure 17.

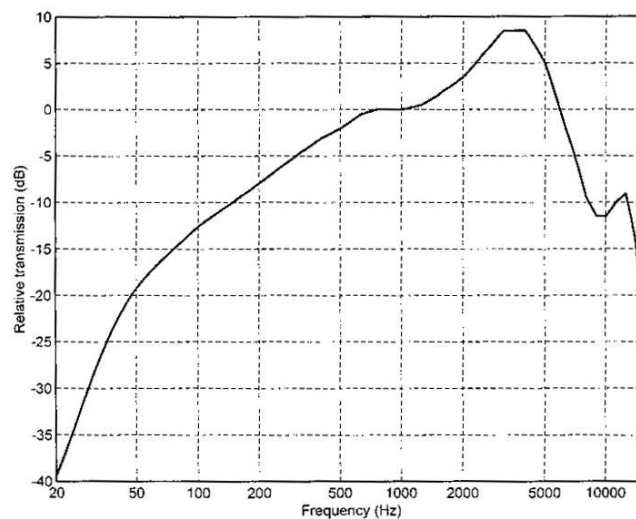


Figure 17: Fonction de transfert du filtre numérique à réponse impulsionnelle finie utilisée pour simuler les effets de l'oreille externe et de l'oreille moyenne. Le gain est calé pour être de 0 dB à 1000 Hz (Glasberg et Moore, 2002).

La cochlée est censée fonctionner comme un banc de filtres passe-bande dont les fréquences centrales sont étalées de 50 à 15000 Hz. La largeur des filtres augmente avec leurs fréquences centrales. Par exemple, pour un filtre centré sur 100 Hz, la largeur de bande rectangulaire équivalente (ERB) est de 35 Hz. Pour une fréquence centrale de 10 kHz, la largeur de bande rectangulaire est de 1100Hz.

La forme des filtres dépend du niveau. En effet, les pentes des filtres à basse fréquence deviennent moins raides quand le niveau augmente. La courbe donnant l'amplitude de sortie de chaque filtre, pour un niveau donné, est censée modéliser le diagramme d'excitation.

La méthode utilisée par Moore pour calculer le diagramme d'excitation est basée sur une analyse spectrale multi-résolution par transformées de Fourier. Afin d'obtenir une résolution fréquentielle à basse fréquence comparable à celle du système auditif, l'analyse se fait sur des segments relativement longs (64 ms). Pour les hautes fréquences, les auteurs utilisent des fenêtres d'analyse plus courtes (2 ms). Ainsi, six FFT (Fast Fourier Transform) sont effectuées en parallèle, pour des fenêtres temporelles de durées 64, 32, 16, 8, 4 et 2 ms. Ces calculs de FFT sont respectivement utilisés pour mesurer les niveaux dans les bandes de fréquences suivantes : 20-80 Hz, 80-500 Hz, 500-1250 Hz, 1250-2540 Hz, 2540-4050 Hz et 4050-15000 Hz. Le diagramme d'excitation étant calculé à partir du spectre toutes les 1 ms, il y a donc un recouvrement temporel dans l'analyse.

L'étape suivante du modèle est de calculer ce que Glasberg et Moore (2002) appellent la sonie « instantanée ». Les auteurs supposent que la sonie instantanée n'est pas quelque chose de conscient dans la perception. Elle correspondrait à l'activité totale dans le nerf auditif mesurée sur un intervalle de temps très court, de l'ordre de la milliseconde. Le calcul de la sonie instantanée se fait à partir du diagramme d'excitation, de la même manière que dans le modèle stationnaire (Moore et coll., 1997). Le diagramme d'excitation est ensuite transformé en sonie spécifique, et l'aire contenue sous le diagramme de sonie spécifique donne la sonie instantanée pour le temps considéré. Ce calcul vaut pour une prise de son



monaurale. Pour une prise de son binaurale, la sonie instantanée est sommée sur les 2 voies pour donner la sonie instantanée globale.

Glasberg et Moore (2002) calculent ensuite la sonie à court terme (STL : Short-Term Loudness) à partir de la sonie instantanée. La sonie à court terme représente la sonie ressentie pour un bref segment de son, par exemple une syllabe. Pour ce faire, la sonie instantanée est modifiée par une opération proche du fonctionnement d'un contrôleur automatique de gain, afin de tenir compte du phénomène d'intégration temporelle (avec une constante de temps de l'ordre de 100 ms) et du masquage temporel (décroissance de la sonie après la fin du stimulus). La sonie à court terme représente la sonie perçue à chaque instant. Les auteurs préconisent de prendre le maximum de la sonie à court terme (STLmax) comme valeur indicatrice de la sonie globale d'un son court.

Les auteurs proposent aussi de calculer la sonie à long terme (LTL : Long-Term Loudness). Le LTL est calculé à partir du STL, et correspond à l'impression globale de sonie ressentie pour un son relativement long, par exemple une phrase ou un morceau de musique. A partir de la sonie à court terme, la sonie à long terme (LTL) est calculée en utilisant aussi une sorte d'intégration temporelle similaire cette fois aussi à un contrôleur automatique de gain, mais avec des constantes de temps plus longues que pour le STL. La sonie à long terme est censée représenter le fait que la sonie globale d'un son est gardée en mémoire pendant quelques secondes après la fin de ce dernier en l'absence de nouveau stimulus. Le maximum de la sonie à long terme (LTLmax) est préconisé par Glasberg et Moore pour estimer la sonie globale de sons variant lentement dans le temps.

3.3 Modèle de sonie pour les sons impulsionnels

Boullet (2006)

La norme NF S31-010 (1996) définit un bruit impulsionnel comme un « bruit consistant en une ou plusieurs impulsions d'énergie acoustique, ayant chacune une durée inférieure à environ 1 s et séparée(s) par des intervalles de temps de durée supérieure à 0,2s ».

Boullet (2006) a étudié plus précisément la sonie des sons impulsionnels présentant une attaque rapide immédiatement suivie d'une décroissance d'amplitude de forme exponentielle. L'estimateur développé lors de ces travaux permet de calculer la sonie pour ce type de sons, et son principe est schématisé sur la Figure 18 et décrit dans ce qui suit.

Tout d'abord, le signal acoustique est filtré pour obtenir un signal temporel dans chacune des 24 bandes critiques. L'étape suivante du modèle est de calculer l'énergie et le temps de descente pour chaque bande critique. La sonie de cœur, dont le calcul est basé sur des travaux de Boullet (2005), est le produit de l'énergie à l'exposant alpha et de la durée de l'extinction de l'impulsion, à l'exposant bêta. Enfin, l'effet de masquage fréquentiel est pris en compte de la même manière que Zwicker et Fastl (1999). La densité de sonie ainsi obtenue est intégrée sur les 24 bandes critiques pour obtenir la sonie globale du son impulsionnel.

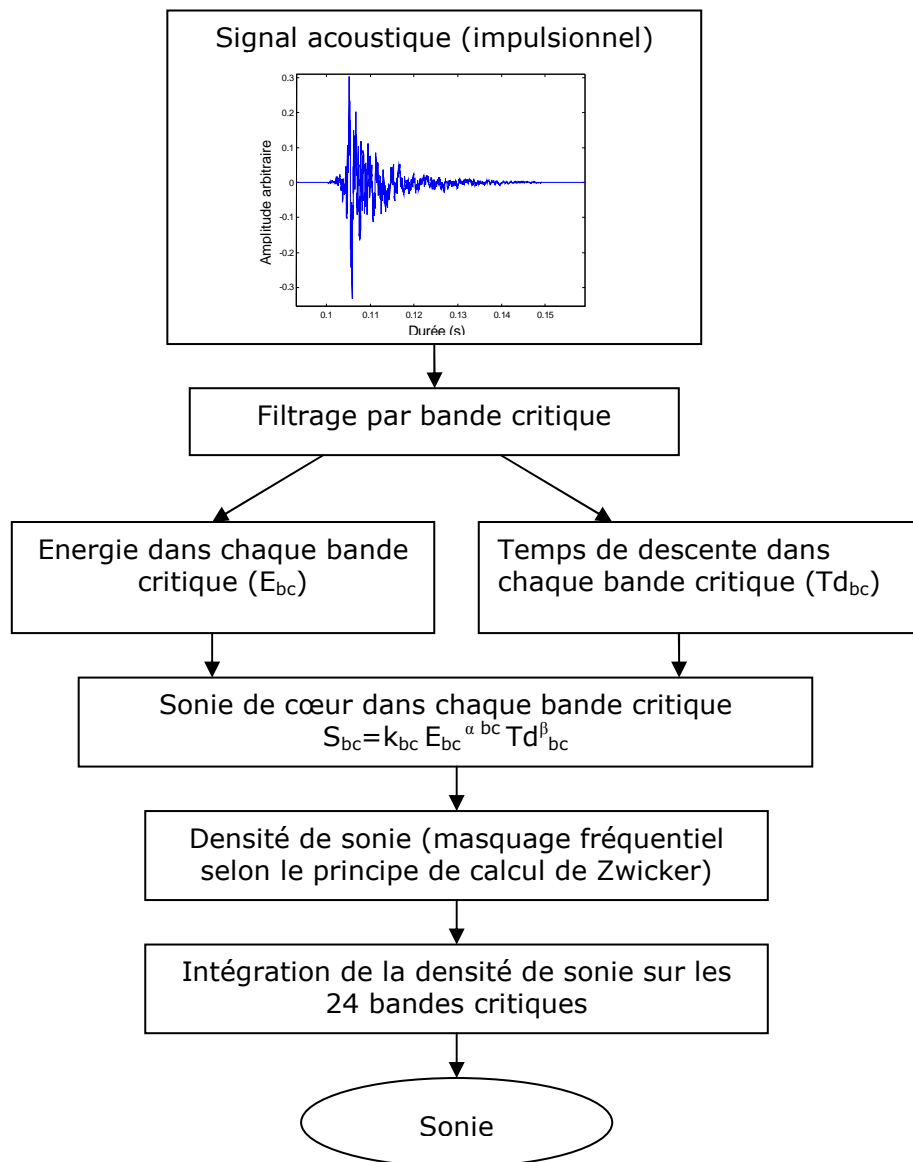


Figure 18: Etapes du calcul de la sonie pour les sons impulsionnels



4 Indicateurs de sonie pour les sons non stationnaires

Cette section présente les indicateurs, issus des modèles présentés précédemment, qui permettent d'estimer la sonie globale pour des sons variant dans le temps et impulsionnels.

4.1 N4, N5 ou N7 selon Zwicker et Fastl (1999)

Le modèle de Zwicker et Fastl (1999), présenté en section 3.2, permet de calculer la sonie (et le niveau d'isophonie) en fonction du temps. Différents indicateurs sont préconisés par les auteurs pour estimer la sonie globale d'un son. Il s'agit d'indices statistiques comme le N7 (utilisé pour des signaux de parole), N5 (bruit de l'environnement) ou N4 (trafic routier).

De manière générique, l'indicateur Nx représente la sonie dépassée pendant x pourcent du temps.

4.2 STLmax et LTLmax selon Glasberg et Moore (2002)

Le modèle de Glasberg et Moore (2002), présenté en section 3.2, permet de calculer la sonie en fonction du temps de sons variant dans le temps. Les auteurs préconisent de calculer la valeur maximale du niveau d'isophonie à court terme (noté STLmax) pour approcher le niveau d'isophonie global de sons variant dans le temps. Le niveau d'isophonie à court terme est le niveau d'isophonie perçu à chaque instant.

Le maximum de la sonie à long terme (LTLmax) est préconisé par Glasberg et Moore pour estimer la sonie globale de sons stationnaires ou variant lentement dans le temps. La sonie à long terme (LTL) traduit le fait que la sonie globale d'un son est gardée en mémoire après la fin de ce dernier pendant quelques secondes et en l'absence de nouveau stimulus.

4.3 LMIS selon Boulet et coll. (2006)

Le LMIS (Loudness Model for Impulsive Sounds), présenté en section 3.2, permet de calculer la sonie de sons impulsionnels. Il a été validé sur une série de sons impulsionnels dont les niveaux crêtes variaient de 56 à 92 dB SPL dont les durées étaient comprises entre 10 et 1620 ms.

ANNEXE 1 : Références bibliographiques

- [1] "American National Standard Procedure for the Computation of Loudness of Steady Sound", 2007, réf. ANSI S3.4-2007
- [2] « Méthode de calcul du niveau d'isosonie », norme internationale, 1975, réf. ISO 532
- [3] Boullet I., "La sonie des sons impulsionnels : Perception, Mesures et Modèles", Thèse de Doctorat de l'Université Aix Marseille II, spécialité Acoustique, 2005.
- [4] Boullet I., Gagneux F., Rabau G., Meunier S. et Boussard P., "Un estimateur de sonie d'impulsion", Congrès Français d'Acoustique, Tours, 2006.
- [5] DIN 45631, "Procedure for calculating loudness level and loudness", 1991
- [6] DIN 45631/A1, "Calculation of loudness level and loudness from the sound spectrum - Zwicker method - Amendment 1: Calculation of the loudness of time-variant sound", 2008
- [7] Glasberg, B. R, and Moore, B. C. J. "Prediction of absolute thresholds and equal-loudness contours using a modified loudness model", J. Acoust. Soc. Am. **120**, 585-588, 2006
- [8] Glasberg, B. R, and Moore, B. C. J. "A model of loudness applicable to time-varying sounds", J. Audio Eng. Soc, Vol.50, n°5, 331-342, 2002
- [9] Levitt H., "Transformed up-down procedures in psychoacoustics" J. Acoust. Soc. Am. **49**, 467-477, 1971.
- [10] Moore B. C. J. et Glasberg B. R., "A revision of Zwicker's loudness model", Acustica, Vol. 82, 335-345, 1996
- [11] Moore B. C. J., Glasberg B. R. et Baer T., "A model for the prediction of thresholds, loudness and partial loudness", J. Audio Eng. Soc, Vol.45, n°4, 224-240, 1997
- [12] Zwicker E., "Program for calculating loudness according to DIN 45631 (ISO 532B)", J. Acoust. Soc. Jpn. 12 (1), 39-42, 1991
- [13] Zwicker E., "Über psychologische und methodische Grundlagen der Lautheit, Acustica", Vol. 8, 237-258, 1958
- [14] Zwicker E., "Procedure for calculating loudness of temporally variable sounds", J. Acoust. Soc. Am., Vol. 62, n°3, p 675-682, 1977
- [15] Zwicker E., "Dependence of post-masking on masker duration and its relation to temporal effects in loudness", J. Acoust. Soc. Am., 75(1), p 219-223, 1984
- [16] Zwicker E. et Fastl H., "Psychoacoustics: Facts and models", 2nd Edition, Springer-Verlag, Berlin, 1999