Electronique et Signal pour la Musique

ENSEA

Sylvain REYNAL, ETIS, CNRS[1]
Hélène-Camille CRAYENCOUR, LSS, CNRS[2]
Bertrand DAVID, Telecom ParisTech
Intervenants et ex-intervenants :
Geoffroy PEETERS, Telecom ParisTech
Dogac BASARAN, IRCAM[3]
Romain HENNEQUIN, Deezer
Andrea VAGLIO, Deezer
Laurent SAID, Augmented Acoustics
Thomas HEZARD[4], Music World Media

2019-2020

Site internet de l'option sur
`http://www-reynal.ensea.fr`
rubrique "Enseignement"

1. + Institut Actes Université Paris 1 Sorbonne
2. CNRS/CentraleSupélec/Université Paris 11
3. Equipe Analyse Synthèse IRCAM/CNRS
4. thomas.hezard@musicworldmedia.com

Table des matières

Chapitre 1

Introduction

1.1 Qu'est ce que le son ?

La réponse physique est la plus simple : c'est une vibration mécanique de fréquence au moins égale à une vingtaine de Hertz, et qui se propage dans un milieu, généralement l'air. Cette vibration se traduit par un écart de la pression instantanée $P(\vec{r}, t)$ au point $\vec{r}$ et à l'instant t par rapport à la pression d'équilibre P_0 (pression statique ou composante DC). Cet écart, qu'on nomme *pression acoustique*, est faible :

$$P(\mathbf{r}, t) = P_0 + p_a(\mathbf{r}, t) \tag{1.1}$$

avec les ordres de grandeur suivants :

— $P_0 = 10^5$ Pa ;

— $p_a^0 = 2.10^{-5}$ Pa, soit $20\mu Pa$, qui est la pression de référence (qui donne le 0dB, cf. définition des dB(SPL) ci-dessous) et correspond au seuil moyen d'audibilité ;

— $p_a = 20$ Pa au seuil de douleur.

Même au seuil de douleur le rapport entre la pression "acoustique" et la pression statique est faible : 2.10^{-4}. Cela signifie qu'on va pouvoir développer les équations au premier ordre en grandeurs acoustiques sans trop de soucis. En outre, il ne faut pas trop s'inquiéter pour nos tympans : un dispositif d'équilibrage de la pression statique est prévu via les trompes d'Eustache, qui permettent à la pression statique de s'équilibrer entre l'extérieur et l'oreille moyenne (qui communique avec le larynx via une trompe, cf. figure 1.1).

On peut également donner quelques autres ordres de grandeur relatifs aux ondes sonore :

— La vitesse de propagation : $c = 340$ ms^{-1}. Sans oublier que cette valeur dépend de la température et de l'humidité.

— L'intensité acoustique au seuil d'audition : $I_a^0 = 10^{-12}$ W/m^2

— Masse volumique $\rho = 1.3$kg/m^3 et impédance acoustique caractéristique de l'air : $\rho c = 400$ S.I.

On peut remarquer que l'impédance caractéristique est homogène au rapport pression/vitesse et que l'intensité acoustique est homogène au produit pression $\times$ vitesse. Le niveau sonore, la pression ou l'intensité acoustique correspondante se mesure habituellement en dB$_{SPL}$[1] définis par :

$$N_{\text{dB}} = 20 \log_{10} \frac{p_a}{p_a^0} = 10 \log_{10} \frac{I_a}{I_a^0} \tag{1.2}$$

L'écart de pression s'accompagne d'un écart en température et d'un écart en masse volumique par rapport à l'équilibre (P_0, T_0, ρ_0). On peut alors facilement comprendre le mode

1. Sound Pressure Level

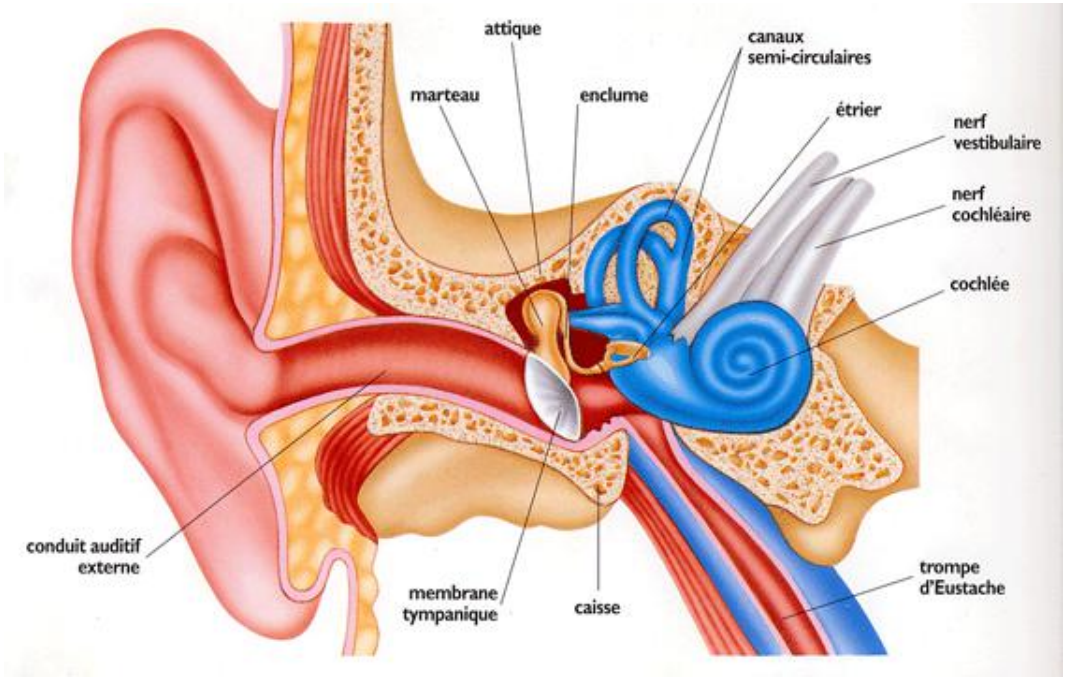


FIGURE 1.1 – Anatomie de l'oreille montrant l'oreille externe, l'oreille moyenne connectée au larynx via les trompes d'Eustache, l'oreille interne composée du système vestibulaire (centre de l'équilibre) et du système cochléaire (centre de l'audition).

de propagation "de proche en proche" du son, représenté figure 1.2. L'analogie est souvent donnée avec une chaîne de ressorts de même raideur accrochés à des masses toutes égales. L'expérience de la pompe à vélo dont on obstrue la valve offre d'ailleurs un parallèle avec un ressort comprimé.

La compression/dépression qui se produit au passage de l'onde acoustique correspond à un déplacement oscillant de petites tranches d'air. On ne s'intéressera pas ici au cas où il existe un écoulement, c'est à dire un déplacement d'ensemble superposé à cette oscillation (c'est par exemple le cas dans l'embouchure des flûtes). Pour nous, la vitesse acoustique sera la seule grandeur acoustique qui n'est pas une perturbation.

Cette définition du son, accrochée à la mécanique vibratoire et à la mécanique des fluides est cependant un peu réductrice. Le son, c'est aussi *ce qu'on entend*. Et étant donnée la faiblesse des perturbations mises en jeu, notre perception auditive est pratiquement la seule manifestation de sa réalité physique. Le son existerait-il si nos oreilles n'étaient pas là pour le traduire en influx nerveux dans le cerveau ? A moins que ce soit le contraire : la sélection aurait favorisé les espèces capables de percevoir ces faibles écarts, c'est-à-dire, possédant un système auditif... Si l'on s'en tient à cette approche anthropocentriste, le son voit son existence réduite à la tranche de fréquence $20 - 20000$Hz. Ces valeurs sont bien sûr une moyenne et sont affectées en général par l'âge du sujet (cf. presbyacousie par exemple). Notre perception des sons varie en outre avec la fréquence comme on peut le voir sur le *tracking de Békésy* représenté figure 1.3.

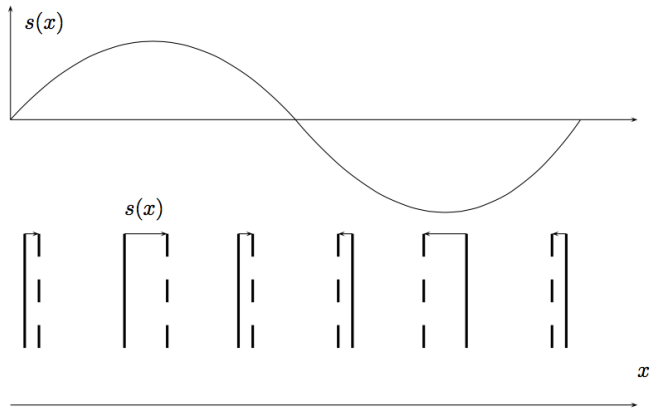


FIGURE 1.2 – Le son : une compression qui se propage de proche en proche ("photo" instantanée à l'instant t)

1.2 Acoustique et musique

Ces deux champs sont assez intimes dés le développement de la culture humaine sur terre. Ainsi :

— Préhistoire : on a retrouvé dans des grottes à la fois des phalanges de rennes servant de sifflet et des signes gravés à des emplacements correspondant à des résonances acoustiques de la grotte

— Pythagore découvre les rapports simples entre les modes de vibration d'une corde tendue.

Actuellement les deux disciplines continuent d'entretenir des rapports privilégiés. L'acoustique musicale est une recherche fondamentale dont l'objectif est de comprendre les mécanismes de fonctionnement des instruments et/ou de recréer ou de créer des sons musicaux.

La facture instrumentale s'est affinée par empirisme au cours des siècles. La compréhension des mécanismes de production du son dans les instruments permet de saisir le sens de son évolution. Cette étude physique ouvre des possibilités pour optimiser les procédés de fabrication ou les systématiser, pour concevoir de nouvelles lutheries (composite par exemple) ou encore pour rationaliser des phases d'enregistrement ou de sonorisation.

Enfin, l'acoustique des salles est devenu un élément incontournable de l'architecture des lieux d'écoute (auditorium, salle de concert, studio de répétition et d'enregistrement...).

1.3 Exemples

1.3.1 Mettons-nous au diapason

Petite expérience : on met en vibration le diapason en le frappant contre une surface dure. On peut alors l'écouter dans 3 conditions différentes :

— bras tendu : c'est faiblement audible

— en approchant les branches de l'oreille : on entend lorsqu'on est très proche et l'intensité varie lorsque l'on fait tourner le diapason

— avec le pied (du diapason !) posé sur une table : rayonnement acoustique beaucoup plus important

Explication : Le diapason émet un son quasi-pur à 440Hz. La longueur d'onde dans l'air

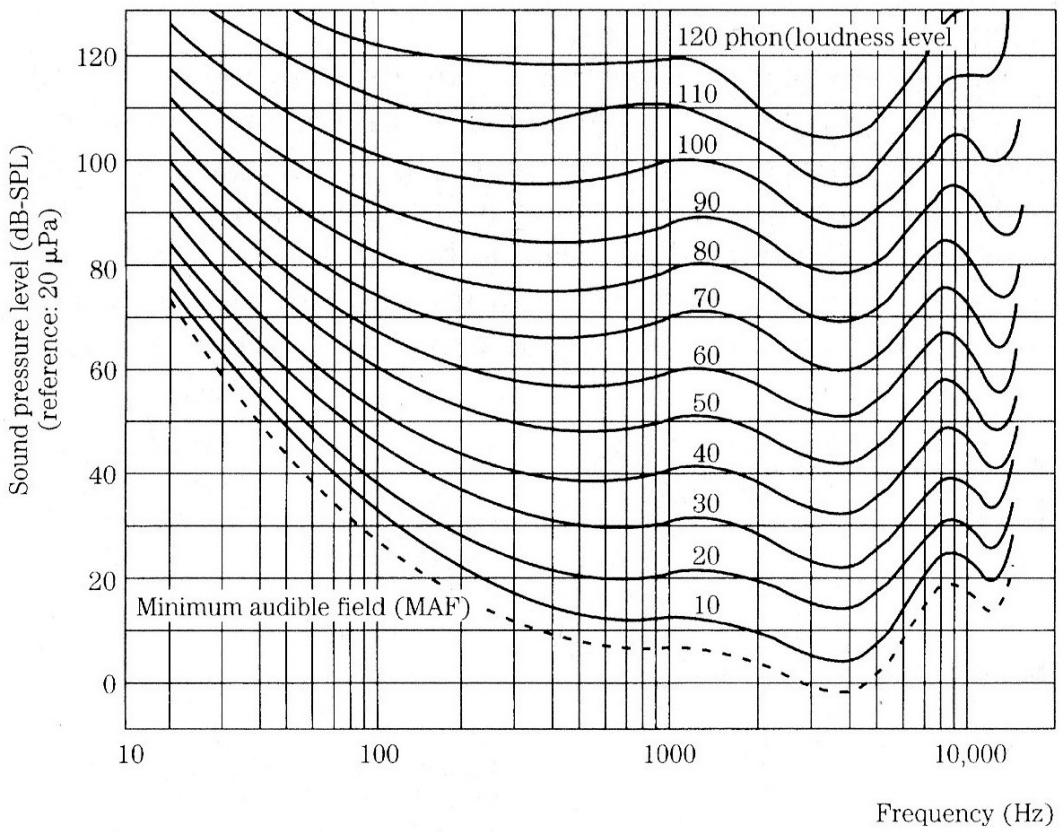


FIGURE 1.3 – Tracking de Békésy obtenu en laboratoire : le sujet ajuste en permanence le volume au seuil d'audibilité et la fréquence varie lentement et continûment. L'échelle des phonies indique l'intensité subjective (perçue) tandis que l'échelle SPL indique le niveau de pression en dB (0dB = 20 micro Pascal). On constate un écart de dynamique de près de 70dB entre 20Hz et 1kHz. C'est énorme et justifie d'ailleurs tout l'intéret des compresseurs dynamiques en mixage audio, mais également du codage psycho-acoustique de type MP3.

pour cette vibration est de l'ordre de 80 cm. La partie vibrante du diapason peut donc être considérée comme ponctuelle. Lorsqu'une branche de diapason vibre comme indiqué sur la figure 1.4, elle produit à chaque oscillation une compression à l'avant et une dépression à l'arrière de la branche. On a deux sources ponctuelles en opposition de phase, presque co-localisées vis à vis des dimensions caractéristiques de l'onde : c'est un **dipôle** acoustique. On comprend alors pourquoi à grande distance le rayonnement est peu efficace alors qu'en champ proche (longueurs largement inférieures à la longueur d'onde) on peut mieux entendre. Outre l'effet géométrique (l'intensité est plus forte quand on est plus près), on peut "séparer" les deux sources (l'intensité de la source la plus proche est prépondérante). Enfin, lorsque le diapason est couplé à la table, les vibrations (=l'énergie cinétique) se transmettent à celle-ci et ses dimensions sont supérieures à la longueur d'onde : elle peut rayonner cette fréquence efficacement.

Complément : le diapason comporte deux branches auxquelles correspond un dipôle. Ce qui nous fait deux dipôles côte-à-côte c'est-à-dire un quadripôle.

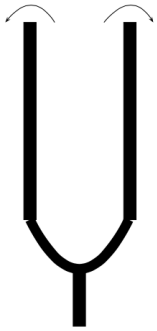


FIGURE 1.4 – Vibration d'un diapason. Les deux branches vibrent en opposition de phase (centre de gravité immobile). La source équivalente est un quadripôle (cf. chapitre 3).

1.3.2 C'est dans nos cordes

Remarque préliminaire : une introduction à la théorie de la musique à l'attention des non-musiciens est présentée dans le chapitre "Traitement du signal audio", en section 7.2.

L'étude des cordes vibrantes est le point de départ de l'histoire de l'acoustique (Pythagore) et, dans une certaine mesure, celui de la musique occidentale (création de la gamme diatonique). Comme une branche de diapason, une corde est incapable de rayonner efficacement, eut égard à son rayon R, faible devant la longueur d'onde λ. Pour obtenir une puissance sonore suffisante pour jouer de la musique (nécessité d'une projection du son vers les auditeurs et d'une dynamique de jeu), il faut, comme dans le cas du diapason, coupler la source de vibration avec un "radiateur" acoustique. C'est pourquoi tous les instruments à corde comprennent une table d'harmonie, à laquelle on adjoint souvent un résonateur (cavité).

Avec Pythagore est apparue la notion de *consonance*. Si on fait vibrer des cordes tendues avec la même tension, de même nature et dont les longueurs sont dans un rapport simple ($L_2/L_1 = p/q$ avec p et q petits), alors les sons obtenus sont "agréables à entendre". Une explication de cette consonance apparaît lorsqu'on considère les spectres obtenus lorsqu'on met en vibration une corde tendue entre deux points fixes. Ce spectre est composé du fondamental, qui correspond à une demi-longueur d'onde entre les deux points fixes ($L = \lambda/2$ soit $k_0 = \pi/L$) et de ses multiples, ou partiels ou encore harmoniques. Selon le point de pincement (au mediator, au plectre ou avec le doigt), on élimine les partiels qui correspondent à un nœud de vibration en ce point. C'est pourquoi le spectre sera plus riche (donc plus agressif, contenant plus d'aiguës) si on pince la corde près du chevalet d'une guitare.

Considérons maintenant une corde de longueur L donnée. La longueur d'onde du fondamental vaut, comme on l'a vu, $\lambda = 2L$ et comme la vitesse des ondes dans la corde vaut $c = \sqrt{\frac{T_0}{\mu}}$ (T_0 est la tension et μ la masse linéique) on obtient une fréquence fondamentale

$$f(Hz) = \frac{1}{2L}\sqrt{\frac{T_0}{\mu}}$$

Pour le mi aigu de la guitare (mi$_3$, le 3 indiquant le numéro de l'octave, cf. figure 1.5) par exemple, on a les ordres de grandeurs suivants : $T_0 = 80$ N, $\mu = 4.3$ kg/m^3 et $L = 65$ cm soit une fréquence d'environ 330 Hz. Les dix premiers partiels (ou harmoniques) sont situés

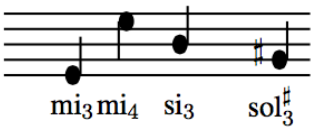


FIGURE 1.5 – Construction des notes les plus consonantes de mi_3 (par ordre décroissant de consonance), par génération de l'octave (2), de la quinte (3/2) et de la tierce majeure(5/4)

aux fréquences multiples de 330Hz, i.e.[2] :

partiels mi_3	1	2	3	4	5	6	7	8	9	10
f(Hz)	330	660	990	1320	1650	1980	2310	2640	2970	3300

Si on considère le mi_4, joué à l'octave (660 Hz), et qui correspond à une longueur de corde divisée par deux, il possède les partiels suivants :

partiels mi_4	1	2	3	4	5	...
f(Hz)	<u>660</u>	<u>1320</u>	<u>1980</u>	<u>2640</u>	<u>3300</u>	...

Ces deux notes possèdent un maximum de partiels communs (en fait tous les partiels de mi_4 sont dans mi_3, ce sont les partiels soulignés dans le tableau). L'octave (8 notes ou 12 demi-tons) est la note la plus consonante. En poursuivant dans cette voie (cf. figure 1.5), on trouve dans l'ordre :

— la quinte, correspondant à 5 notes ou trois tons et demi (ici mi-fa-sol-la-si, ce qui donne un si_3) : le rapport est de 3/2 sur les fréquences, ce qui donne :

partiels si_3	1	2	3	4	5	6	7	...
f(Hz)	495	<u>990</u>	1485	<u>1980</u>	2475	<u>2970</u>	3465	...

— puis la tierce majeure, située à 3 notes ou deux tons du fondamental (ici un $sol\sharp_3$), rapport 5/4 :

partiels $sol\sharp_3$	1	2	3	4	5	6	7	8	...
f(Hz)	412.5	825	1237.5	<u>1650</u>	2062.5	2475	2887.5	<u>3300</u>	...

Si la note étudiée est un Do, les notes de la gamme les plus consonantes seront le sol puis le mi. Do-mi-sol est l'accord "parfait", tel qu'il est référencé dans la nomenclature musicale (accord parfait majeur). Une autre manière de mettre la consonance en évidence est d'ailleurs de noter que, dans les premiers partiels de Do, on retrouve les (fondamentaux des) notes mi et sol jusqu'au sixième partiel. Si on considère Do_1 par exemple, on a :

partiels de Do_1	1	2	3	4	5	6	7	8	...
notes	Do_1	Do_2	Sol_2	Do_3	Mi_3	Sol_3	Si♭, "dissonant"	Do_4	...

De fait, c'est de cette manière que la gamme diatonique puis la gamme chromatique ont été créées, en allant de consonances en consonances, c'est à dire de quinte en quinte. On trouve ainsi la gamme de Pythagore, construite à partir de fréquences multipliées par la fraction 3/2. On peut, en parcourant le *cycle des quintes*, passer par toutes les notes de la gamme chromatique et revenir à la note de départ, ainsi qu'il est décrit sur la figure 1.6. Le

2. Exceptionnellement, je numérote le fondamental comme "partiel 1", contrairement à la convention utilisée en théorie du signal où le partiel 1 est en fait la première harmonique à $2f$.

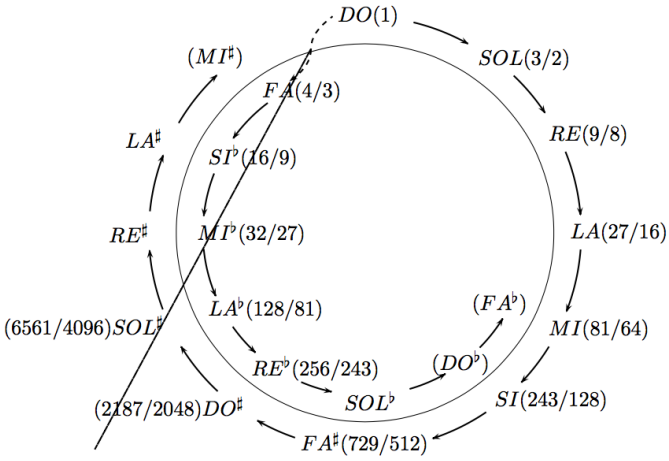


FIGURE 1.6 – Le cycle des quintes (gamme de Pythagore) : chaque note est obtenue à partir de la précédente en multipliant (sens horaire) ou divisant (sens trigonométrique) sa fréquence par $3/2$, puis en ramenant cette fréquence dans l'octave de départ par multiplication/division par 2^n. On notera que cette gamme (sans les notes diésées/bémolisées, i.e. la gamme majeure) est assez simple puisqu'elle ne possède que deux types d'intervalles, le ton (DO à RE = 9/8) et le **demi-ton diatonique** (SI à DO = $256/243 \sim 1.053$). Mais elle est relativement fausse : a) deux demi-tons diatoniques font moins qu'un ton ; b) la tierce (81/64) est supérieure (légèrement) à la tierce naturelle (5ème harmonique, i.e. 5/4), l'écart entre ces deux tierces étant appelé **comma naturel**. L'histoire se complique si on tient compte des notes diésées et bémolisées, qui de toute évidence sont légèrement décalées (par exemple La♯ et Si♭), l'écart étant ici appelé **comma de Pythagore**. Le **demi-ton chromatique** est l'écart entre, par exemple, La et La♯ ($3^7/2^{11} \sim 1.068$), et est légèrement plus élevé que le demi-ton diatonique.

rapport r entre la fréquence du Do de départ et celle du Do d'arrivée après le parcours du cycle peut se calculer de deux manières : on a parcouru soit 12 quintes, soit 7 octaves. Les deux calculs aboutissent à des résultats différents,

$$r_{quintes} = (3/2)^{12} = 129.7 \quad \text{et} \quad r_{octaves} = 2^7 = 128$$

En accordant sur les quintes, on ajuste les notes un peu trop haut pour que les octaves sonnent justes ! Ce problème d'accord n'a été résolu qu'avec l'apparition du tempérament égal, au XVIIème siècle. Ce tempérament réparti également l'erreur d'accord sur les douze demi-tons. Ces demi-tons doivent être dans un rapport de fréquence tel que $r_{1/2}^{12} = 2$ (pour avoir l'octave au douzième demi-ton) et vaut donc :

$$r_{1/2} = \sqrt[12]{2} \sim 1.059$$

Le demi-ton de la gamme tempérée est donc situé entre le demi-ton diatonique (1.053) et le demi-ton chromatique (1.068), cf. figure 1.6 pour la définition de ces deux termes.

1.3.3 Un exemple d'évolution de la facture : la guitare

◇ *Origines* Elles sont médiévales : la *guitarra moresca*, qui ressemble à un luth mais sonne plus aigu et *guitarra latina* qui ressemble à une viole avec un long manche et qui sonne entre

le luth et la guitarra moresca. Le vrai ancêtre de la guitare moderne est cependant la *vihuela* espagnole ou la *viola* italienne. Elle peut se jouer à l'archet, avec un plectre ou à la main (*vihuela de mano*). Elle possède déjà une table d'harmonie en épicéa mais elle est sujette à des difficultés d'accordage (gamme de Pythagore diatonique puis placement "à l'oreille" des demi-tons) et ce d'autant plus qu'elle comporte deux cordes par note.

Vers la fin du XVIème siècle guitarra et vihuela se confondent, la guitarra étant en général considérée comme une petite vihuela.

◇ *La guitare à 5 cordes* La vihuela devient obsolète au début du XVIIème. Elle est remplacée par la guitare à 5 cordes (doublées), dont la fameuse *chitarra battente* dont les cordes sont en métal. A cette époque, se développent de célèbres écoles de Luthiers en Europe, hors des espagnols et des italiens. Les écoles de Paris et de Hambourg sont les plus connues.

◇ *La guitare à 6 cordes* Vers la fin du XVIIIème, la cinq cordes (doublées) a disparu au profit de la 6 cordes simples. A partir du milieu de ce siècle, de gros changements dans les techniques de constructions vont mener à la guitare actuelle au début du XIXème. D'une part les italiens et les français imposent le modèle à 6 cordes simples accordées au tempérament égal et d'autre part les espagnols réalisent les plus gros progrès en utilisant de nouvelles essences (bois de rose, cyprès) pour le fond et les éclisses et surtout en adoptant les premiers un barrage de la table d'harmonie "en éventail". Celui-ci conserve une bonne rigidité à la table tout en améliorant notablement l'efficacité acoustique. Les frets du manche descendent à présent jusqu'à la rose.

◇ *Antonio de Torres et la guitare moderne* Il est considéré comme le "stradivarius" de la guitare. C'est de lui que viennent la plupart des caractéristiques de construction utilisées aujourd'hui au niveau des dimensions et des techniques d'assemblage. Les principales difficultés dans la facture de la guitare sont d'obtenir un instrument d'une puissance convenable et qui "chante" bien (clarté des aigus notamment). A. de Torres a fait et démontré beaucoup pour cela :

- le corps de l'instrument est agrandi et ces proportions fixées ; en particulier la hauteur de la table et la profondeur du manche. Le volume de la cavité est ainsi mieux adapté.
- Il améliore les dimensions du manche (plus large, plus épais) ce qui lui confère une plus grande facilité de jeu
- Il améliore sensiblement les techniques de barrage de la table, avec le barrage à 7 barres en éventail.

1.4 Acousto-devinettes

- Pourquoi faut-il construire des enceintes autour des HP ?
- Quel est le rôle de la caisse de résonance d'une guitare ?
- Pourquoi les tables d'harmonie comportent-elles un barrage ?
- Pourquoi des grands HP pour les basses et des petits pour les aiguës ?

Chapitre 2

Les équations de base

Elles sont le résultat combiné des équations de la mécanique des milieux continus (dont le jeu complet avec dissipation est appelé équations de Navier-Stokes) et des relations thermodynamiques pour la transformation adiabatique de l'air considéré comme un gaz parfait. Le caractère adiabatique des échanges est une hypothèse fondée sur la faible célérité des ondes thermiques par rapport à la célérité acoustique.

2.1 Un peu de méca flotte

2.1.1 Forces volumiques de pression

On appellera $\vec{f}$ la résultante des forces volumiques de pression :

$$\vec{f} = -\vec{\nabla} p$$

Pour trouver ce résultat, il faut faire un bilan des forces de pression s'exerçant sur un cube élémentaire de fluide de dimensions dx, dy, dz. Par exemple, selon l'axe Oz, la résultante des forces de pression s'écrit

$$dF_z = -p(z+dz)dS + p(z)dS = -\frac{dp}{dz}dzdS = -\frac{dp}{dz}dxdydz$$

(le signe moins provenant du fait que la force de pression s'exerce vers l'intérieur du cube).

2.1.2 Équation d'Euler

C'est simplement la transcription en mécanique des fluides de la relation fondamentale de la dynamique. On étudie ici l'acoustique linéaire non dissipative. On n'aura donc pas de terme de frottement. De plus les termes de pesanteur sont négligés devant les termes de forces de pression. Un moyen de l'obtenir est d'utiliser le paragraphe précédent et d'écrire l'expression de la dérivée particulaire de la vitesse $\vec{v}(M,t)$.
Rappels : La dérivée particulaire correspond à une dérivée de la grandeur en suivant le mouvement de la particule. Autrement dit cela correspond à une variation intrinsèque de la grandeur. La dérivée $\frac{\partial}{\partial t}$ quant à elle est une dérivée locale (à M constant). La description des vitesses par $\vec{v}(M,t)$ est une description par un champ de vecteur (la vitesse au point M à l'instant t). C'est ce qu'on appelle une description Eulérienne par opposition à description Lagrangienne, qui elle, s'intéresse à la vitesse d'une particule au cours de son mouvement.

On obtient après développement le résultat :

$$\vec{f} = -\vec{\nabla}p = \rho\frac{d\vec{v}}{dt} = \rho(\frac{\partial\vec{v}}{\partial t} + (\vec{v}\cdot\vec{\nabla})\vec{v}) \tag{2.1}$$

◇ *Conservation de la masse* On cherche une relation intégrale. On écrit que la variation de masse contenue dans un volume $\mathcal{V}$ fixe défini par un surface fermée $\mathcal{S}$ correspond d'une part à la variation interne de masse et d'autre part au flux de masse à travers la surface. On peut ainsi relier la variation de la masse contenue dans $\mathcal{V}$ au débit massique à travers la surface. Comme la relation obtenue doit être valable quelque soit le volume $\mathcal{V}$ on obtient en utilisant Ostrogradski

$$\frac{\partial\rho}{\partial t} + \mathrm{div}(\rho\vec{v}) = 0 \tag{2.2}$$

◇ *Transformation adiabatique d'un gaz parfait* Le but de ce calcul est de relier la pression p et la masse volumique ρ. Une relation simple entre ces deux grandeurs est l'équation d'état du gaz considéré comme parfait :

$$p = \rho\frac{RT}{\mathcal{M}} \tag{2.3}$$

où $\mathcal{M}$ est la masse molaire du gaz. On la prendra donc comme point de départ, en faisant intervenir la température T. Mais la température ne peut pas être considérée comme constante puisque l'onde acoustique se caractérise par une alternance de compressions-dilatations qui entraînent respectivement une élévation et une baisse locale de température, par rapport à la température d'équilibre T_0 . Du fait de la lenteur des échanges thermiques (ordre de grandeurs de la célérité d'une onde thermique dans l'air : 0,5m/s) on pourra en revanche, supposer en première approximation que les phénomènes acoustiques sont adiabatiques. On trouve alors une relation thermodynamique liant les variables d'état précitées :

$$p\rho^{-\gamma} = c^{te} = p_0\rho_0^{\gamma} \tag{2.4}$$

2.2 Linéarisation, équation d'Euler linéarisée

Les équations de l'acoustique linéaire non dissipative s'obtiennent par un calcul de perturbation à l'ordre 1. On a vu que les perturbations acoustiques étaient très faibles comparées aux valeurs statiques. On développe donc toutes les variables comme somme d'une grandeur au repos et d'une grandeur "acoustique" qui sera un infiniment petit d'ordre 1. Soient :

$$\begin{aligned} p &= p_0 + p_a \\ \rho &= \rho_0 + \rho_a \\ \vec{v} &= 0 + \vec{v_a} \end{aligned}$$

On suppose ici qu'il n'y a pas de mouvement des particules au repos. En se limitant au premier ordre, on trouve notamment **l'Equation d'Euler Linéarisée** :

$$\boxed{\vec{f} = -\vec{\nabla}p = \rho_0\frac{\partial\vec{v}}{\partial t}} \tag{2.5}$$

Cette équation sera utilisée presque systématiquement dans la suite de ce cours, notamment dans le chapitre suivant dédié aux impédances. En effet, pour une onde plane monochromatique, le gradient de la pression se traduit par une multiplication par jk , et la dérivée par

rapport au temps par une multiplication par $j\omega$. La notion d'impédance acoustique apparaît alors naturellement comme étant le rapport de p et v en régime harmonique.

Pour obtenir les équations de propagation, il suffit ensuite de développer les équations du paragraphe précédent (conservation de la masse, équation thermodynamique) sous leur forme locale, au premier ordre en grandeurs acoustiques. On en déduit alors les équations d'ondes suivantes :

$$(\frac{\partial^2}{\partial t^2} - c^2\vec{\nabla}^2)p_a = 0 \tag{2.6}$$

$$(\frac{\partial^2}{\partial t^2} - c^2\vec{\nabla}^2)\rho_a = 0 \tag{2.7}$$

$$(\frac{\partial^2}{\partial t^2} - c^2\vec{\nabla}^2)(\vec{\nabla}.\vec{v_a}) = 0 \tag{2.8}$$

Dans le cas où le champ des vitesses est irrotationnel à l'instant initial, il reste irrotationnel en permanence (il suffit de prendre le rotationnel de l'équation d'Euler) et alors on montre que la vitesse satisfait également une équation du type 2.6 pour chacune de ces composantes.

2.3 Énergie acoustique

Comme en mécanique du solide ou en mécanique du point, on peut faire une analyse énergétique en multipliant scalairement la relation fondamentale de la dynamique par la vitesse. Ainsi on fait apparaître le travail des forces et la dérivée de l'énergie mécanique, soit ici, en partant de l'équation 2.1 (équation d'Euler) on trouve :

$$\frac{\partial \mathcal{E}}{\partial t} + \operatorname{div} \vec{I} = 0 \tag{2.9}$$

avec la définition de deux grandeurs :

$$\mathcal{E} = \frac{\rho v_a^2}{2} + \frac{p_a^2}{2\rho_0 c^2} \quad \text{qui est un terme d'énergie volumique } (e_c + e_p)$$

et

$$\vec{I} = p_a\vec{v_a} \quad \text{qui est un terme de densité de courant d'énergie}$$

On en déduit que l'équation 2.9 précédente est simplement la traduction locale de la conservation de l'énergie. Par intégration sur le volume, on retrouve un terme de variation interne et un terme de flux, qu'on obtient par la formule d'Ostrogradsky.

Chapitre 3

Propagation et rayonnement

Ce chapitre s'intéresse aux solutions simples de l'équation des ondes acoustiques en pression de manière à introduire la notion d'impédance acoustique de rayonnement. Les deux cas étudiés — unidimensionnel et sphérique — sont la base de l'étude des sources acoustiques et permettent également de se faire une idée du spectre (modes propres de tuyau) de la plupart des instruments à vent, qu'on peut ramener à des formes simples (cylindres ou cônes).

3.1 Solutions des équations

3.1.1 Cas unidimensionnel

L'équation de propagation, supposée selon un axe Ox, s'écrit :

$$(\frac{\partial^2}{\partial t^2} - c^2 \frac{\partial^2}{\partial x^2}) p_a = 0 \tag{3.1}$$

C'est le cas archi-classique de l'équation de D'Alembert. Les solutions générales s'obtiennent par un changement de variable, qu'on trouve plus évident après factorisation de l'opérateur :

$$\frac{\partial^2}{\partial t^2} - c^2 \frac{\partial^2}{\partial x^2} = (\frac{\partial}{\partial t} - c\frac{\partial}{\partial x})(\frac{\partial}{\partial t} + c\frac{\partial}{\partial x})$$

Après avoir posé $X = x + ct$ et $Y = x - ct$ l'équation d'onde prend la forme simple :

$$\frac{\partial^2 p_a}{\partial X \partial Y} = 0$$

La solution générale est une somme de deux ondes planes progressive et rétrograde de la forme :

$$p(x,t) = f(x - ct) + g(x + ct)$$

(où p représente désormais la pression acoustique, *i.e.* la différence de pression par rapport à la pression atmosphérique).

◇ *Relation de dispersion* On cherche la relation entre le nombre d'onde k et la pulsation ω d'une onde *plane progressive sinusoïdale*, soit en notation complexe $p = p_0 e^{j\omega t - kx}$. En reportant cette expression dans l'équation d'onde on trouve pour ce type d'onde une relation

$$k = \frac{\omega}{c}$$

Cette relation est la *relation de dispersion* des ondes acoustiques et elle est ici linéaire : le milieu n'est pas dispersif. Elle relie les variations spatiales de l'onde (longueur d'onde $2\pi/k$) et les variations temporelle (période $2\pi/\omega$).

◇ *Impédance caractéristique* Le rapport $Z_a = \dfrac{p}{v}$ est appelé *impédance acoustique spécifique* et son inverse est l'admittance acoustique spécifique. Dans le cas d'une onde plane progressive sinusoïdale se rapport s'obtient facilement à partir de l'équation d'Euler et des expressions de la pression et de la vitesse en notation complexe. Il est constant et vaut :

$$Z_c = \rho_0 c$$

Le signe devient négatif pour une onde rétrograde. Z_c est *l'impédance caractéristique* du milieu. Pour l'air sa valeur est de l'ordre de 400 S.I. Comme en électronique le rapport des impédances entre deux milieux règle le transfert d'énergie de l'un vers l'autre. C'est pourquoi cette grandeur est importante pour la caractérisation d'un milieu du point de vue acoustique. La fameuse voix "Donald Duck" des plongeurs est un effet d'une variation de l'impédance caractéristique.

3.1.2 Cas sphérique

Ce deuxième cas simple permet de décrire par superposition de nombreuses sources acoustiques. En symétrie sphérique on aura aucune dépendance angulaire des grandeurs et l'expression du laplacien scalaire de la pression se réduit à

$$\vec{\nabla}^2 = \frac{\partial^2}{r\partial r^2}(rp)$$

On retrouve une équation de D'Alembert en rp dont les solutions générales sont :

$$p(r,t) = \frac{f(r-ct)}{r} + \frac{g(r+ct)}{r}$$

C'est la somme d'une onde sphérique progressive et d'une onde sphérique rétrograde. Le dénominateur montre un affaiblissement en $1/r$ qui traduit la conservation de l'énergie (qui est proportionnelle au carré de p). La solution sphérique progressive sinusoïdale correspondante s'écrit :

$$p(r,t) = p_0 \frac{e^{j\omega t - kr}}{kr}$$

pour rendre p_0 homogène à une pression (kr étant sans dimension).

La relation de dispersion s'obtient en reportant cette expression dans l'équation d'onde et on trouve toujours $k = \omega/c$.

◇ *Admittance spécifique* Le calcul de $Y(r)$ à partir de l'équation d'Euler de l'expression de la pression donne

$$Y(r) = Z_c^{-1}(1 - \frac{j}{kr}) \tag{3.2}$$

On peut alors séparer deux domaines. Lorsque la distance est grande devant la longueur d'onde, on se trouve en *champ lointain* et l'admittance spécifique est la même que celle en onde plane. Lorsque r est faible devant λ on se trouve en *champ proche* et alors $Y(r)$ est imaginaire pur (énergie instantanée qui tend vers l'infini en $r = 0$ mais énergie rayonnée nulle en moyenne, cf. paragraphe suivant sur l'intensité acoustique des sources).

3.2 Sources acoustiques

L'étude du monopôle et du dipôle acoustiques est la base du calcul du rayonnement acoustique dans le cas général, qui se ramène à une distribution de monopôles et de dipôles. D'autre part, par réciprocité, on peut transformer l'étude de ces sources en étude des récepteurs acoustiques (microphones). La source monopôlaire devient alors le micro-omnidirectionnel et le dipôle devient le micro-bidirectionnel.

3.2.1 Le monopôle (ou sphère pulsante)

Il s'agit simplement d'une source quasi-ponctuelle sinusoïdale, qu'on peut se représenter par une sphère pulsante dont le rayon a est petit devant la longueur d'onde(fig. 3.1), et qui génère un débit volumique $q_a(t) = 4\pi a^2 v_a(t)$. Ce modèle élémentaire nous permettra, par la suite, de construire un modèle de rayonnement pour le haut-parleur, appelé "modèle du piston plan encastré".

Le monopôle produit un champ de pression sphérique progressif et sinusoïdal tel que décrit par la solution sphérique précédente. On écrira de manière générale (lorsque c'est possible) :

$$p(r,\theta) = p(r)h(\theta)$$

avec ici

$$p(r) = p_0 \frac{e^{j(\omega t - kr)}}{kr} \text{ et } h(\theta) = 1$$

La fonction $h(\theta)$ est appelée *directivité* de la source. Ici le tracé de h ou **diagramme de directivité** est un cercle.

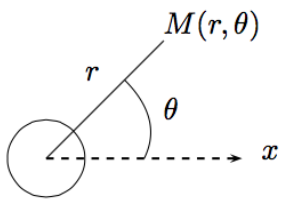


FIGURE 3.1 – Modèle de la sphère pulsante

La solution de l'équation de d'Alembert pour une géométrie à symétrie sphérique donne pour le débit :

$$\boxed{q(r,\theta) = 4\pi r^2 v(r,\theta) = q_a(1 + jkr)e^{-jkr}} \tag{3.3}$$

et pour le champ de pression :

$$\boxed{p(r,\theta) = jk\rho_0 c q_a \frac{e^{-jkr}}{4\pi r}} \tag{3.4}$$

qui est bien une solution sphérique de l'équation de propagation, avec $p_0 = jk^2\rho_0 c q_a/4\pi$ (cette condition étant indirectement fixée par la contrainte $q(r = 0) = q_a$ dans l'équation précédente).

Ces deux grandeurs admettent respectivement pour expression temporelle :

$$p(r,\theta,t) = \frac{k\rho_0 c q_a}{4\pi r} sin(\omega t - kr)$$

et

$$q(r,\theta,t) = q_a[\cos(\omega t - kr) + kr\sin(\omega t - kr)]$$

Pour passer de q à p et réciproquement, on peut utiliser l'équation d'Euler :

$$-\vec{\nabla}p = \rho_0 \frac{\partial \vec{v}}{\partial t}$$

soit, en sphérique :

$$-\frac{\partial p}{\partial r} = \rho_0 \frac{\partial v}{\partial t} = \rho_0 j\omega v = \frac{\rho_0 j\omega q}{4\pi r^2}$$

3.2.2 Dipôle acoustique

Lorsqu'on deux sources monopôlaires sont placées à une distance faible devant la longueur d'onde et en opposition de phase, on obtient un *dipôle acoustique*. En utilisant les notations données dans la figure 3.2 on trouve par un développement à l'ordre 1 en $k\delta r$ l'expression pour la pression :

$$p(r,\theta,t) = jp_0 k\delta r[1 - \frac{j}{kr}]\frac{e^{j\omega t - kr}}{kr} \tag{3.5}$$

avec les conditions $ka \ll 1$ et $a \ll r$. La dépendance angulaire apparaît dans l'expression de $\delta r = a\cos\theta$. On en déduit que la directivité du dipôle est

$$h(\theta) = \cos\theta$$

Ce qui, d'après la figure 3.3, justifie l'appellation de *bidirectionnelle* donnée à ce type de source.

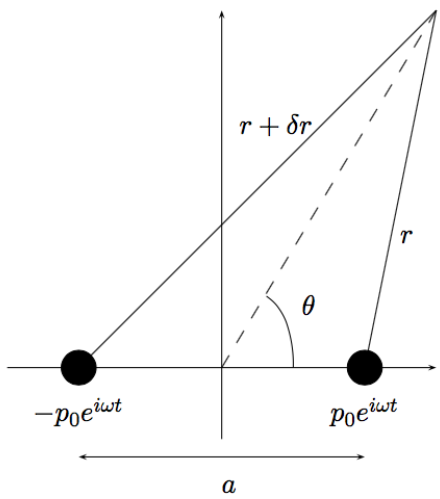


FIGURE 3.2 – Dipôle acoustique.

3.2.3 Source cardioïde

L'association d'un monopôle et d'un dipôle tels que décrits précédemment, placés tous deux à l'origine permet de réaliser une source cardioïde. Si on appelle p_0 l'amplitude des sources monopôlaires qui constituent le dipôle et qu'on prend

$$p_1 = -jkdp_0$$

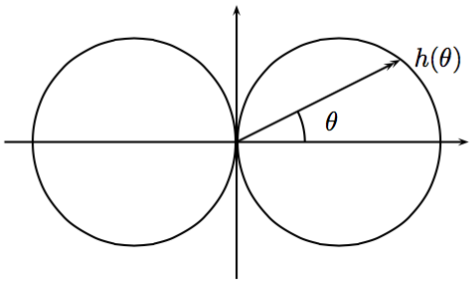


FIGURE 3.3 – Directivité du dipôle acoustique.

pour amplitude du monopôle alors la fonction de directivité en champ lointain pour l'association des deux sources prend la forme :

$$h(\theta) = 1 + cos\theta$$

La courbe correspondante, représentée figure 3.4 à l'allure caractéristique qui justifie son appellation.

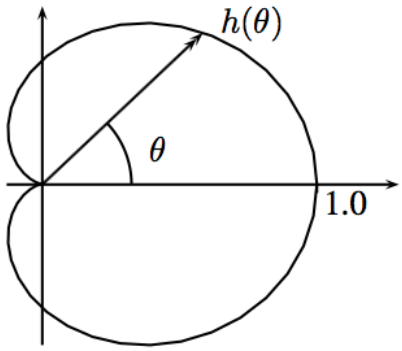


FIGURE 3.4 – Directivité cardioïde

3.2.4 Puissance et intensité acoustique d'une source

La **puissance acoustique** d'une source est définie par la valeur moyenne temporelle de la puissance instantanée

$$P(t) = F(t)v(t) = Sp(t)v(t) = p(t)q(t)$$

calculée au niveau de la source ($r = a$ dans le cas de la sphère pulsante (monopôle). Soit, dans le cas d'une source monochromatique,

$$\boxed{\mathcal{P}_{ac} = \langle p(t)q(t)\rangle = \frac{1}{2}Re[\underline{p}\ \underline{q}^*]} \tag{3.6}$$

Dans le cas du monopôle, on trouve donc comme puissance acoustique moyenne

$$\mathcal{P}_{ac} = k^2 q_a^2 \frac{\rho_0 c}{8\pi} \sim 33k^2 q_a^2\ (W) \tag{3.7}$$

ce qui montre que la sphère pulsante rayonne une énergie proportionnelle :

— au carré du débit ;
— au carré de la fréquence.

Par ailleurs, on notera que $\mathcal{P}_{ac}$ ne dépend pas de r, *i.e.* la puissance acoustique se conserve (et l'énergie aussi par la même occasion...). Ceci n'est vrai en toute rigueur, naturellement, qu'en négligeant les pertes dans le milieu (viscosité de l'air entre autre).

On définit également l'**intensité acoustique** par la relation :

$$I_{ac} = \frac{\mathcal{P}_{ac}}{S} = \rho_0 c (ka)^2 v_a^2$$

L'intensité acoustique est l'analogue acoustique du vecteur de Pointing défini en électromagnétisme, à la différence près qu'ici on travaille en puissance et non en énergie (la puissance acoustique est d'ailleurs le flux de l'intensité acoustique).

3.2.5 Impédance de rayonnement d'une source

L'**impédance de rayonnement** d'une source est définie par la relation :

$$\boxed{Z_R = \left.\frac{F}{v}\right|_{r=0} = S^2 \left.\frac{p}{q}\right|_{r=0} = S Z_a(r=0)} \tag{3.8}$$

où Z_a est l'**impédance acoustique spécifique** de l'onde rayonnée,

$$Z_a(r,\theta) = \frac{p(r,\theta)}{v(r,\theta)}$$

Pour la sphère pulsante, on a vu que l'impédance acoustique de l'onde rayonnée s'écrit

$$Z_a = Z_c \frac{k^2 r^2 + jkr}{1 + k^2 r^2} = R_a + jX_a$$

où $Z_c = \rho_0 c$.

En champ lointain ($kr \gg 1$), Z_a admet l'approximation $Z_a \sim \rho_0 c$, qui est l'impédance caractéristique de l'air (*i.e.* l'impédance d'une plane progressive monochromatique, ce qui est logique puisqu'en champ lointain, l'onde possède une structure d'onde plane) ; la réactance acoustique étant nulle, la puissance acoustique est donc uniquement active, ce qui correspond à une puissance effectivement transportée (*i.e.* vecteur de Pointing non-nul). En champ proche en revanche, $Z_a \sim \rho_0 c jkr$ et la puissance acoustique est purement réactive (vecteur de Pointing nul en moyenne, *i.e.* pas d'énergie transportée).

De l'expression de l'impédance acoustique, on déduit l'impédance de rayonnement

$$Z_R = 4\pi a^2 \rho_0 c \frac{k^2 a^2 + jka}{1 + k^2 a^2} = R_R + jX_R \tag{3.9}$$

En basse fréquence ($ka \ll 1$), $Z_R \sim 4\pi a^2 \rho_0 c jka$: la puissance rayonnée est purement réactive. En haute fréquence en revanche, $Z_R \sim 4\pi a^2 \rho_0 c$, et la puissance est essentiellement active : de l'énergie est donc effectivement rayonnée par la source.

Remarque 3.1 *On définit également l'**impédance acoustique spécifique réduite**, $z_a = Z_a/\rho_0 c$. Cette forme est souvent utilisée lors du calcul des ¡¡ impédances ramenées ¿¿.*

3.2.6 Facteur de directivité d'une source

Le facteur de directivité d'une source est défini par le rapport

$$\boxed{Q = \frac{p_{ax}^2(r)}{p_{sph}^2(r)}} \tag{3.10}$$

où $p_{ax}(r)$ est la pression émise par la source sur son axe de référence, et $p_{sph}(r)$ la pression émise par une sphère pulsante de même puissance. En général, on calcule Q à une distance $r = 1m$ de la source.

Un facteur de directivité de 1 correspond donc à la sphère pulsante. Un facteur élevé correspond à une source très directive.

Pour calculer Q, on peut utiliser le fait que $p_{sph}^2(1) = \rho_0 c\mathcal{P}_{ac}/4\pi$. On a alors :

$$Q = \frac{4p_{ax}^2(r)}{\rho_0 c\mathcal{P}_{ac}}$$

Il suffit de connaître la pression sur l'axe, à 1 mètre de la source, en fonction de sa puissance acoustique, pour déterminer Q.

Une autre méthode consiste à utiliser la **fonction de directivité en pression** $h(\theta)$. Elle est définie par la relation

$$p(1, \theta) = p_{ax}(1)h(\theta) \tag{3.11}$$

avec bien évidemment $h(0) = 1$. On montre alors que

$$Q = \frac{1}{\langle h(\theta)\rangle} = \left(\frac{1}{2}\int_0^\pi h^2(\theta)\sin\theta\, d\theta\right)^{-1}$$

Exemples :

— pour une source omnidirectionnelle, $h(\theta) = 1$ et $Q = 1$;

— pour un dipôle ou une source cardioïde, on a respectivement $h(\theta) = \cos\theta$ et $h(\theta) = 1 + \cos\theta$, et dans les deux cas $Q = 3$.

3.3 Propagation dans un tube

Le concept d'impédance ramenée permet, dans les problèmes de propagation unidimensionnelle, de calculer l'impédance acoustique spécifique en un point quelconque du système à partir de la connaissance de l'impédance acoustique spécifique en un point donné. Nous nous en servirons lors de la modélisation des microphones et haut-parleurs.

3.3.1 Impédance ramenée en un point quelconque d'un tube

Considérons un tube cylindrique de section S, dans lequel le champ des vitesse est donné par :

$$v(x, t) = Ae^{i(\omega t+kx)} + Be^{i(\omega t-kx)} = v_+ + v_-$$

Il s'agit de la superposition d'une onde progressive v_+ et d'une onde régressive v_-, c'est donc une onde stationnaire.

En chaque point d'abscisse x, la pression est donnée par l'équation d'Euler,

$$-\frac{\partial p}{\partial x} = \rho_0\frac{\partial v}{\partial t}$$

soit, pour l'onde progressive de pression,

$$p_+(x,t) = \rho_0 c v_+$$

et pour l'onde régressive,

$$p_-(x,t) = -\rho_0 c v_-$$

L'onde de pression s'écrit donc :

$$p(x,t) = \rho_0 c(v_+ - v_-) = \rho_0 c(Ae^{i(\omega t + kx)} - Be^{i(\omega t - kx)})$$

En chaque point d'abscisse x, on peut déterminer l'impédance acoustique spécifique de l'onde :

$$Z(x) = \frac{p(x,t)}{v(x,t)} = \frac{\rho_0 c(v_+ - v_-)}{v_+ + v_-}$$

soit, en remplaçant par l'expression de $v(x,t)$ et en simplifiant par $e^{i\omega t}$,

$$Z(x) = \rho_0 c \frac{Ae^{-ikx} - Be^{ikx)}}{Ae^{-ikx} + Be^{ikx)}}$$

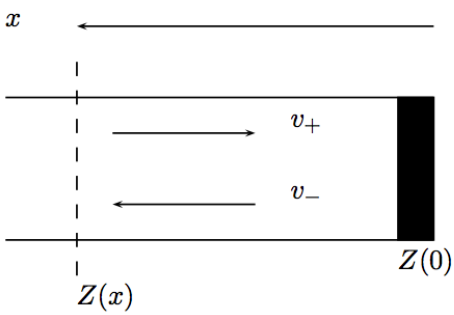


FIGURE 3.5 – Impédance de l'extrémité du tube ramenée en x.

Dans le cas d'un milieu non-dissipatif avec réflexion totale à l'extrémité (fig 3.5), on a naturellement $A = -B$: $Z(x)$ est imaginaire pure (pas de puissance active dissipée).

Notons $Z(0)$ l'impédance acoustique spécifique de l'extrémité située en $x = 0$:

$$Z(0) = \rho_0 c \frac{A - B}{A + B}$$

On peut alors exprimer $Z(x)$ en fonction de $Z(0)$:

$$\boxed{Z(x) = \rho_0 c \frac{Z(0) + i\rho_0 c \tan kx}{\rho_0 c + iZ(0) \tan kx}} \tag{3.12}$$

Cette équation permet donc de déterminer l'impédance de l'extrémité **ramenée** en n'importe quel point du tube.

On présente parfois cette formule sous la forme d'impédances réduites, $z = Z/\rho_0 c$:

$$z(x) = \frac{z(0) + i \tan kx}{1 + iz(0) \tan kx}$$

Impédance adaptée : si $Z(0) = \rho_0 c$, on voit que $Z(x) = \rho_0 c, \forall x$. On dit alors que l'extrémité du tube est **adaptée**. C'est ce que l'on cherche à faire dans le cas des **chambres de compression** (en effet - voir plus bas à propos des coefficients de réflexion - si l'impédance dans le tube est la même que dans l'air ¡¡ libre ¿¿, il n'y a pas de réflexion à l'interface haut-parleur/air libre, et la puissance acoustique est intégralement rayonnée).

Application à la détermination de modes propres : si le tube est parfaitement étanche en $x = 0$, la vitesse de l'onde en ce point est nulle, d'où $Z(0)$ est infini. L'impédance ramenée en x vaut donc :

$$Z(x) = \frac{i\rho_0 c}{\tan kx}$$

Si le tube est ouvert en $x = L$ (autre extrémité), l'impédance en ce point vaut $Z(L) = 0$ car la pression acoustique y est nulle (si l'on néglige le rayonnement du tube). Cette condition impose :

$$\frac{i\rho_0 c}{\tan kL} = 0$$

soit

$$\cos kL = 0$$

soit finalement

$$k_m = (2m+1)\frac{\pi}{L}, \; f_m = \frac{k_m c}{2\pi} = (2m+1)\frac{c}{2L}$$

Les fréquences f_m définissent les **modes propres** du tube (en négligeant le rayonnement à l'extrémité ouverte).

3.3.2 Changement de section dans un tube

Lorsqu'il y a changement de section (fig. 3.6), la conservation du débit impose **à l'interface** que :

$$v_1 S_1 = v_2 S_2$$

Au contraire, les pressions p_1 et p_2 sont égales à l'interface. On en déduit que l'impédance

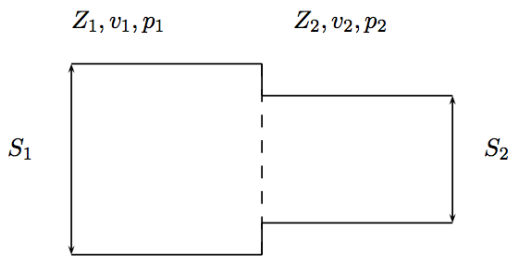


FIGURE 3.6 – Changement de section dans un tube.

Z_1 immédiatement à gauche de l'interface s'exprime en fonction de l'impédance Z_2 immédiatement à droite selon l'expression :

$$\boxed{Z_1 = \frac{S_1}{S_2} Z_2} \tag{3.13}$$

3.3.3 Réflexion/transmission d'une onde à l'interface entre deux milieux

Lorsque le milieu dans lequel se propage l'onde change brutalement d'impédance acoustique spécifique (par exemple : changement de densité ρ_0), il y a réflexion partielle ou totale de l'onde à l'interface. On considère, dans le milieu 1 d'impédance Z_1, une onde incidente

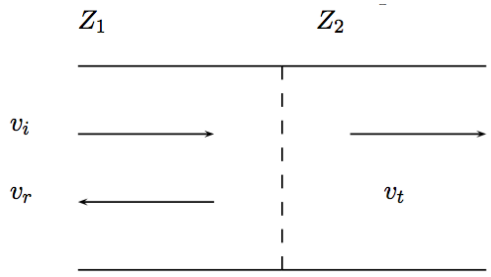


FIGURE 3.7 – Réflexion/transmission d'une onde à l'interface entre deux milieux.

(*i.e.* progressive) v_i (ou p_i pour la pression) et une onde réfléchie v_r, et une onde transmise v_t dans le milieu 2 d'impédance Z_2 (fig. 3.7).

A l'interface, la conservation des débit impose, **s'il n'y a pas simultanément changement de section** :

$$v_i + v_r = v_t$$

De même, l'égalité des pression à l'interface impose :

$$p_i + p_r = p_t$$

Or, on a pour chaque onde progressive, $p = Z_a v$, *i.e.*

$$p_i = Z_1 v_i,\ p_t = Z_2 v_t$$

et pour l'onde régressive,

$$p_r = -Z_1 v_r$$

(pour s'en convaincre, calculer p_r en fonction de v_r à l'aide de l'équation d'Euler : il y a bien un signe moins si l'onde est régressive). Compte-tenu de la relation liant les vitesses à l'interface, on trouve aisément :

$$R_v = \frac{v_r}{v_i} = \frac{Z_1 - Z_2}{Z_1 + Z_2},\ T_v = \frac{v_t}{v_i} = \frac{2Z_1}{Z_1 + Z_2}$$

pour les coefficients **en vitesse**, et

$$R_p = \frac{p_r}{p_i} = \frac{Z_2 - Z_1}{Z_1 + Z_2},\ T_p = \frac{p_t}{p_i} = \frac{2Z_2}{Z_1 + Z_2}$$

Comme la puissance acoustique de l'onde est donnée par $P_{ac} = p.v$, on en déduit les coefficients de transmission/réflexion en puissance (ou en intensité, ceux sont les mêmes) :

$$R_I = \left(\frac{Z_2 - Z_1}{Z_1 + Z_2}\right)^2,\ T_I = \frac{4Z_1 Z_2}{(Z_1 + Z_2)^2} \tag{3.14}$$

Quelques cas particuliers :

— Naturellement, on retrouve l'absence de réflexion si $Z_1 = Z_2$!

— Si $Z_2 \to \infty$ (milieu 2 parfaitement rigide), la réflexion est totale.

Chapitre 4

Oscillateurs mécaniques et acoustiques, analogies électriques

On ne considère, dans cette partie, que des longueurs d'onde suffisamment grandes pour que les systèmes étudiés puissent être modélisés par des circuits à constantes localisées, comme en électronique basse fréquence. On se propose d'étudier des modèles génériques d'oscillateurs mécaniques (système masse-ressort) et acoustiques (résonateur de Helmoltz) constituant les briques élémentaires pour la modélisation des transducteurs électromécaniques.

L'objectif de cette partie est de proposer, pour chaque oscillateur mécanique ou acoustique, un modèle électrique analogue : ainsi, l'intégralité du transducteur peut être modélisée comme un circuit électrique, pour lequel on dispose de puissantes méthodes d'analyse (fonctions de transfert, impédance). L'intérêt d'une telle approche est aussi (surtout ?) de permettre une compréhension intuitive du comportement du transducteur.

4.1 Le système masse-ressort

Un système masse-ressort est représenté figure 4.1. Si l'on ne tient pas compte de la présence du champ de pesanteur (dont l'effet se résume à modifier la position d'équilibre), la mise en équation donne :

$$m\frac{d^2x}{dt^2} + f'\frac{dx}{dt} + kx = F_{ex}(t)$$

où m est la masse suspendue, f' le coefficient de frottement fluide, et k la raideur du ressort. On appelle $1/k$ la **compliance** du ressort, et nous verrons que, dans l'analogie électrique, une compliance est remplacée par une capacité. En régime harmonique (régime sinusoïdal forcé), l'équation précédente s'écrit :

$$m(j\omega)^2x + f'j\omega x + kx = F_{ex}$$

Pour mettre en évidence la notion d'impédance, on exprime le premier membre de l'équation en fonction de la vitesse, *i.e.*

$$j\omega mv + f'v + \frac{k}{j\omega}v = F_{ex}$$

On voit que cette formulation permet de mettre en place l'analogie électrique suivante :

— une force est représentée par une d.d.p. (*i.e.* une **force** électromotrice),

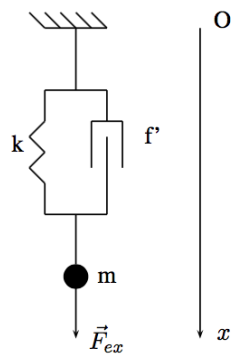


FIGURE 4.1 – Système masse-ressort.

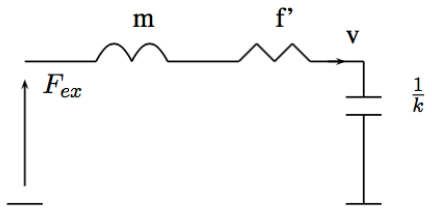


FIGURE 4.2 – Schéma électrique équivalent du système masse-ressort.

— une vitesse est représentée par un courant (*i.e.* une vitesse à un facteur de densité près),
— une masse est représentée par une inductance,
— une frottement fluide est représenté par une résistance,
— un ressort est représenté par un condensateur de capacité égale à la compliance du ressort.

ainsi que l'illustre le schéma électrique équivalent, figure 4.2.

On peut alors définir l'impédance mécanique du système, par la relation

$$\boxed{F_{ex} = Z_m v} \tag{4.1}$$

avec

$$\boxed{Z_m = jm\omega + f' + \frac{k}{j\omega}} \tag{4.2}$$

Cette impédance est donc le rapport d'une force (ce qui crée le mouvement) sur une vitesse (ce qui traduit le mouvement) : une impédance mécanique élevée caractérise donc un système qui résiste fortement au mouvement, soit parce qu'il possède une forte inertie (terme inductif de masse) ou des frottements élevés (terme résistif), soit parce que le ressort oppose une force de rappel importante (terme de compliance faible, donc raideur élevée). Comme l'inductance en électricité, la masse joue un rôle primordial à haute fréquence (caractéristique essentielle de l'inertie), tandis que le ressort, qui, en HF, n'a pas le temps de se comprimer, joue un rôle essentiel en basse fréquence. Ces trois paramètres sont donc fondamentaux pour la modélisation d'un transducteur.

Si l'on prend l'exemple du haut-parleur, la masse de l'équipage mobile constituée de la membrane, de la bobine et des fils, représente le terme inductif, tandis que la raideur de la membrane représente le terme de compliance (capacitif) : de la masse de l'équipage mobile

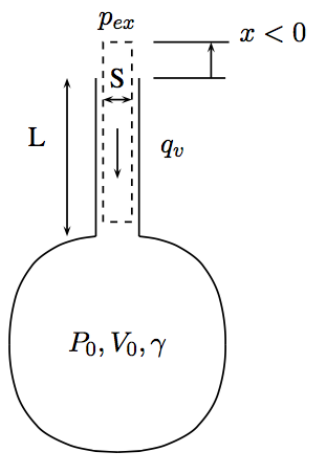


FIGURE 4.3 – Résonateur de Helmholtz.

dépendra donc la réponse en haute fréquence, et notamment la fréquence de coupure haute du haut-parleur. Toute ceci est naturellement intuitif dans ce cas simple : la technique utilisant l'analogie électrique prouvera réellement son intérêt dans le cas d'un HP bafflé, où le modèle fait intervenir plusieurs oscillateurs couplés.

Résumons les différentes analogies mécano-électriques :

F	$\leftrightarrow$	U
v	$\leftrightarrow$	i
x	$\leftrightarrow$	q
a	$\leftrightarrow$	$\frac{di}{dt}$
m	$\leftrightarrow$	L
f'	$\leftrightarrow$	R
$\frac{1}{k}$	$\leftrightarrow$	C

Remarque 4.1 *x, à l'instar de q, joue bien le rôle d'un potentiel : on a, pour le ressort $E_p = \frac{1}{2}kx^2$ et $E_p = \frac{1}{2}Q^2/C$ pour le condensateur. De même, v joue le rôle d'un courant, puisque $E_c = \frac{1}{2}mv^2$ pour la masse admet comme analogue $E_p = \frac{1}{2}Li^2$ pour l'inductance.*

4.2 Le résonateur de Helmoltz

Le résonateur de Helmoltz est l'analogue acoustique du système masse-ressort (cf. figure 4.3). Dans l'hypothèse basse fréquence, le résonateur peut-être modélisé par un circuit à constantes localisées, la masse de l'air contenu dans le goulot jouant le rôle d'une inductance, tandis que la compression de l'air contenu dans la cavité induit un effet capacitif. L'hypothèse *basse fréquence* consiste évidemment à négliger le temps de propagation des ondes acoustiques dans la cavité. D'autre part, on suppose que le volume de la cavité est suffisamment grand devant celui du goulot pour que la masse d'air contenue dans la cavité ait une vitesse négligeable (pas de terme inductif) ; symétriquement, on néglige la compression de l'air contenu dans le goulot. La masse d'air contenue dans le goulot vaut :

$$m = \rho_0 LS$$

où ρ_0 est la masse volumique de l'air, et L et S sont respectivement la longueur et la section du goulot. La première chose à faire est de modéliser l'effet "ressort" dû à la compression de

l'air de la cavité. En supposant la transformation adiabatique, *i.e.* $PV^\gamma = P_0V_0^\gamma = C^{te}$, on a, en notant dP et dV les variations de pression et de volume par rapport à P_0 et V_0,

$$\frac{dP}{P_0} + \gamma\frac{dV}{V_0} = 0$$

soit encore

$$dP = -\frac{\gamma P_0}{V_0}dV$$

où $dV = -Sx$ et $-dPS$ est la force qu'exerce l'air de la cavité (force comptée positivement vers le bas) sur la masse d'air du goulot.

Cet équation constitue l'analogue de la force de rappel pour un ressort. Si on appelle x le déplacement de la masse d'air du goulot par rapport à sa position d'équilibre (cf. figure 4.3), et p_{ex} la pression extérieure à l'extrémité du goulot, le principe fondamental de la dynamique appliqué à la masse d'air contenue dans le goulot s'écrit :

$$\rho_0 LS\frac{d^2x}{dt^2} + \frac{\gamma P_0 S^2}{V_0}x = p_{ex}S$$

Il s'agit bien de l'équation d'un oscillateur non-amorti, de fréquence propre

$$\omega_0 = \sqrt{\frac{\gamma P_0 S}{\rho_0 L V_0}} \tag{4.3}$$

En régime harmonique, on obtient, en utilisant le débit volumique (entrant) $q_v = vS$ à la place de la vitesse v de l'air, et en divisant par la section S du goulot,

$$\frac{\rho_0 L}{S}j\omega q_v + \frac{\gamma P_0}{V_0}\frac{q_v}{j\omega} = p_{ex}$$

En appelant $m_a = \frac{\rho_0 L}{S}$ la masse acoustique, et $C_a = \frac{V_0}{\gamma P_0}$ la compliance acoustique, on obtient finalement

$$m_a j\omega q_v + \frac{q_v}{j\omega C_a} = p_{ex}$$

Nous pouvons désormais établir des règles de correspondance (pour obtenir l'analogie électrique) sensiblement identiques au cas précédent, où seules les constantes "géométriques" différeront (la pression remplaçant la force, le débit volumique q_v remplaçant la vitesse, etc.) :

$m_a = \frac{\rho_0 L}{S}$	$\leftrightarrow$	L
f'	$\leftrightarrow$	R
$C_a = \frac{V_0}{\gamma P_0}$	$\leftrightarrow$	C
$q_v = vS$	$\leftrightarrow$	i
x	$\leftrightarrow$	q
p	$\leftrightarrow$	U

Le terme f' représente les frottements dus à la viscosité de l'air, et le frottement fluide de l'air sur les parois du goulot : il donne bien une force proportionnelle à la vitesse, ainsi que l'exige une frottement fluide. La figure 4.4 propose un schéma équivalent du résonateur ("un" parce qu'on peut inverser les tensions, et placer p_{ext} à droite au lieu de $-p_{ext}$ à gauche[1]). A noter que si l'oscillateur possède deux ouvertures opposées, respectivement soumises aux pressions p_1 et p_2, le circuit équivalent devient celui proposé figure 4.5.

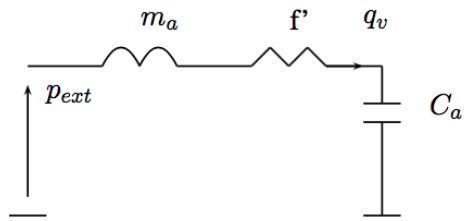


FIGURE 4.4 – Schéma équivalent du résonateur de Helmholtz.

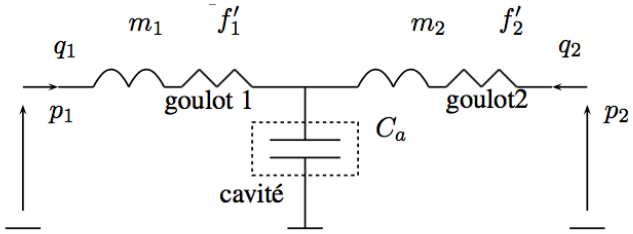


FIGURE 4.5 – Schéma équivalent du résonateur de Helmholtz. à deux goulots. Les débits sont définis entrant positivement.

Dans la lignée de l'exemple précédent, on définit l'impédance acoustique du résonateur par la relation :

$$\boxed{\underline{Z_a} = \frac{p_{ext}}{q_v} = jm_a\omega + f' + \frac{1}{jC_a\omega}} \tag{4.4}$$

où q_v est ici le débit entrant (sinon, changer le signe de Z_a).

4.3 Couplage d'oscillateurs

De nombreux dispositifs de transduction électro-mécano-acoustique peuvent être modélisés par un ensemble d'oscillateurs électriques, mécaniques (masse-ressort), et acoustique (résonateur) couplés. Nous allons tout d'abord dégager quelques grandeurs importantes relatives au couplage *génériques* de deux oscillateurs électriques : impédance libre, facteur de réciprocité, etc.

4.3.1 Couplage d'oscillateurs électriques

On considère le circuit électrique représenté figure 4.6. La mise en équation donne deux équations couplées :

$$\begin{aligned} U_1 &= Z_1 i_1 + Z_c(i_1 + i_2) &= Z_{11} i_1 + Z_c i_2 \\ U_2 &= Z_2 i_2 + Z_c(i_1 + i_2) &= Z_{22} i_2 + Z_c i_1 \end{aligned}$$

où $Z_{11} = Z_1 + Z_c$ est l'**impédance d'entrée à sortie ouverte**, et $Z_{22} = Z_2 + Z_c$ l'**impédance de sortie à entrée ouverte**. On définit également deux impédances importantes pour la suite,

1. Si vous avez tout compris, ça veut dire, en résumé, qu'il faut veiller au sens du débit : sur ma figure, il est positif en sortant.

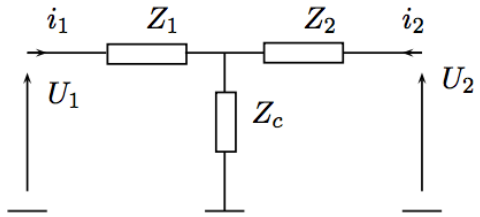


FIGURE 4.6 – Deux oscillateurs électriques couplés via Z_c.

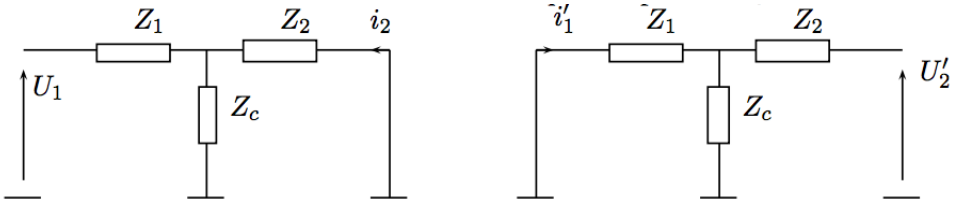


FIGURE 4.7 – Grandeurs caractéristiques pour la définition du facteur de réciprocité.

l'**impédance libre en entrée** (*i.e.* à sortie en court-circuit),

$$Z_e^* = \left.\frac{U_1}{i_1}\right|_{U_2=0} = Z_{11} - \frac{Z_c^2}{Z_{22}}$$

et l'**impédance libre en sortie** (*i.e.* à entrée en court-circuit),

$$Z_s^* = \left.\frac{U_2}{i_2}\right|_{U_1=0} = Z_{22} - \frac{Z_c^2}{Z_{11}}$$

Enfin, le **facteur de réciprocité**, définit selon les notations de la figure 4.7,

$$F = \left.\frac{i_2}{u_1}\right|_{U_2=0} = \left.\frac{i'_1}{u'_2}\right|_{U_1=0} = -\frac{Z_c}{Z_{11}Z_2^*} = -\frac{Z_c}{Z_{22}Z_1^*}$$

permet d'évaluer le "gain" du système lorsqu'il est utilisé en mode réciproque, *i.e.* lorsque l'entrée devient la sortie.

$$0 = Z_{11}i_1 + Z_c i_2$$
$$0 = Z_{22}i_2 + Z_c i_1$$

Ce système n'admet de solutions non-triviales que si son déterminant est nul, *i.e.*

$$Z_{11}Z_{22} = Z_c^2$$

En explicitant les diverses impédances en fonctions de ω, on obtient directement les pulsations propres et l'amplitude relative des modes propres (ici, i_2/i_1 est donnée par le rapport Z_c/Z_{11} pris pour chaque pulsation propre).

On considère le système masse ressort représenté figure 4.8, dont le comportement est régi par les deux équations couplées :

$$m_1\frac{d^2x_1}{dt^2} = -(k_1+k)x_1 + kx_2 + F_{ex,1}$$
$$m_2\frac{d^2x_2}{dt^2} = -(k+k_2)x_1 + kx_1 + F_{ex,2}$$

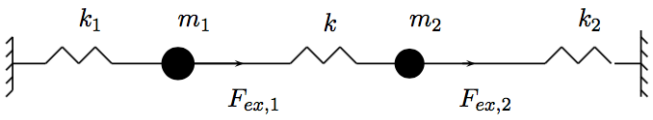


FIGURE 4.8 – Deux oscillateurs masse-ressort couplés par un ressort.

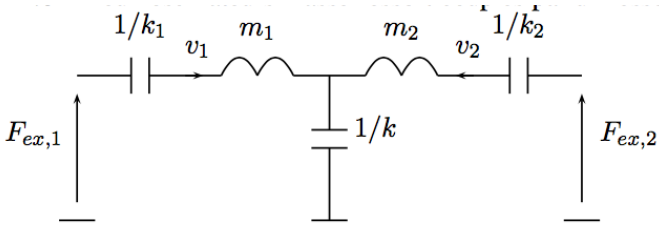


FIGURE 4.9 – Schéma équivalent au couplage de deux oscillateurs mécaniques couplés par un ressort.

Ce système mécanique admet pour schéma électrique équivalent le circuit de la figure 4.9, et on peut extraire de ce schéma les impédances suivantes :

$$Z_1 = j\omega m_1 + \frac{k_1}{j\omega}$$
$$Z_2 = j\omega m_2 + \frac{k_2}{j\omega}$$
$$Z_c = \frac{k}{j\omega}$$

On en déduit l'impédance d'entre à sortie ouverte,

$$Z_{11} = \left.\frac{F_{ex,1}}{v_1}\right|_{v_2=0} = j\omega m_1 + \frac{k_1 + k}{j\omega}$$

qui correspond au blocage de la masse 2 (par exemple, on l'attache à un support rigide) : dans le cas d'un haut-parleur, il s'agit de l'impédance d'entrée (électrique en l'occurrence) lorsque la membrane est bloquée. A titre d'exemple, exprimons également l'impédance libre en entrée, Z_e^* :

$$\begin{aligned} Z_e^* &= \left.\frac{F_{ex,1}}{v_1}\right|_{F_{ex,2}=0} \\ &= (j\omega m_1 + \frac{k + k_1}{j\omega}) + \frac{\left(\frac{k}{\omega}\right)^2}{j\omega m_2 + \frac{k+k_2}{j\omega}} \end{aligned}$$

Elle correspond à une sortie en court-circuit, *i.e.* à une force extérieure nulle sur m_2 : concrètement, dans le cas du haut-parleur, cela signifierait qu'aucune force extérieure n'agirait sur la membrane, ce qui imposerait de placer le haut-parleur dans le vide afin que l'impédance de rayonnement soit nulle (celle-ci traduit la "force" de résistance exercée par l'air sur la membrane lorsque celle-ci rayonne).

Un point à mettre en évidence dans l'allure de cette impédance, est la forme du couplage entre les deux oscillateurs : Z_e^* comporte deux termes, le premier est simplement l'impédance

d'entrée Z_{11} (m_2 bloquée), le second fait intervenir l'impédance Z_{22} via un terme de couplage (au numérateur) en $(1/\omega)^2$, qu'on qualifie de couplage à raideur prépondérante : à haute fréquence, le couplage est donc négligeable, et le système se comporte comme si m_2 était bloquée. Il existe d'autre types de couplage, par exemple à frottement prépondérant (le terme de couplage est constant en fréquence, le dispositif de couplage est un amortisseur), à inertie prépondérante (le couplage est en ω^2, le dispositif de couplage est par exemple une masse d'air déformable sans frottement), ou comportant un mélange de ces trois types de couplages (par exemple, un ressort de masse non-nulle, se déformant avec amortissement). Pour achever la caractérisation de ce couplage d'oscillateurs, il serait indispensable de déterminer les pulsations propres. Elles sont fournies par l'équation

$$Z_{11}Z_{22} = Z_c^2$$

vue plus haut, soit en explicitant,

$$(j\omega m_1 + \frac{k_1}{j\omega})(j\omega m_2 + \frac{k_2}{j\omega}) = \left(\frac{k}{j\omega}\right)^2$$

Ce qu'il faut retenir ici, c'est que ces pulsations propres détermineront directement la réponse fréquentielle du système, les résonances d'amplitude ayant lieu autour de ces deux pulsations : évidemment, déplacer ces résonances hors du spectre d'utilisation d'un haut-parleur est indispensable...

Chapitre 5

Le haut-parleur électrodynamique

Le schéma d'un haut-parleur électrodynamique à bobine mobile est indiqué figure 5.1. Le bobinage, alimenté par l'amplificateur de puissance, est solidaire de la membrane, qui est reliée, via la **suspension périphérique**, au **saladier** constituant le "socle" du haut-parleur.

5.1 Impédance mécanique

On pose $l = NL$, longueur totale de fil sur le bobinage. La bobine est soumise aux forces suivantes :

- à la force de Laplace : $F = Bli$;
- à une force de rappel due à la raideur de la suspension : $F = -kx$;
- à une force de frottement, que l'on modélise en première approximation par un frottement fluide de coefficient f' : $F = -f'v$;
- à une force résultant des surpressions acoustiques de part et d'autre de la membrane : $F = pS = -Z_R v$ où Z_R est l'**impédance de rayonnement** de la membrane (cf. détail plus loin).

L'équation mécanique s'écrit alors :

$$m\frac{dv}{dt} = Bli - kx - f'v - Z_R v$$

soit, en régime harmonique,

$$\boxed{Bli = (Z_m + Z_R)v} \tag{5.1}$$

où l'**impédance mécanique** du haut-parleur s'écrit

$$\boxed{Z_m = jm\omega + \frac{k}{j\omega} + f'} \tag{5.2}$$

Remarque 5.1 *Le terme Bl est appelé **facteur de force** du haut-parleur. C'est lui qui détermine l'intensité du couplage électro-mécanique.*

5.2 Impédance électrique

Une f.e.m. d'induction est crée lorsque la bobine se déplace à la vitesse $\vec{v}$ dans le champ magnétique crée par l'aimant : d'après la loi de Lenz, cette f.e.m. possède un signe tel qu'elle s'oppose, par ses effets, à la cause qui lui a donné naissance, *i.e.* le courant i (via la force de Laplace).

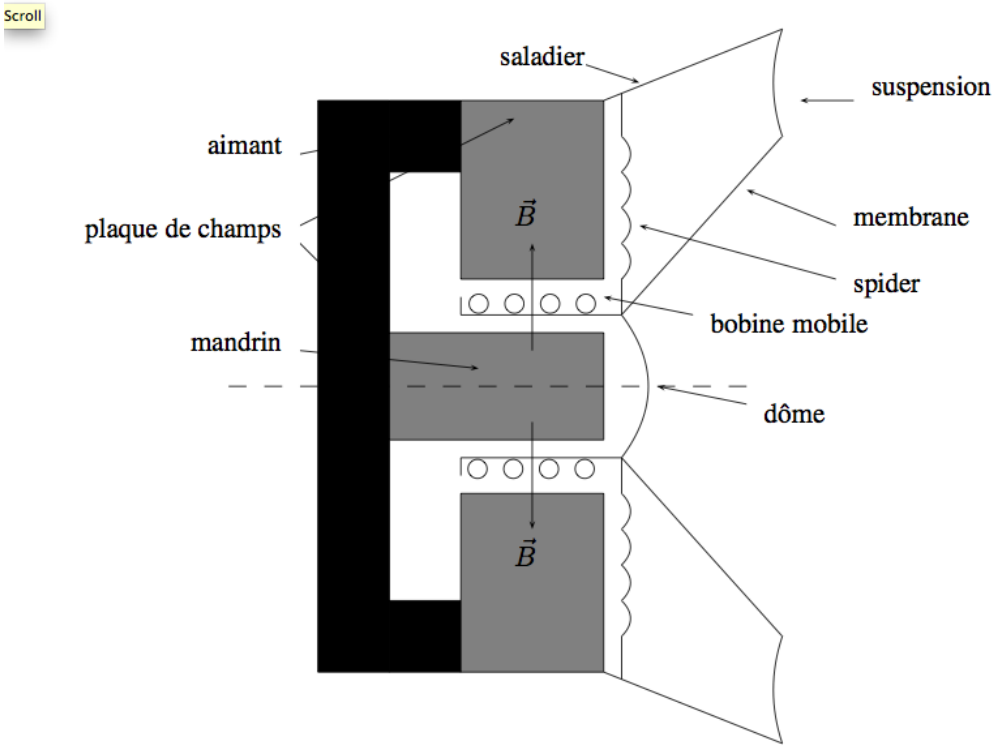


FIGURE 5.1 – Schéma d'un haut-parleur à bobine mobile.

On peut obtenir la valeur absolue de cette f.e.m. en utilisant la force de Lorenz $\vec{F} = q\vec{v} \times \vec{B}$ (on peut aussi utiliser la règle du flux coupé, pour ceux qui connaissent). $\vec{v}$ et $\vec{B}$ étant perpendiculaire (l'induction magnétique est — idéalement — radiale), on a en module $F = qvB$, et cette force est dirigée, en tout point, tangentiellement au bobinage : elle est donc équivalente à un champ **électromoteur** $\mathcal{E} = vB$, dont l'intégrale étendue à tout le bobinage (de longueur l) donne la d.d.p.

$$e = Blv$$

Le circuit électrique s'en déduit aisément (cf. figure 5.2) : il inclut les éléments électriques du bobinage (capacités parasites, inductance de bobinage, résistance des fils) et la f.e.m. induite (on notera qu'avec notre choix d'orientation de la source de tension associée à la f.e.m. induite, cette dernière s'oppose bien à la cause qui lui a donné naissance, *i.e.* le courant i).

En régime harmonique, on a

$$\boxed{U = Z_e i + Blv} \tag{5.3}$$

où

$$Z_e = R + jL\omega + \frac{1}{jC\omega}$$

est l'impédance électrique du haut-parleur.

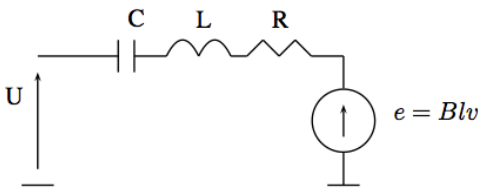


FIGURE 5.2 – Circuit électrique du haut-parleur à bobine mobile.

5.3 Impédance électrique libre

Les équations 5.1 et 5.3 forment un système de deux équations couplées, *i.e.* un système de deux oscillateurs couplés via la force de Laplace d'une part, la f.e.m. d'induction d'autre part, comprenant un oscillateur électrique constitué par L et C, et un oscillateurs mécanique de type masse-ressort constitué par la masse de l'équipage mobile et la raideur de la membrane :

$$\begin{aligned} U &= Z_e i + Blv \\ 0 &= -Bli + (Z_m + Z_R)v \end{aligned}$$

On notera que l'expression de Z_R dépend de l'environnement du HP : baffle et type d'enceinte, environnement de la salle, ...

L'impédance électrique libre s'obtient en combinant ces deux équations,

$$U = \left(Z_e + \frac{B^2 l^2}{Z_m} \right) i$$

C'est cette impédance que l'on mesure aux bornes électriques du HP, sauf si la membrane est maintenue immobile, auquel cas on ne mesure que Z_e.

On reconnaît dans cette impédance :

— le terme électrique Z_e dû au bobinage,

— un terme qui résulte directement du mouvement de la bobine, que l'on appelle **impédance motionnelle**, et dont l'importance est directement proportionnelle au facteur de force du HP.

5.4 Schéma en T

On peut aussi essayer de mettre les deux équations précédentes sous une forme telle que l'on retrouve notre circuit en "T" (cf. figure 4.6) correspondant au couplage de deux oscillateurs. Si l'on pose

$$Z_c = jBl$$

nos deux équations deviennent

$$\begin{aligned} U &= Z_e i + Z_c \frac{v}{j} \\ -Z_R \frac{v}{j} &= Z_c i + Z_m \frac{v}{j} \end{aligned}$$

On peut donc faire les correspondance suivante :

$$\begin{aligned} i_1 &\to i \\ i_2 &\to v/j \\ U_1 &\to U \\ U_2 &\to -Z_R \frac{v}{j} = -p_{ac}S/j \end{aligned}$$

ce qui montre au demeurant que la sortie est en court-circuit tant qu'on ne s'occupe pas de l'effet du rayonnement. On retrouve sous une forme très simple l'expression des impédances définies précédemment,

$$\begin{aligned} Z_{11} &= Z_1 + Z_c = Z_e \\ Z_{22} &= Z_2 = Z_c = Z_m \end{aligned}$$

ainsi que l'illustre la figure 5.3.

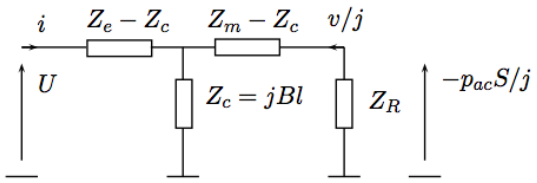


FIGURE 5.3 – Schéma en T du haut-parleur

Ce qui est intéressant dans ce formalisme, c'est que l'impédance libre en entrée Z_e^* s'identifie alors avec l'impédance d'entrée électrique calculée plus haut :

$$Z_e^* = Z_e + \frac{B^2 l^2}{Z_m}$$

5.5 Réponse en fréquence

A la résonance mécanique ($\omega^2 = k/m$), l'impédance mécanique Z_m s'annule, et l'impédance électrique libre Z_e^* passe par un maximum : autour de la résonance mécanique (toujours plus basse en fréquence que la résonance électrique, vue la valeur des éléments du circuit équivalent), l'impédance d'entrée est donc essentiellement d'origine mécanique. Une courbe typique de réponse en fréquence est présentée figure 5.4. On note la résonance mécanique autour de 20 Hz, et la prépondérance de l'inductance de bobinage aux hautes fréquences. Egalement aux hautes fréquence, **l'effet de peau** s'exerçant sur les fils du bobinage devient non-négligeable : il a pour effet d'augmenter la résistivité du bobinage.

A partir de nos deux équations, on peut également exprimer la condition de **linéarité en fréquence** de la réponse du HP. En l'occurrence, **l'efficacité** du dispositif est donnée par

$$M = \frac{q_v}{U} = \frac{BlS}{Z_e Z_m + B^2 l^2}$$

puisque c'est le débit volumique $q_v = vS$ généré par la membrane qui détermine l'intensité acoustique rayonnée (cf. rayonnement plus loin). Si l'on souhaite un fonctionnement **en**

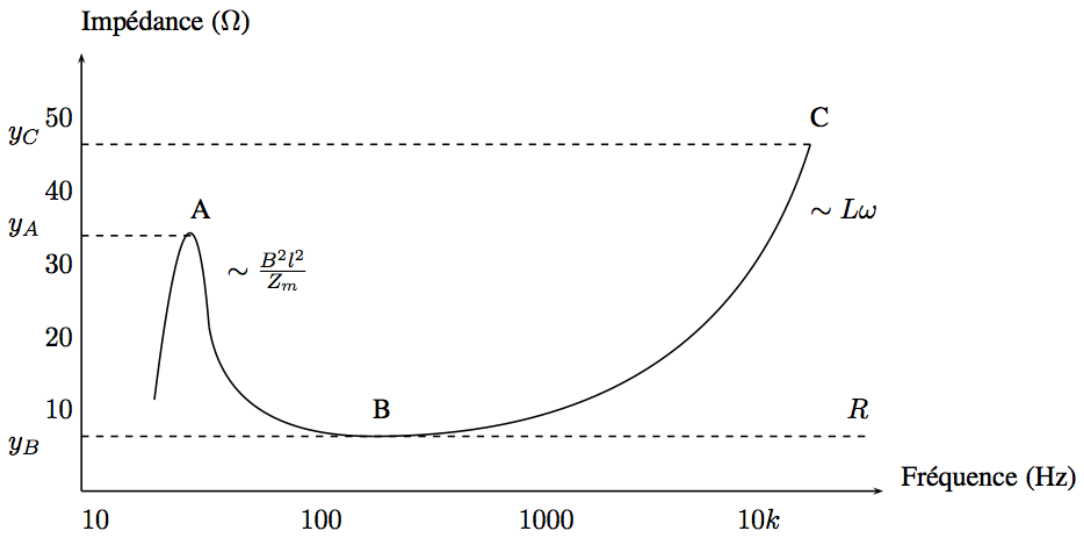


FIGURE 5.4 – Impédance électrique d'un haut-parleur électrodynamique.

vitesse du haut-parleur (*i.e.* v ne dépend pas de la fréquence), le facteur précédente doit être **réel** sur toute la plage de fréquence utile, ce qui implique :

$$Z_e Z_m \in \mathbb{R}$$

Si l'on admet que l'impédance électrique Z_e est essentiellement résistive dans la plage de fréquence utile, les conditions de linéarité se réduisent à :

$$Z_e \sim R, Z_m \sim f$$

soit en pratique

$$R \gg L\omega$$

et

$$f \gg m\omega \text{ et } k/\omega$$

On dit alors que le dispositif est à **frottement prépondérant**, puisque, dés lors que f domine l'expression de Z_m, c'est lui qui fixe également la valeur de l'efficacité :

$$M \sim \frac{BlS}{Rf + B^2l^2}$$

On pourra alors jouer sur les paramètres suivants :

— facteur de force Bl : à B fixé (généralement inférieur à $1,5T$), on doit augmenter la longueur de bobinage pour améliorer la linéarité ; simultanément, cela augmente R, ce qui est bénéfique à la linéarité en fréquence, mais élève l'impédance nominale, et donc impose à l'amplificateur de puissance une tension crête-à-crête plus importante (nécessité d'augmenter la tension d'alimentation, et de dimensionner les transistors en conséquent) ;

— un frottement f minimal (incompatible avec la condition de linéarité, un compromis doit donc être trouvé) ;

— une surface S importante (augmente le débit) ; toutefois, la surface est limitée par les phénomènes de directivité qui apparaissent lorsque la longueur d'onde acoustique devient inférieure au diamètre de la membrane (cf. rayonnement).

5.6 Mesures expérimentales

5.6.1 Mesures à partir du diagramme de Bode de l'impédance

On exploite les 3 points A, B et C du diagramme 5.4.

— au point C, on a $|Z|_c \sim \sqrt{R^2 + L^2\omega^2}$ puisque l'impédance motionnelle est négligeable à cette fréquence ;
— au point B, on a $|Z|_b \sim R$ (le problème est que le point B est difficile à localiser) ;
— au point A (résonance), on a $|Z|_a \sim R + B^2l^2/f$ car l'inductance est négligeable à cette fréquence.

On en déduit :

— $R \sim y_B$
— $L \sim \sqrt{y_C^2 - y_B^2}/\omega$
— $f \sim B^2l^2/(y_A - y_B)$
— $k/m = \omega_m^2$ (résonance mécanique)

Reste donc à déterminer le facteur de force et, au choix, la raideur de la suspension ou la masse de l'équipage mobile.

Les déterminations de la masse ou de la raideur peuvent être réalisées selon différentes méthodes :

— coller une masse additionnelle m' sur le dôme du haut-parleur, et mesurer le décalage de la fréquence de résonance : on a alors la masse recherchée m qui vaut à une bonne approximation $m = \frac{m'}{(\omega_m/\omega'_m)^2-1}$;
— coller une masse additionnelle m' sur le dôme, haut-parleur posé horizontalement, et mesurer le déplacement Δx de la membrane ; on récupère alors la valeur de $k = m'g/\Delta x$.

Enfin, le facteur de force Bl peut-être déterminé via la force de Laplace : la méthode de la balance consiste à mesurer la force nécessaire (masse M à poser sur le plateau de la balance) pour ramener la membrane à sa position d'équilibre après application d'un courant i dans le bobinage. On a alors $Bl = Mg/i$.

5.6.2 Données constructeur

Le tableau 5.5 indique les valeurs numériques correspondant à un haut-parleur de grave de fabrication moderne.

5.7 Rayonnement du HP

Le piston plan encastré (fig. 5.6) est un modèle simplifié de rayonnement du haut-parleur, valable dans une bande de fréquence où la membrane vibrante peut effectivement être approximée par un plan (*i.e.* pour des fréquences largement inférieures au mode fondamental de la membrane).

La membrane supposée ici circulaire, de rayon a, se déplace à une vitesse $v_a(t)$, et crée un débit $q_a = \pi a^2 v_a$. Le plan est supposé infini, et correspond au cas d'un haut-parleur bafflé

Paramètre généraux	Masse du HP	$10\,kg$
	Q total	$0,12$
	Niveau d'efficacité caractéristique (1W bruit rose pondéré)	$101\,dB(SPL)$
	Puissance nominale	$200\,W$
Paramètres électriques	Z_e nominale à 300 Hz	8Ω
	$\|Z_e\|$ minimale	7Ω à 250Hz
	R	5.5Ω
	Q_e	$0,13$
Paramètres mécaniques	Fréq. de résonance	15 ± 3 Hz
	Résistance mécanique	$2,49\,kg/s$
	Compliance suspension	$0,97.10^{-3}\,m/N$
	Masse mobile	$116\,g$
	Q_m	$4,39$
Rayonnement	Surface émissive	$0,0880\,m^2$
	Diamètre émissif	$33,5\,cm$
Paramètres magnétiques	Diamètre bobine	$10\,cm$
	Hauteur bobinage	$14\,mm$
	Induction dans l'entrefer	$1,5\,T$
	Flux dans l'entrefer	$3,3\,mWb$
	Facteur de force (Bl)	$26,04$
	Volume de l'entrefer	$3,298.10^{-6}\,m^3$
	Hauteur de l'entrefer	$7\,mm$
	Diamètre de l'aimant ferrite	$220\,mm$
	Hauteur de l'aimant	$23\,mm$
	Masse de l'aimant	$3,14\,kg$

FIGURE 5.5 – Haut-parleur de grave (Boomer), HD 38 S 100, 38 cm de diamètre.

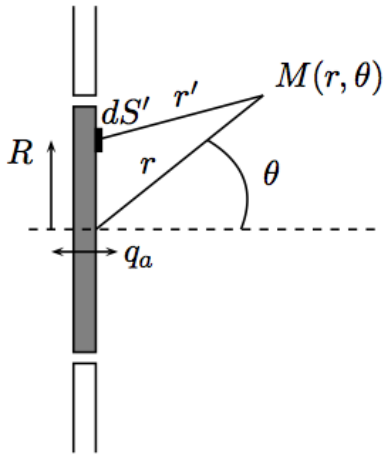


FIGURE 5.6 – Modèle du piston plan encastré

dans un baffle infini. On montre que l'encastrement du haut-parleur dans un tel baffle double la pression rayonné et la puissance acoustique.

Le principe du calcul du champ rayonné en tout point de l'espace consiste à décomposer le piston en sources ponctuelles. Chaque élément de surface dS' du piston créé un débit $dq' = v_a dS'$, et est assimilable à une source ponctuelle de type "sphère pulsante" de débit identique. Le champ de pression est alors la somme des pressions rayonnées par chacune de ces sources élémentaires :

$$p(r,\theta) = jk\rho_0 c v_a \iint_{S=\pi a^2} \frac{e^{-jkr'}}{4\pi r'} dS'$$

On distingue généralement, pour des raisons d'approximation, le champ acoustique sur l'axe du champ rayonné en tout à grande distance.

5.7.1 Champ acoustique sur l'axe

Lorsque M est sur l'axe, l'expression du champ de pression se simplifie, compte-tenu de $r' = \sqrt{a^2 + r^2}$, en :

$$p_{ax}(r) = \rho_0 c v_a \sin\left(\frac{kr}{2}(\alpha - 1)\right)$$

avec $\alpha = \sqrt{1 + a^2/r^2}$. Le champ rayonné par un HP de 38cm de diamètre ($a = 19cm$) est illustré pour deux fréquences, figures 5.7 et 5.8.

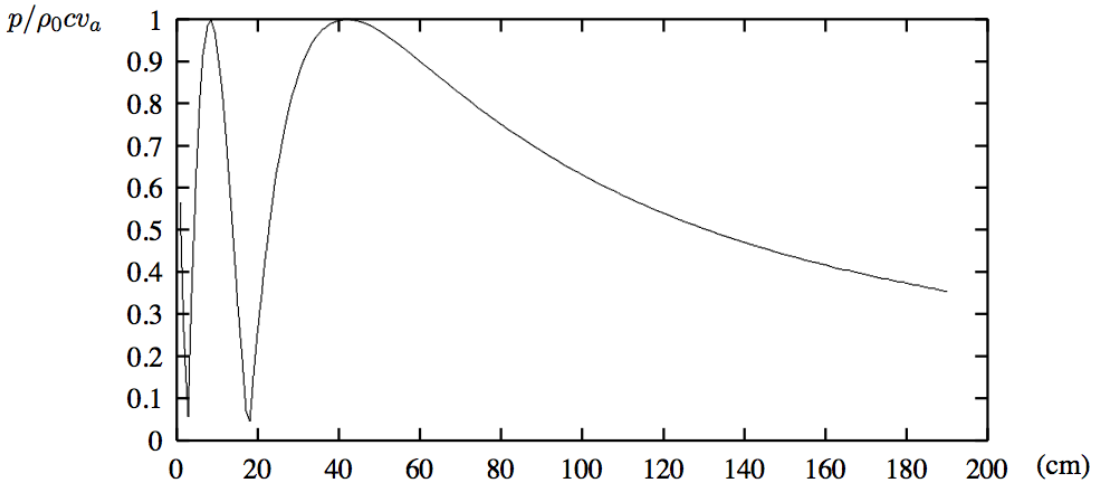


FIGURE 5.7 – Champ rayonné sur l'axe d'un HP de 38 cm de diamètre, à 4 kHz, soit $\lambda = 8,25cm$.

Les minima de pression sont situées à

$$r_m = \frac{a^2}{2m\lambda} - \frac{m\lambda}{2}$$

et les maxima en

$$r_n = \frac{a^2}{(2n+1)\lambda} - \frac{(2n+1)\lambda}{4}$$

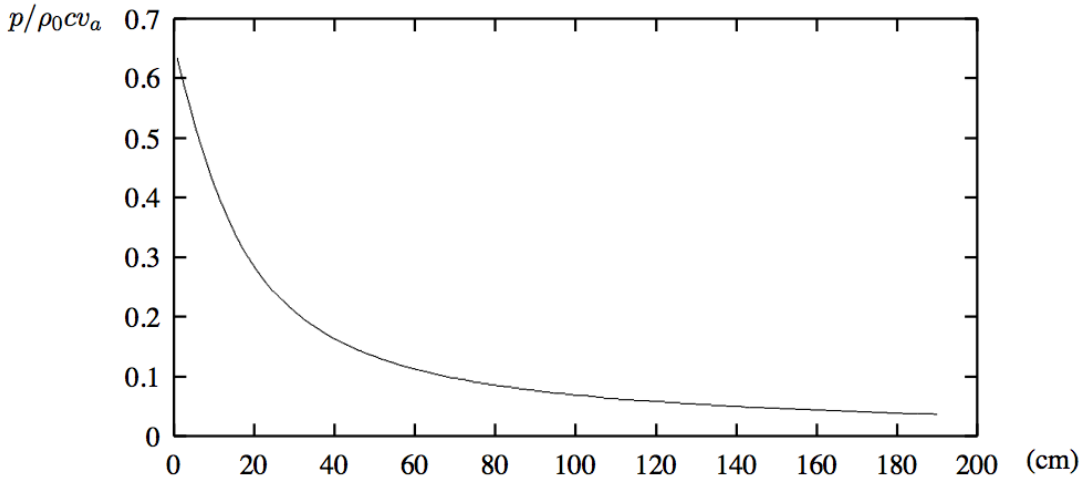


FIGURE 5.8 – Champ rayonné sur l'axe d'un HP de 38 cm de diamètre, à 400 Hz, soit $\lambda = 82.5cm$.

On notera qu'en basse fréquence (*i.e.* $\lambda > a$), il n'y a aucun minimum. En revanche, plus la fréquence augmente, plus le champ de pression varie rapidement en champ proche, ce qui rend, à haute fréquence, toute mesure de la pression sur l'axe extrêmement délicate.

En champ lointain ($r \gg a$), l'expression de $p_{ax}(r)$ est approximativement :

$$p_{ax}(r) \sim \rho_0 c v_0 \frac{ka^2}{4r}$$

Le champ de pression varie en $1/r$ (onde sphérique) : c'est un moyen de remonter à la pression rayonnée en chaque point de l'axe *en champ lointain*.

5.7.2 Rayonnement en champ lointain

Le rayonnement en tout point de l'espace n'admet une forme analytique qu'en champ lointain ($r \gg a$). Dans cette approximation, on peut écrire

$$r' \sim r - R \sin\theta \ \cos\phi$$

L'expression du champ de pression devient alors :

$$p(r,\theta) \sim jk\rho_0 c \frac{v_0}{2\pi r} e^{-jkr} \int_0^a \int_0^{2\pi} e^{jkR\sin\theta\cos\phi} d\phi \ R \, dR$$

soit, après calcul des intégrales :

$$\boxed{p(r,\theta) \sim \frac{1}{2}\rho_0 c v_0 \frac{ka^2}{r} \frac{2J_1(ka\sin\theta)}{ka\sin\theta}} \tag{5.4}$$

où $J_1(x)$ est la fonction de Bessel de première espèce et d'ordre 1.

On peut également extraire de cette expression la fonction de directivité en pression du HP *en champ lointain* :

$$h(\theta) = \frac{2J_1(ka\sin\theta)}{ka\sin\theta}$$

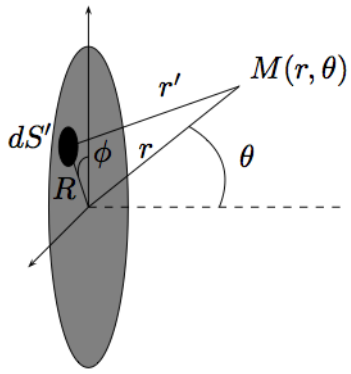


FIGURE 5.9 – Notations pour le calcul du champ rayonnée à grande distance.

Plusieurs diagrammes de rayonnement, pour différentes fréquences, sont représentés figures 5.10 et 5.11.

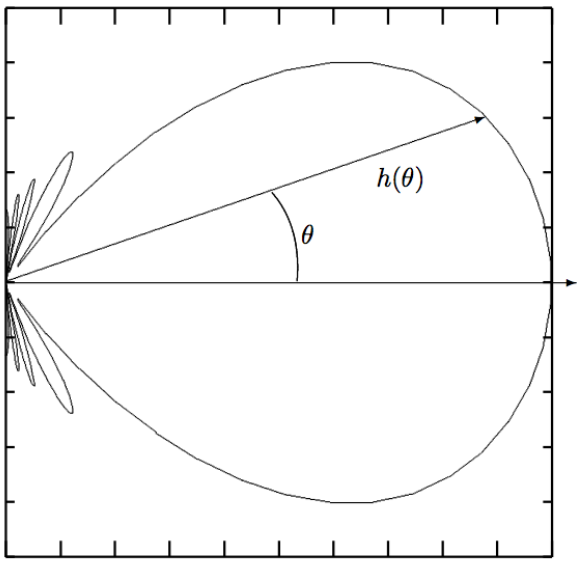


FIGURE 5.10 – Diagramme de rayonnement d'un HP de 38cm de diamètre, à 4 kHz.

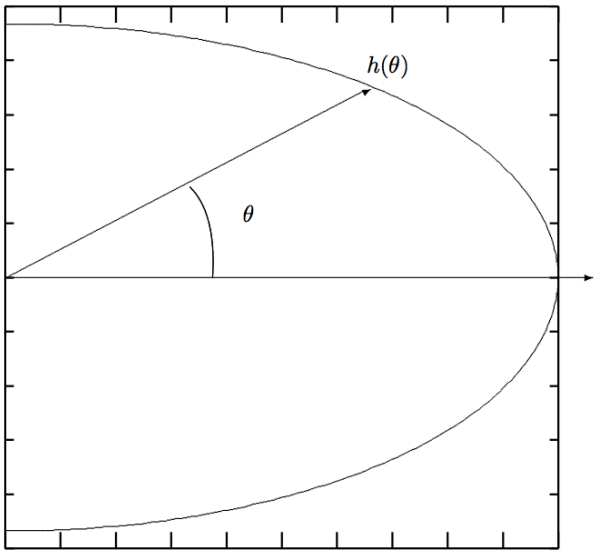


FIGURE 5.11 – Diagramme de rayonnement d'un HP de 38cm de diamètre, à 200 Hz. Le rayonnement est quasiment isotrope.

Chapitre 6

Les microphones

6.1 Le microphone électrodynamique

La figure 6.1 montre un schéma de principe du microphone électrodynamique à bobine mobile. Il s'agit exactement d'un haut-parleur utilisé en mode réciproque, et par conséquent, les équations seront rigoureusement les mêmes. Le déplacement de la bobine (solidaire de la membrane du microphone) dans le champ magnétique crée par l'aimant permanent engendre une force électromotrice induite $e = Blv$ et la circulation d'un courant dans le bobinage est à l'origine d'une force de Laplace $F = Bli$. L'équation mécanique s'écrit donc :

$$pS = m\frac{d^2x}{dt^2} + f'\frac{dx}{dt} + Bli + kx$$

et l'équation électrique

$$Blv = Ri + L\frac{di}{dt} + \frac{1}{C}\int_0^t i(\tau)d\tau$$

En régime harmonique, on pourra encore écrire :

$$\begin{aligned} pS &= Z_m v + Bli \\ Blv &= Z_e i \end{aligned}$$

avec les même expressions des impédances que dans le cas du haut-parleur. On peut directement en déduire l'impédance mécanique de la membrane :

$$Z_m^* = \frac{pS}{v} = Z_m + \frac{B^2l^2}{Z_e}$$

qui traduit, à pression acoustique donnée (de l'onde incidente), la vigueur avec laquelle la membrane répondra. Comme dans le cas du haut-parleur, on note que le second terme de cette impédance est lié à la circulation d'un courant dans le bobinage, et résulte donc directement du **couplage électromécanique**. On peut aussi chercher un schéma électrique équivalent (cf. figure 4.6) pour l'ensemble du microphone, en posant toujours

$$Z_c = jBl$$

pour l'impédance de couplage, et en choisissant les correspondances

$$\begin{aligned} i_1 &\to v \\ i_2 &\to i/j \\ U_1 &\to pS \\ U_2 &\to 0 \end{aligned}$$

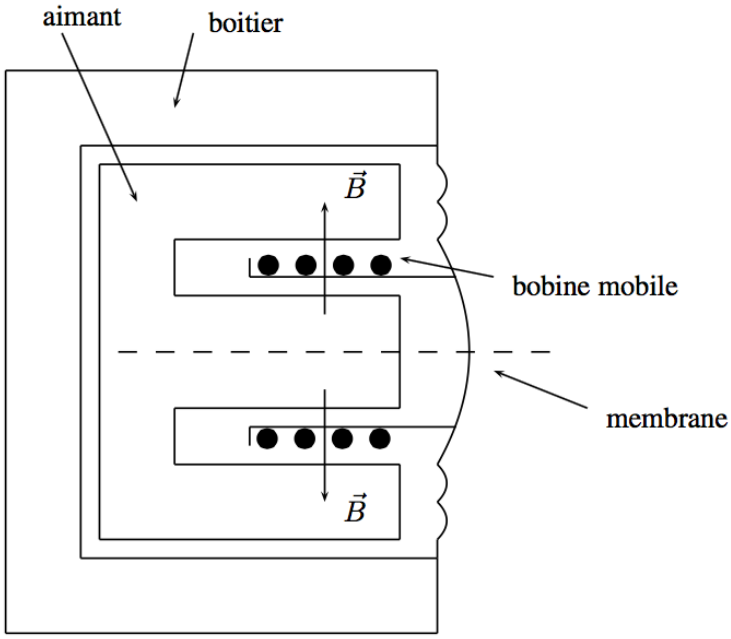


FIGURE 6.1 – Schéma de principe d'un microphone électrodynamique à bobine mobile.

On retrouve en particulier

$$Z_{11} = Z_m$$
$$Z_{22} = Z_e$$
$$Z_e^* = Z_m + \frac{B^2 l^2}{Z_e}$$

6.2 Le microphone électrostatique

Le principe du microphone électrostatique (ou microphone à condensateur) est basé sur la variation de capacité d'un condensateur lorsque la distance inter-armatures varie. Ce microphone offre généralement une réponse en fréquence bien plus plate que le microphone à bobine mobile, du fait du moindre poids de l'équipage mobile, qui se résume à la membrane. La figure 6.2 montre le schéma de principe et le dispositif de polarisation : la membrane constitue l'armature mobile du condensateur, qui doit être polarisé[1] puisque c'est la modulation des charges stockées qui génère le courant de sortie du microphone. La capacité du condensateur à armatures est donnée par la relation

$$C(x) = \frac{\epsilon S}{d - x}$$

où S est la surface d'une armature, et $d - x$ la distance inter-armature (qui varie donc à l'image de la membrane). On note

$$C_0 = \frac{\epsilon S}{d}$$

1. Le microphone à **électret** utilise un matériau ferroélectrique et est donc prépolarisé.

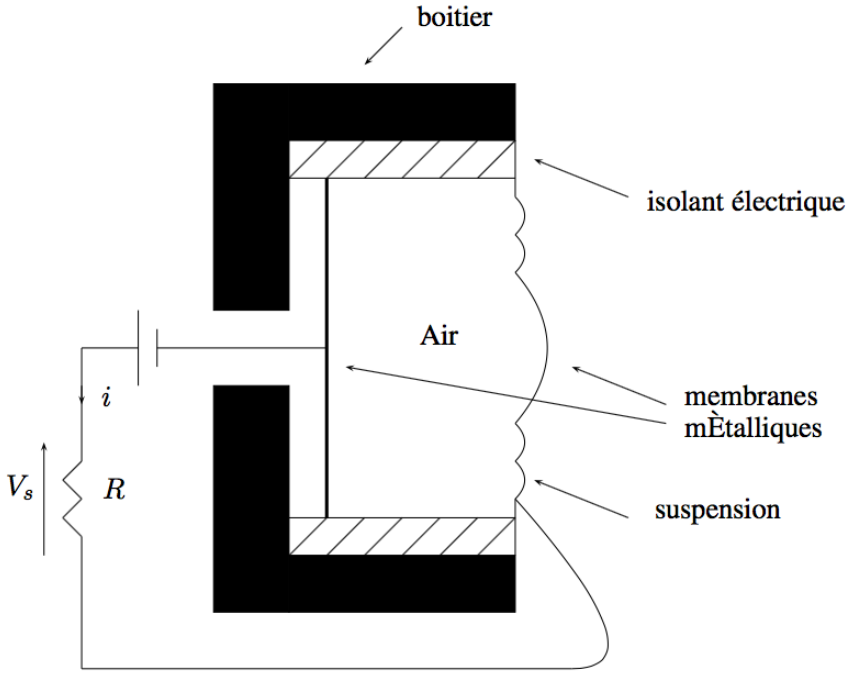


FIGURE 6.2 – Schéma de principe d'un microphone électrostatique et son dispositif de polarisation.

la capacité du condensateur, membrane au repos. On a donc, au premier ordre en x,

$$C(x) = \frac{C_0}{1 - \frac{x}{d}} \sim C_0 \left(1 + \frac{x}{d}\right)$$

En première analyse, la charge stockée vaut sensiblement $Q = C(x)V_0$, donc toute variation de capacité engendre un courant proportionnel à la vitesse de la membrane. Pour une modélisation plus précise, nous devons tenir compte de la présence de la résistance R (sans laquelle on ne peut d'ailleurs pas récupérer de tension variable au borne du condensateur...). Avec les notations de la figure 6.2, la charge stockée au repos vaut :

$$Q_0 = C_0 V_0$$

et l'on sépare Q en sa composante continue Q_0 et sa composante variable q :

$$Q(t) = Q_0 + q(t)$$

La loi des mailles s'écrit, au premier ordre en x (donc en q) :

$$\begin{aligned} V_0 &= Ri + \frac{Q}{C} \\ &= Ri + \frac{Q_0 + q}{C_0}\left(1 - \frac{x}{d}\right) \\ &\sim Ri + V_0 + \frac{q}{C_0} - \frac{V_0 x}{d} \end{aligned}$$

et se simplifie en

$$Ri + \frac{q}{C_0} = \frac{V_0 x}{d}$$

En régime harmonique, on a alors

$$Z_e i = (R + \frac{1}{jC_0\omega})i = \frac{\mathcal{E}_0}{j\omega}v$$

en notant $\mathcal{E}_0$ le champ électrique régnant dans le condensateur, au repos. Enfin, si l'on s'intéresse plutôt à la tension de sortie $V_s = Ri$, on a, toujours en régime harmonique,

$$\frac{V_s}{v} = \frac{RC_0\mathcal{E}_0}{1 + RC_0 j\omega}$$

On note que la sensibilité du microphone est proportionnelle à

$$RC_0\mathcal{E}_0 = \frac{R\epsilon S\mathcal{E}_0}{d}$$

et la coupure haute à lieu à la pulsation

$$\omega_c = \frac{1}{RC_0} = \frac{d}{R\epsilon S}$$

On en déduit le produit gain-bande,

$$G \times BP = \mathcal{E}_0$$

qui montrer tout l'intérêt de polariser le microphone sous une forte tension. Déterminons maintenant l'équation mécanique. Rappelons tout d'abord que la force électrostatique d'attraction entre les électrodes d'un condensateur plan vaut[2]

$$F = \frac{Q^2}{2\epsilon S}$$

ce qui donne, au premier ordre en q,

$$F \sim \frac{Q_0^2}{2\epsilon S} + \frac{Q_0}{\epsilon S}q$$

Le premier terme est un terme statique : il représente simplement la force statique qui s'exerce entre armatures du fait de la polarisation sous V_0, et cette force est bien entendu équilibrée par la force de rappel résultant de la compression du diélectrique. Reste la composante alternative,

$$F = \mathcal{E}_0 q$$

C'est un terme positif qui vient donc **renforcer** l'effet de la pression due à l'onde incidente[3]. L'équation mécanique s'écrit finalement

$$m\frac{d^2x}{dt^2} = pS + \mathcal{E}_0 q - f'\frac{dx}{dt} - kx$$

où k est la raideur mécanique du diélectrique (liée à son coefficient de compressibilité isentropique). En régime harmonique, nous obtenons donc au final deux équations couplées :

$$Z_m v = pS + \frac{\mathcal{E}_0}{j\omega}i$$
$$Z_e i = \frac{\mathcal{E}_0}{j\omega}v$$

2. Cette expression s'obtient par exemple en prenant le gradient de l'énergie potentielle $E_p = 1/2CU^2$.
3. J'insiste : dans le cas du micro à bobine mobile, la force de Laplace joue un rôle de contre-réaction, et augmente donc l'amortissement. Ici c'est l'inverse, et la réaction est positive.

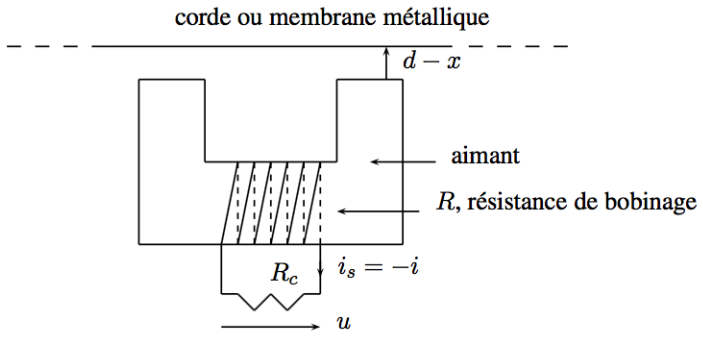


FIGURE 6.3 – Schéma de principe d'un microphone à réluctance variable pour guitare électrique.

On en déduit par exemple l'impédance mécanique libre de la membrane,

$$Z_m^* = \frac{pS}{v} = Z_m - \frac{Z_c^2}{Z_e}$$

en posant pour impédance de couplage,

$$Z_c = \frac{\mathcal{E}_0}{j\omega}$$

On voit que, contrairement au microphone à bobine mobile, le couplage est à raideur prépondérante, et diminue donc avec la fréquence.

6.3 Les microphones pour guitare électrique (à réluctance variable)

Basé sur le principe du couplage électromagnétique, ce microphone possède deux utilisations essentielles, dans les écouteurs téléphoniques d'une part, pour son faible encombrement et son coût peu élevé, dans les guitares électriques d'autre part, parce qu'il utilise alors la corde métallique comme composant du circuit magnétique (ce serait naturellement impossible avec une corde en nylon...). Le schéma de principe est illustré figure 6.3. La vibration de la corde (ou de la membrane métallique dans le cas de l'écouteur) provoque une variation de largeur d'entrefer : la variation de flux qui en résulte crée une f.e.m. induite, que l'on récupère sur le bobinage de sortie. Par construction, ces microphones ont donc, comme les micros à bobine mobile, une impédance de sortie élevée due au nombre élevé de spire nécessaire à l'obtention d'une sensibilité acceptable. En supposant infinie la perméabilité des armatures du circuit magnétique, l'application du théorème d'Ampère conduit à

$$\int \vec{H}.\vec{dl} \sim 2H_0(d-x) = 2\mathcal{R}_0\Phi_0\left(1-\frac{x}{d}\right)$$

où :

— $\Phi_0 = \mu_0 H_0 S$ est la composante continue du flux dans le circuit magnétique (et notamment dans l'entrefer)

— $\mathcal{R}_0 = \frac{d}{\mu_0 S}$ est la **réluctance** d'entrefer au repos ($x = 0$)

On pose, comme pour le microphone électrostatique,

$$\Phi(t) = \Phi_0 + \phi(t)$$

et

$$I(t) = I_0 + i(t)$$

où I_0 correspond, soit au courant magnétisant circulant dans la bobine (destiné à créer le champ), soit au courant équivalent dans le cas d'un aimant permanent. L'inductance de bobinage s'écrit :

$$L = \frac{\Phi}{I} = \frac{N\mu_0 HS}{I} = \frac{N^2\mu_0 S}{d-x} = \frac{L_0}{1-x/d}$$

On procède désormais comme pour le microphone électrostatique, en développant toutes les expressions à l'ordre 1 en x (donc en i) :

$$I = I_0 + i = \frac{\Phi}{L} \sim I_0 + \frac{\phi}{L_0} - \frac{I_0 x}{d}$$

qui se simplifie en

$$i \sim \frac{\phi}{L_0} - \frac{I_0 x}{d}$$

La loi des mailles s'écrit :

$$u = e + Ri = \frac{d\phi}{dt} + Ri$$

En régime harmonique, on obtient finalement :

$$u - Z_e i = \frac{\Phi_0}{d} v = -(R + R_c + jL_0\omega)i$$

où R_c est la résistance de charge (par ex., la résistance d'entrée de l'amplificateur). L'équation mécanique fait intervenir la force magnétique d'entrefer (analogue à la force électrostatique inter-armature), obtenue en prenant le gradient de l'énergie potentielle magnétique, $E_p = 1/2Li^2$:

$$F = \frac{\Phi^2}{2L_0 d} \sim \frac{\Phi_0^2}{2L_0 d} + \frac{\Phi_0}{L_0 d}\phi$$

On reconnaît là encore un terme statique, compensé par la (légère) compression des armatures[4], et un terme variable, qui se réécrit

$$F = \frac{\Phi_0}{L_0 d}\left(L_0 i + \frac{L_0 I_0 x}{d}\right) = \frac{\Phi_0}{d} i + \frac{\Phi_0^2}{L_0 d^2} x$$

Au final, l'équation mécanique prend donc la forme suivante :

$$F_{ex} + \frac{\Phi_0}{d} i + \frac{\Phi_0^2}{L_0 d^2}\frac{v}{j\omega} = Z_m v$$

où F_{ex} représente, soit la pression acoustique de l'onde incidente (pS s'exerçant sur la membrane de l'écouteur), soit la force d'excitation exercée sur la corde au moment du pincement (mais si l'on veut prévoir l'apparition de **Larsen**, on doit aussi prendre en compte la

4. C'est là l'origine des vibrations des transformateurs à 50Hz

mise en vibration de la corde par l'onde provenant du haut-parleur de l'ampli). L'impédance mécanique de la membrane (ou de la corde métallique) s'écrit

$$Z_m = f' + jm\omega + \frac{K_m}{j\omega}$$

en négligeant les phénomènes des propagation sur la corde, puisque dans le cas d'une guitare électrique, une infime partie de l'énergie est transférée de la corde vers le corps de la guitare. En posant $K_{mag} = -\frac{\Phi_0^2}{L_0 d^2}$, on a finalement

$$F_{ex} + \frac{\Phi_0}{d} i = \left(Z_m + \frac{K_{mag}}{j\omega} \right) v$$

On remarque que, conformément à la loi du flux maximum, la raideur magnétique est négative, *i.e.* elle renforce l'effet de la force excitatrice, comme pour le microphone électrostatique. Déterminons pour conclure l'impédance mécanique **libre** de la corde :

$$\begin{aligned} Z_m^* &= \frac{F_{ex}}{v} \\ &= Z_m + \frac{K_{mag}}{j\omega} + \frac{(\Phi_0/d)^2}{R + R_c + jL_0\omega} \end{aligned}$$

On peut donc choisir pour impédance de couplage,

$$Z_c = \frac{j\Phi_0}{d}$$

Cette expression montre que le couplage dépend très logiquement :

— du flux crée par l'aimant permanent,

— de la distance entre la corde et l'aimant (d'où la présence de vis permettant d'approcher plus ou moins le micro de la corde afin d'équilibrer le volume entre les 6 cordes).

Chapitre 7

Éléments de traitement du signal musical (H. Papadopoulos)

7.1 Introduction

Depuis quelques années, on assiste à l'apparition d'un nombre gigantesque et toujours croissant de collections de musique en ligne. Grâce au développement des techniques de compression (tels que les formats MP3, WMA, AAC ou OGG), chacun peut avoir à portée de main sa collection personnelle composée de milliers de morceaux de musique. Ce phénomène a attiré l'attention de nombreux scientifiques ces dernières années. En effet, il est devenu urgent de développer des outils et des méthodes qui permettent d'interagir avec ces énormes bibliothèques de musique numériques. Professionnels ou amateurs souhaitent maintenant avoir accès à des outils qui leur permettent d'écouter ou d'apprendre la musique de façon personnalisée. Le domaine de la recherche d'information musicale (Music Information Retrieval, MIR) est ainsi devenu très actif depuis une dizaine d'années. L'indexation musicale est un aspect très important de ce domaine. C'est une discipline qui a pour but d'aider au stockage, à la diffusion et à la consultation des gigantesques collections de musique en ligne.

Ne vous est-il jamais arrivé de vous réveiller avec un air de musique en tête et d'être agacé de ne pas savoir d'où il venait ? Et si vous pouviez chantonner cet air à un ordinateur, et que celui-ci vous suggérait que c'est le groupe "Pink Noise Party" qui l'a composé et vous en proposait plusieurs versions trouvées sur le Web ? Imaginez qu'en plus il vous propose également d'autres chansons similaires qui pourraient également vous intéresser. Votre irritation disparaîtrait probablement bien vite. Un tel système n'existe pas encore aujourd'hui. Mais les récents développements en Music Information Retrieval (MIR, en français, Recherche d'information en musique) montrent qu'un tel système sera disponible dans les années à venir. Cependant, il reste un grand nombre de difficultés à surmonter.

Une oeuvre musicale peut être caractérisée par un certain nombre d'attributs musicaux comme la mélodie, la progression des accords, l'instrumentation ou le tempo. Un des aspects les plus importants du domaine MIR est l'extraction et le traitement de descripteurs spécifiques au signal audio de musique. Ce problème est lié au problème plus général que constitue la transcription de musique. L'extraction automatique de descripteurs musicaux a de nombreuses applications directes telles que la classification des morceaux de musique, l'identification d'artistes, la classification par humeur ou par genre, la segmentation structurelle, ou, de manière générale, toutes les applications basées sur l'indexation par le contenu.

L'annotation manuelle du contenu de pièces musicales est une tâche difficile et extrêmement coûteuse en temps. Il est donc essentiel de développer des techniques qui permettent d'extraire de manière automatique des informations sur le contenu musical des signaux audio. Bien que ce soit une tâche relativement aisée pour un musicien un peu entraîné, les travaux de recherche effectués jusqu'à présent ont montré que l'extraction automatique d'information de contenu musical s'avère être une tâche extrêmement complexe.

Les algorithmes qui permettent d'extraire des informations sur le contenu musical d'un signal audio commencent par une phase d'analyse du signal. Le traitement d'un signal audio de musique nécessite de développer des outils qui lui sont adaptés. On va par exemple s'intéresser au contenu harmonique d'un morceau de musique, c'est-à-dire qu'on va chercher à en extraire les notes, les accords, éventuellement la mélodie. L'analyse du contenu harmonique d'une partition de musique est un problème complexe, mais cette tâche est encore plus compliquée lorsque l'on travaille directement sur le signal audio. En effet, en plus des difficultés inhérentes à la syntaxe musicale (la grammaire, la langue), la première difficulté est d'obtenir des informations sur la hauteur des notes qui sont présentes dans le signal audio.

Le but de ce chapitre est de présenter deux représentations qui sont spécifiquement utilisées pour l'analyse des signaux audio de musique. Des sujets de mini-projet permettront d'approfondir cette étude.

7.2 Pré-requis musicaux à l'attention des scientifiques non-musiciens

Lorsque l'on travaille sur des signaux audio de musique, on doit se référer à des notions et des termes empruntés au langage musical théorique. C'est pourquoi nous commencerons par aborder quelques notions générales de la théorie de la musique.

7.2.1 Notes et intervalles

Notes

Avant de poursuivre, nous rappelons quelques notions sur les notes, qui sont les signes employés pour écrire la musique. Les notes représentent des durées et des hauteurs de son. Les méthodes de division de l'octave en intervalles ont depuis très longtemps donné naissance à la gamme dite heptatonique, c'est-à-dire comprenant sept notes dans un intervalle d'octave. Ces notes ont été nommées, dans le sens ascendant :

français	do	re	mi	fa	sol	la	si	do
anglo-saxon	C	D	E	F	G	A	B	C

Intervalles

L'oreille identifie des intervalles. Un intervalle est une grandeur additive que nous percevons comme une différence de hauteur, quand la physique identifie des rapports de fréquences.

On appelle *octave* l'intervalle entre deux sons dont l'un est à la fréquence f et l'autre à la fréquence $2f$. L'octave correspond aussi à l'intervalle qui sépare la fréquence fondamentale de la première harmonique.

7.2.2 Son fondamental et harmoniques

Une vibration sonore est une fonction mathématique périodique. Elle peut donc être décomposée en série de Fourier, c'est-à-dire en une somme de fonctions sinusoïdales élémentaires. Les sons musicaux sont formés d'un son fondamental et des harmoniques appelées aussi partiels, dont les rapports de fréquence avec la fondamentale sont des quotients de nombres entiers[1]. La hauteur d'un son est mesurée par la fréquence du fondamental.

Lorsque l'on entend un Do, on entend aussi la première harmonique qui est le Do de l'octave supérieure. Une note de la gamme est ainsi déterminée modulo la multiplication par une puissance de 2 qui détermine l'octave où elle se trouve.

Par exemple l'échelle des A (=La), en Hz est la suivante :

Fréquence (Hz)	55	110	220	440	880	1760
note	$A1$	$A2$	$A3$	$A4$	$A5$	$A6$

Lorsque l'on entend un C (=Do) de fréquence f, on entend aussi les harmoniques de fréquences $2f$, $3f$ ... :

Fréquence (Hz)	f	$2f$	$3f$	$4f$	$5f$	$6f$
note	C	C	G	C	E	G

7.2.3 Gamme et tonalité

Une gamme est une série de sons conjoints. La gamme tempérée est de nos jours utilisée de façon presque universelle dans la musique occidentale. Elle est obtenue en divisant l'octave en douze intervalles égaux.

note	C	D	E	F	G	A	B	C
fréquence	f	a^2f	a^4f	a^5f	a^7f	a^9f	$a^{11}f$	$2f$

avec $a = 2^{1/12}$.

On appelle *demi-ton* l'intervalle défini par $(f, 2^{1/12}f)$. Un *ton* peut se diviser en 2 demi-tons.

Pour former une gamme, on utilise sept notes de noms différents. Chacune ayant un rôle déterminé, on lui donne le nom de degré que l'on écrit en chiffres romains. Chaque degré a un nom particulier qui caractérise la position qu'il occupe dans la gamme.

On utilise les termes *tonique*, qui correspond au premier degré, ainsi que les termes *médiante*, *sous-dominante*, *dominante* et *octave* qui correspondent respectivement aux $3^{\text{ème}}$, $4^{\text{ème}}$, $5^{\text{ème}}$ et $8^{\text{ème}}$ degrés.

7.2.4 Classes de hauteur ou pitch class

La perception de la hauteur d'un son par l'oreille humaine est fondée sur la périodicité du signal sonore. Les hauteurs séparées par un nombre entier d'octaves sont perçues comme "sonnant" de manière équivalente, leur consonance est maximale et c'est pourquoi on leur donne le même nom. On dit qu'elles partagent le même *chroma*. L'ensemble des notes partageant un même chroma est appelé *classe de hauteurs* ou *pitch class*.

1. En réalité les partiels au-dessus de la fréquence fondamentale ne sont pas harmoniques parfait dans de nombreux instruments, comme le piano ou la guitare. Chaque partiel est en fait un peu faux, ou inharmonique

Les théoriciens de la musique se réfèrent en général aux différentes classes de hauteur en utilisant des nombres. On peut transformer la fréquence fondamentale f d'un son en un nombre réel p selon l'équation suivante (conversion en échelle midi) :

$$p = 69 + 12log_2\frac{f}{440} \tag{7.1}$$

Le Do4 correspond à la note midi 60[2]

Dans l'espace des classes de hauteur, il n'y a pas de distinction entre les notes qui sont séparées par un nombre entier d'octaves (p, $p + 12$, $p + 2 * 12$...). Par exemple Do4, Do5, Do6 appartiennent à la même classe de hauteur.

7.3 Comparaison de représentation d'un signal acoustique

Les algorithmes pour l'analyse automatique des signaux audio de la musique reposent en général sur une représentation spectrale du signal. La transformée de Fourier est la représentation la plus couramment utilisée. Bien qu'elle soit très populaire, un désavantage majeur de cette représentation est que les composantes fréquentielles sont espacées de manière égale et ont donc une résolution constante. Il faut donc faire un compromis résolution temporelle et résolution fréquentielle. Les approches multi-résolution ont été proposées comme alternative à la transformée de Fourier.

7.3.1 Transformée de Fourier

Depuis son introduction au 18e siècle, la transformée de Fourier et ses extensions sont devenues les représentations du signal les plus couramment utilisées en traitement du signal.

En pratique, nous traitons les signaux de mesure finie. La transformée de Fourier discrète (DFT) calcule un spectre de fréquences d'un signal $x(n)$ à temps discret de longueur finie N :

$$X(k) = \sum_{n=0}^{N-1} x(n)e^{-j2\pi kn/N} \tag{7.2}$$

On obtient ainsi une représentation spectrale discrète du signal échantillonné $x(n)$.

Le coût de calcul d'une DFT peut être très élevé. Un algorithme beaucoup plus rapide a été développé par Cooley et Tukey en 1965, appelée Fast Fourier Transform (FFT) et est utilisé en général dans le traitement du signal.

En pratique, en traitement du signal, une fenêtre d'analyse ou de pondération $w(n)$ est appliquée aux données. La transformée de Fourier à court terme (Short Time Fourier Transform, STFT) représente le contenu fréquentiel d'un court segment (d'une durée limitée) du signal. Ce segment d'une durée limitée est supposé être stationnaire. La STFT d'un signal discret $x(n)$ peut être calculée selon :

2. La norme MIDI, née dans les années 70, établit un ensemble de protocoles de communication entre instruments de musique électroniques. Des messages sont échangés entre instruments (par exemple entre Live Ableton ou Cubase et un synthétiseur externe) à chaque fois qu'une note doit être jouée. Dans ces messages MIDI, les notes sont encodées sous forme d'un nombre de 7 bits.

$$X(k) = \sum_{n=0}^{N-1} w(n)x(n)e^{-j2\pi kn/N} \tag{7.3}$$

La longueur de la fenêtre N détermine la résolution en temps et en fréquence. La précision dans le domaine fréquentiel augmente avec la longueur de la fenêtre. Toutefois, cela se produit au détriment de la résolution temporelle.

Le spectrogramme est un outil de visualisation du son standard qui montre la répartition de l'énergie en temps et en fréquence. Il s'agit d'une image obtenue en utilisant la transformée de Fourier à court-terme (STFT). Le temps est indiqué horizontalement et la fréquence est indiquée verticalement. Ce type de représentation est utilisé très fréquement par la plupart des algorithmes de traitement des signaux audio. On peut accéder à un grand nombre d'indications sur le son à partir du spectrogramme. Par exemple on peut repérer les sons percussifs qui sont indiqués par des traces d'énergie assez brèves dans le domaine temporel. On peut aussi distinguer les différents partiels d'un son émis par un instrument. La figure 7.1 montre le spectrogramme d'un extrait de *I Me Mine* des Beatles.

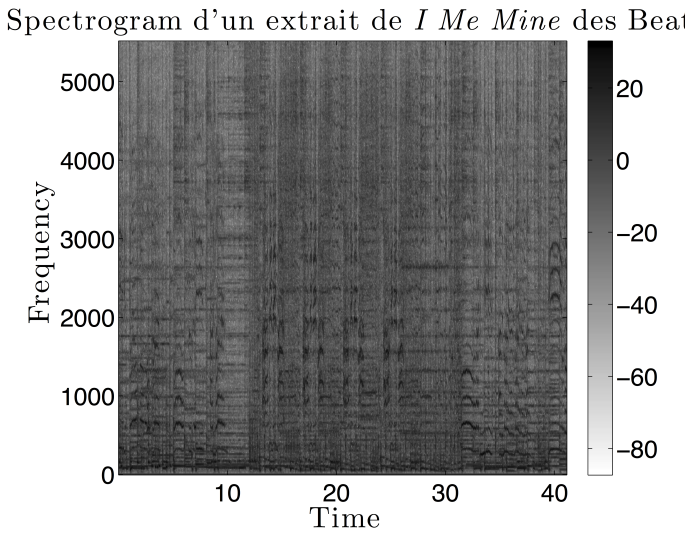


FIGURE 7.1 – Spectrogramme d'un extrait de *I Me Mine* des Beatles.

7.3.2 Résolution fréquentielle versus résolution temporelle

Lors de l'analyse des signaux de musique, le choix de la longueur de la fenêtre d'analyse est un élément clé. Il détermine le compromis entre résolution fréquentielle et résolution temporelle, qui affecte la régularité du spectre et la détectabilité des composantes sinusoïdales. D'une part, une bonne résolution temporelle, et donc une longueur de fenêtre courte, est nécessaire afin de détecter des changements rapides dans le signal (par exemple les attaques des notes). D'autre part, il faut que la longueur de la fenêtre d'analyse soit suffisamment grande afin que l'on puisse distinguer les sinusoïdes rapprochées correspondant à des hauteurs voisines.

Considérons un signal audio de musique avec deux sinusoïdes de fréquences f_1 et f_2 qui correspondent à deux notes adjacentes. Notons $\Delta f = f_2 - f_1$. Dans un signal de musique,

lorsque deux sinusoïdes correspondant à deux notes adjacentes ont des fréquences voisines séparées par Δf Hz, il est nécessaire que la longueur de la fenêtre N soit assez grande pour que l'on puisse distinguer les deux pics dans le spectre (voir figure 7.2).

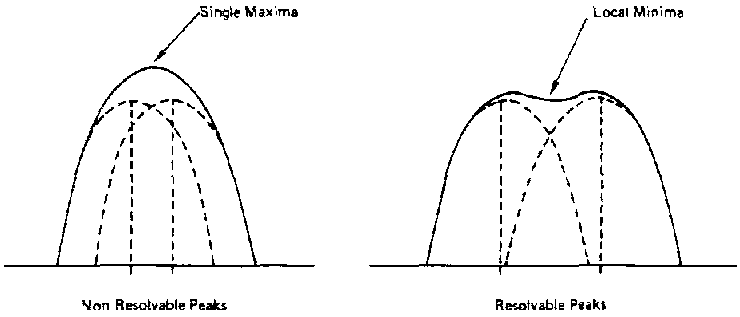


FIGURE 7.2 – Résolution spectrale de pics proches. à gauche : pics non resolvables. à droite : pics resolvables. Tiré de l'article Harris, *On the use of windows for harmonic analysis with the discrete Fourier transform*, IEEE 1978.

Comme les fréquences de la gamme tempérée de la musique occidentale suivent une échelle logarithmique, les fréquences correspondant à deux notes adjacentes sont d'autant plus proches que l'on se rapproche des basses fréquences. Deux notes adjacentes correspondant à un demi-ton sont séparés par 6 % de la fréquence de la note la plus basse. En effet, si f_k et f_{k+1} désignent respectivement la fréquences de la note la plus basse et la plus haute, par construction de l'échelle de la musique tonale occidentale, nous avons :

$$\frac{f_{k+1}}{f_k} = 2^{\frac{1}{12}} soit\ f_{k+1} \approx 1.059 f_k \tag{7.4}$$

Par exemple, à un Do1 correspond une fréquence $f_{Do1} = 32.7$Hz. La fréquence de la note suivante Do#1 est : $f_{Do\#1} = 32.7 + 0.06 * 32.7 = 34.6$Hz.

Dans la musique tonale occidentale, il est peu probable que deux notes adjacentes de basse fréquence soient jouées simultanément, parce que c'est généralement désagréable pour les oreilles. Toutefois, des notes chromatiques qui sont jouées successivement peuvent être superposées pendant une certaine durée. C'est par exemple le cas dans la plupart de la musique pour piano de Chopin, où l'utilisation intensive de la pédale résulte dans le mélange de notes adjacentes dans les basses fréquences. Lorsque l'on analyse de tels signaux, il faut donc tenir compte de la contrainte liée à la taille de fenêtre d'analyse minimum nécessaire pour avoir une résolution fréquentielle satisfaisante.

Les approches multi-résolution ont été proposées comme alternative à la fréquence linéaire classique et constante résolution de la TFD. Dans la suite, on va voir une approche multi-résolution couramment utilisée dans l'analyse de la musique audio : la transformée Constant-Q (Constant-Q Transform, CQT).

7.3.3 Constant-Q Transform

Une approche commune pour résoudre le dilemme résolution temps/fréquence est de réaliser une analyse fréquentielle multi-résolution. Dans le cas d'une analyse à résolution fixe (transformée de Fourier discrète par exemple), les fréquences sont espacées de manière égale

et on a donc une résolution fréquentielle constante. Dans le cas d'une analyse fréquentielle multi-résolution, le spectre de fréquences est divisé en sous-bandes et chacune est traitée indépendamment des autres. Ceci permet l'utilisation de fenêtres d'analyse plus courtes aux fréquences les plus élevées tandis que dans les basses fréquences on peut encore avoir la résolution en fréquence nécessaire pour séparer des sinusoïdes rapprochées. Une approche intéressante a été présentée en 1991 par Brown qui a proposé d'utiliser la *constant-Q transform* pour l'analyse des signaux de musique. La constant-Q transform est une représentation spectrale où les canaux de domaine de fréquence ne sont pas linéairement espacés, comme dans le cas de la transformée de Fourier, mais géométriquement espacés (le rapport entre fréquence centrale et résolution $Q = \frac{f}{\Delta f}$ demeure constant).

La CQT est étroitement liée à la transformée de Fourier, mais donne une meilleure représentation des données spectrales d'un signal de musique. Les fréquences centrales qui sont distribuées géométriquement suivent celles des notes de la gamme tempérée utilisée dans la musique occidentale. Notez que la CQT a été introduite plus tôt, en dehors du contexte musical.

La constant-Q transform est très proche de la transformée de Fourier. Cependant, contrairement à celle-ci, ses composantes fréquentielles sont séparées de manière géométrique. Elle est donc plus adaptée aux signaux de musique. En effet, une note de musique jouée par un instrument produit un son fondamental ainsi que plusieurs harmoniques dont la fréquence est multiple de la fondamentale. Une propriété qui en découle est que la position des harmoniques les unes par rapport aux autres est indépendante de la fréquence fondamentale si elles sont tracées en log-fréquence. Ceci est illustré sur la Figure 7.3 où on représente le spectre hypothétique d'un son de fréquence fondamentale f. La distance entre les deux premières harmoniques est $log(2)$, entre la seconde et la troisième est $log(\frac{3}{2})$ etc. La position absolue des harmoniques sur l'échelle des fréquences dépend de la fréquence fondamentale mais la position relative des différentes harmoniques d'une note est constante et forme un "pattern".

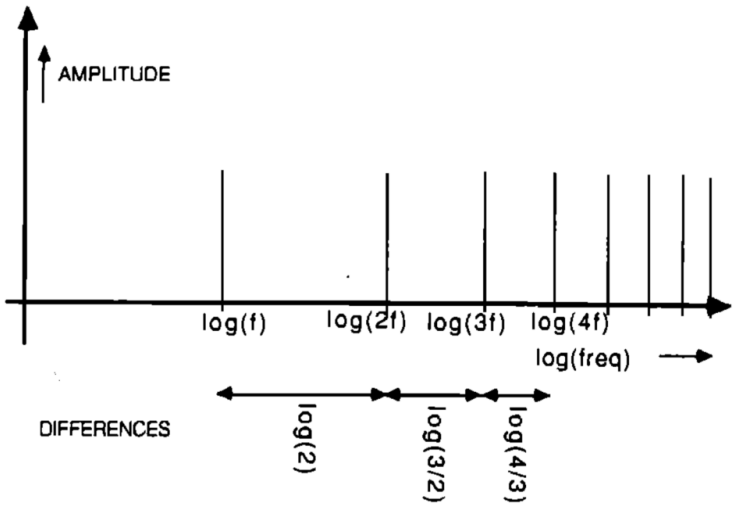


FIGURE 7.3 – Pattern de la transformée de Fourier d'une note de musique tracé selon une échelle des fréquences logarithmique. Adapté de l'article Brown, *Calculation of a constant Q spectral transform*, JASA 1991.

Une autre propriété de la constant-Q transform est que la résolution temporelle augmente avec la fréquence. Ainsi, une grande taille de fenêtre d'analyse est utilisée dans les basses

fréquences et, plus la fréquence augmente, plus la taille de fenêtre diminue.

Lorsqu'il est appliqué aux signaux de musique, le calcul de la CQT est basé sur des fréquences qui correspondent aux notes de la gamme tempérée occidentale. La constant-Q transform d'un signal temporel échantillonné $x(n)$ peut être directement calculée par :

$$X^{cq}(k) = \sum_{n=0}^{N(k)-1} w(n,k)x(n)e^{-j2\pi f_k n} \tag{7.5}$$

où $X^{cq}(k)$ est la k^{eme} composante de la constant-Q transform. Pour chaque valeur de k, la fenêtre d'analyse $w(n,k)$ varie proportionnellement à la fréquence centrale f_k. On note Q le rapport constant fréquence à résolution : $Q = \frac{f_k}{\Delta f_k}$, et f_s le taux d'échantillonnage. La longueur en échantillons de la fenêtre $w(n,k)$ à la fréquence f_k est $N(k) = \frac{Q.f_s}{f_k}$. $N(k)$ dépend de la fréquence et donc de la position du "bin" k.

La figure 7.4 représente la taille de fenêtre N (en secondes) par rapport à la fréquence (en Hertz), pour un espacement de $\frac{1}{2}$-ton ($Q = (2^{\frac{1}{12}} - 1)$). Par exemple, à une fréquence de 104Hz correspond une fenêtre de $0,5$s. On peut voir que la résolution temporelle augmente avec la fréquence : plus la fréquence est élevée, plus la taille de fenêtre est petite.

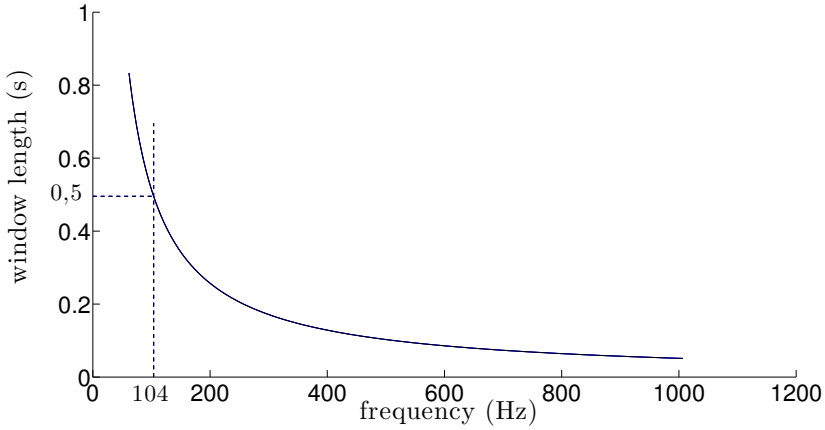


FIGURE 7.4 – Longueur de la fenêtre de la constant-Q transform en secondes par rapport à la fréquence en Hertz.

7.4 Chromagram

Les chromas sont une représentation intéressante et puissante des signaux audio de musique. On peut voir cette représentation comme une représentation intermédiaire entre une représentation "signal" telle que la transformée de Fourier et une représentation symbolique telle que la partition. Cette représentation du signal permet d'en extraire de manière satisfaisante des informations sur le contenu harmonique et mélodique.

7.4.1 Definition

La notion de chroma a été introduite par le psychologiste Roger Shepard dans les années 1960. Les chromas transforment les fréquences en classes d'équivalence d'octaves. Shepard a montré que deux dimensions sont nécessaires pour bien représenter la perception du système auditif humain. Celle-ci peut être reprensentée par une hélice (voir figure 7.5). Cette représentation de la hauteur par une courbe hélicoïdale est assez ancienne puisqu'elle avait déjà été proposée par Drobisch en 1846.

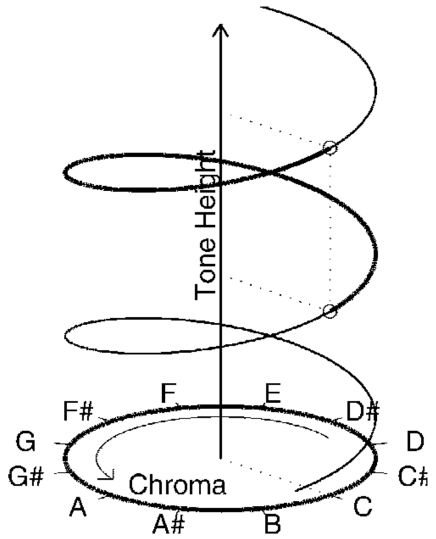


FIGURE 7.5 – Représentation de la perception de la hauteur par l'oreille humaine. B_{n+1} est une octave au-dessus de B_n. Extrait de Harte & Sandler, *Automatic chord identification using a quantised chromagram*, AES 2005.

La hauteur d'une note (pitch (p)) en Hz peut être décrite par deux valeurs : le chroma (ou pitch class (c)) et la hauteur de ton (ou pitch heigh (h)).

$$p = 2^{c+h} \tag{7.6}$$

Les chroma sont traditionnellement des vecteurs à 12 dimensions, chaque dimension correspondant à l'intensité associée à l'une des 12 classes de hauteur (chroma) de la gamme chromatique de la musique tonale occidentale, quelle que soit l'octave. Le chromagram (ou spectre de chroma) est une extension de la notion de chroma qui inclut la dimension temporelle. Il peut être utilisé pour représenter les propriétés de la distribution du spectre d'énergie du signal à travers les fréquences et le temps. Il s'agit d'une représentation compacte de la représentation spectrale (FFT ou CQT) du signal audio.

7.4.2 Calcul

Le chromagram s'obtient en effectuant un mapping entre une représentation spectrale et un vecteur à 12 dimensions représentant l'intensité des 12 demi-tons de la gamme chromatique. La représentation sous forme de chromas est très utilisée dans les travaux relatifs

à l'estimation automatique de la tonalité, de l'harmonie, des accords d'un morceau de musique, mais aussi pour la détection de la structure musicale (lorsque l'on cherche à découper un morceau en intro, couplet, refrain etc.), de similarité musicale (lorsque l'on cherche à déterminer si deux morceaux se ressemblent ou non) etc. Le calcul du chromagram se fait de la manière suivante :

— Mapping des valeurs du spectre aux 12 demi-tons des pitch-classes. Pour chaque fréquence f_k on a :

$$c(f_k) = mod(12log_2(\frac{f_k}{261.62}), 12) \tag{7.7}$$

où $c(f_k)$ est la valeur associée à f_k sur l'échelle des chromas et 261.61 correspond à la fréquence d'un Do4.

— Calcul du vecteur de chroma à 12 dimensions en additionnant les intensités de la transformée de Fourier des fréquences de même valeur $c(f_k)$:

$$\text{Pour } l = 1, \ldots, 12 \quad C(l) = \sum_{f_k \text{telle que} \quad c(f_k)=l} A(f_k) \tag{7.8}$$

La figure 7.6 correspond à un extrait du chromagram obtenu à partir du morceau *I am a Loser* de l'album *Beatles For Sale* des Beatles. Les régions les plus sombres corrrespondent aux intensités les plus importantes. Nous pouvons distinguer ici la suite d'accords La majeur(La Do# Mi), La# majeur, Ré majeur(Ré F# A), Sol majeur(Sol Si Ré), Ré majeur, Sol majeur, Ré majeur, Sol majeur.

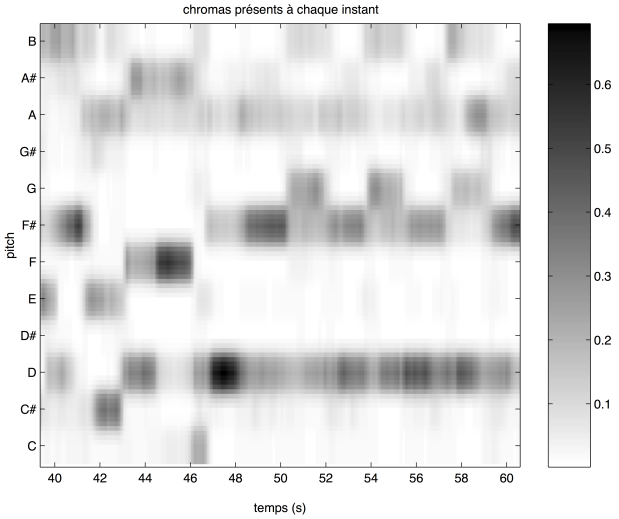


FIGURE 7.6 – Exemple de chromagram extrait de *I am a Loser* de l'album *Beatles For Sale*.

7.4.3 Un problème intéressant : chromas et harmoniques

Comme nous l'avons vu, lorsqu'un instrument émet une note il produit également un ensemble d'harmoniques. Dans une représentation spectrale, nous n'observons pas directement les notes que l'on entend, mais un mélange de leurs harmoniques, ce qui se traduira par un mélange de valeurs non nulles dans le vecteur de chroma. Il est donc important de noter

que le vecteur de chroma d'une note jouée par un instrument ne contient pas seulement une valeur non nulle au "bin" correspondant à la classe de hauteur correspondant à la fréquence fondamentale $f0$ de la hauteur percue (sans tenir compte des considérations d'octave), mais également un mélange de ses harmoniques

La Figure 7.7 montre le chroma d'un Do1 $(65,4$ Hz) joué par un violoncelle en fonction de divers intervalles de fréquences, en fonction de l'intervalle de fréquences considéré dans le spectre pour le calcul du chromagram (de $f_{min} = 60$ Hz à différentes valeurs de f_{max}). Nous pouvons suivre l'apparition des harmoniques de la note Do : Do-Do-Sol-Do-Mi-Sol et ainsi de suite. On observe également l'apparition d'autres composantes liées à la partie résiduelle du signal (surtout dans les hautes fréquences).

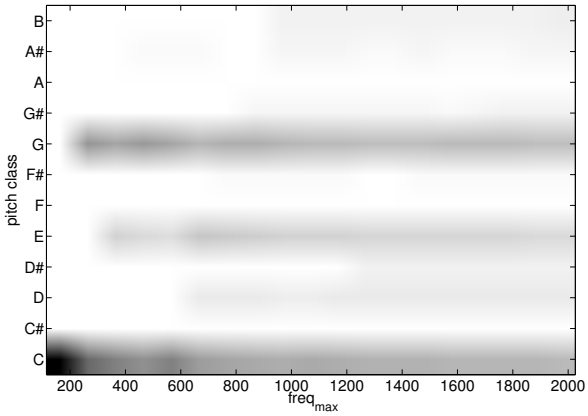


FIGURE 7.7 – Chromas obtenu pour une note Do1 jouée par un violoncelle en considérant divers intervalles de fréquences pour le calcul (de $f_{min} = 60$Hz à $f_{max} \in [100, 2000]$ Hz).

Chapitre 8

Synthèse sonore (slides, Thomas Hézard)

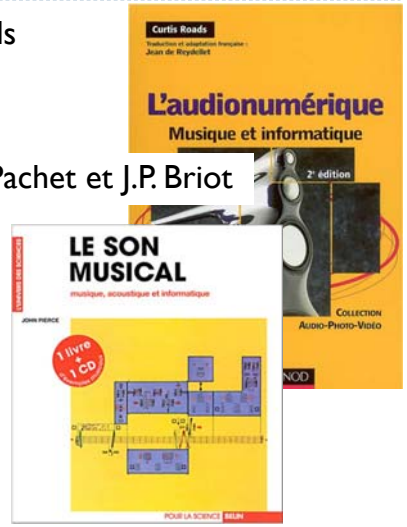
Méthodes de synthèse sonore et applications en informatique musicale

Thomas Hézard
Institut de Recherche et de Coordination Acoustique/Musique
Equipe Analyse/Synthèse de Sons
Equipe Acoustique Instrumentale

thomas.hezard@ircam.fr

Bibliographie

- L'audionumérique de C. Roads
- Informatique musicale de F. Pachet et J.P. Briot
- Le son musical de J. Pierce



Plan de la présentation

- Historique de la synthèse sonore et de l'informatique musicale
- Synthèse additive
- Synthèse soustractive
- TD / TP dirigés Matlab
- Synthèse par modulation
- Synthèse par modèles physiques
- Panorama des autres méthodes de synthèse
- TD / TP dirigés Matlab

Introduction historique

- Précurseurs et fondamentaux

1641

1ère machine à calculer de Blaise Pascal

Introduction historique

- Précurseurs et fondamentaux

1751

Invention de l'électricité par Benjamin Franklin

Introduction historique

- Précurseurs et fondamentaux

1789

Publication du mémoire de Joseph Fourier

Introduction historique

- Précurseurs et fondamentaux

1821

Charles Babbage et Ada Lovelace présentent la « machine à différence »

Introduction historique

- Précurseurs et fondamentaux

1863

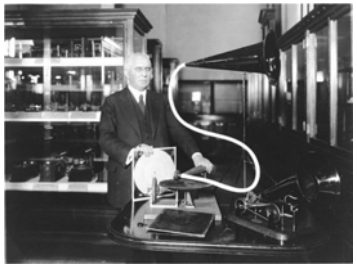
Publication du traité de Hermann von Helmholtz

Introduction historique

- Précurseurs et fondamentaux

1877

Emile Berliner invente le premier microphone

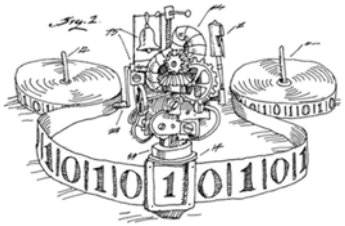


Introduction historique

- Précurseurs et fondamentaux

1937

Machine de Turing



Introduction historique

- Précurseurs et fondamentaux

1944

Mark-1, le premier ordinateur numérique (Harvard)

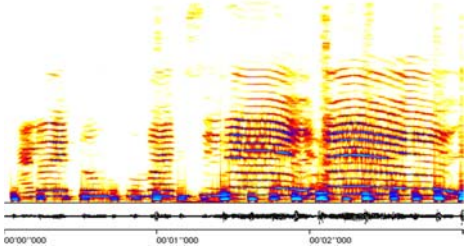
The Harvard Mark I

Introduction historique

- Précurseurs et fondamentaux

1946/1947

Invention du sonogramme



Introduction historique

- Précurseurs et fondamentaux

1947

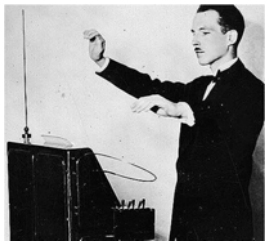
Premier transistor

Introduction historique

- Premières expériences et premières fondations

1917-1918

Thérémin et Ondes Martenot



Introduction historique

- Premières expériences et premières fondations

1951

CSIRAC

Csirac.flv

Introduction historique

- Premières expériences et premières fondations

1957

Max Mathews met au point *MUSIC I* aux Bell Labs

Introduction historique

- Premières expériences et premières fondations

1961

Première synthèse de voix chantée aux Bell Labs

Daisy Bell

Daisy.flv

Introduction historique

- Premières expériences et premières fondations

1964

Premier synthétiseur Moog

Introduction historique

- Premières expériences et premières fondations

1967

John Chowning met au point la synthèse FM

Introduction historique

- Premières expériences et premières fondations

1968

Jean-Claude Risset compose *Computer Suite for Little Boy*

shepard tone(Jean-Claude Risset).flv

Introduction historique

- Premières expériences et premières fondations

1969

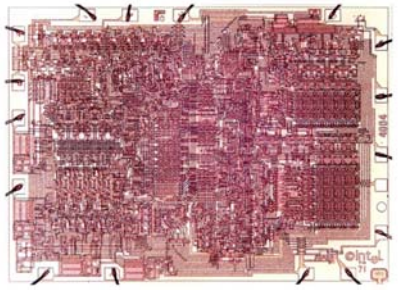
Jean-Claude Risset publie
« An introductory catalog of computer-synthesized Music »

Introduction historique

- Premières expériences et premières fondations

1971

INTEL 4004 : 1er microprocesseur



Introduction historique

- Vers le numérique

1975

Yamaha acquiert le brevet de la synthèse FM
En 1983 : Yamaha DX7

Introduction historique

- Vers le numérique

1975-76

Début des travaux scientifiques à l'IRCAM
(Boulez, Berio, Risset…)

Introduction historique

- Vers le numérique

1976-81

Premiers synthétiseurs numérique (TR, 4A, 4X…)

Introduction historique

- Vers le numérique

1978-79

Premier Walkam de Sony

Premiers CD

Introduction historique

- Vers le numérique

1978-79

Xavier Rodet (IRCAM) : CHANT (FOF)

Introduction historique

- La révolution numérique

1982

MIDI

cf Le livre d'or de la norme MIDI

Introduction historique

- La révolution numérique

1983

Karplus et Strong

Synthèse physique (cordes pincées) par guide d'onde

Introduction historique

- La révolution numérique

1985

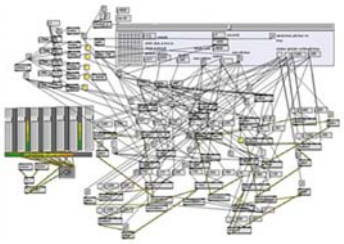
J-M Adrien (IRCAM) : Synthèse modale

Introduction historique

- La révolution numérique

1986

Miller Pucket créé la première version de Max à l'IRCAM

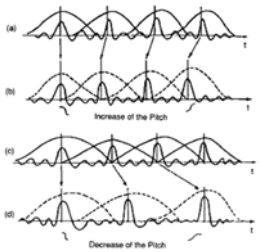


Introduction historique

- La révolution numérique

1986

PSOLA, développé à l'origine par FranceTelecom



Introduction historique

- La révolution numérique

Années 90

2 mouvements parallèles importants :

hardware => software
institutions => home studio

Introduction historique

- La révolution numérique

1993

Sensorband

Atau Tanaka at the Luminaire.flv

Introduction historique

- La révolution numérique

1999-2001

Mauro Lanza utilise massivement
la synthèse par modèles physiques

Introduction historique

- La révolution numérique

De nos jours

Evolutions actuelles :

- Contrôle gestuel
- Intelligence artificielle
- Ensemble électronique
- Musique pour réseaux

Introduction historique

- Avancées théoriques et techniques à retenir :
  - Invention de l'électricité
  - Séries de Fourier
  - « Thèse additive » du son
  - Naissance de l'informatique
  - Premiers transistors et premiers microprocesseurs
- Lieux à retenir
  - Harvard
  - Bell Labs
  - IRCAM

Introduction historique

- Chercheurs/compositeurs à retenir
  - Joseph Fourier (1768-1830)
  - Hermann Ludwig von Helmholtz (1821-1894)
  - Alan Turing (1912-1954)
  - Pierre Schaeffer (1910-1995)
  - Max Vernon Mathews (1926-)
  - John Kelly (1923-1965) & Carol Lochbaum
  - Robert Moog (1934-2005)
  - Pierre Boulez (1925-)
  - John Chowning (1934-)
  - Jean-Claude Risset (1938-)
  - …

- Historique de la synthèse sonore et de l'informatique musicale
- Synthèse additive
- Synthèse soustractive
- TD / TP dirigés Matlab
- Synthèse par modulation
- Synthèse par modèles physiques
- Panorama des autres méthodes de synthèse
- TD / TP dirigés Matlab

La synthèse additive

- Principe fondamental

Il est théoriquement possible d'approcher n'importe quelle forme d'onde complexe en additionnant des formes d'ondes élémentaires.

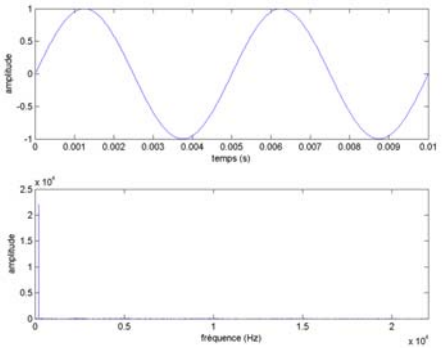
Séries de Fourier !

La synthèse additive

- Synthèse d'une onde stable au cours du temps parfaitement périodique
  - $x(t) = \sum_{k=1}^{N} A_k s_k(t)$
  - $s_k(t) = \sin(2\pi f_k t + \phi_k)$
  - $f_k = k f_0$
  - 2N+1 paramètres : $\begin{cases} A_k \\ \phi_k \\ f_0 \end{cases}$

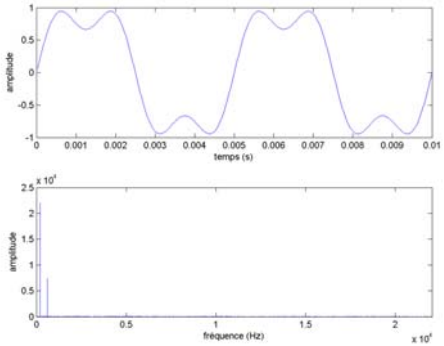
La synthèse additive

- Premier exemple
  - Addition des harmoniques d'ordre impaire avec $A_k = \frac{1}{k}$
  - $f_0 = 200$
  - 1 harmonique



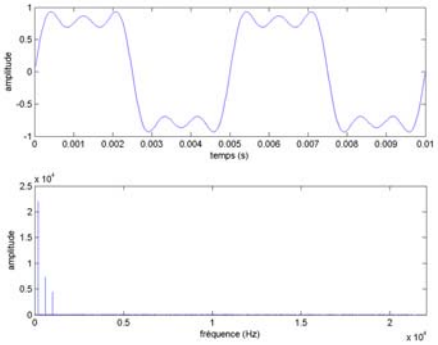
La synthèse additive

- Premier exemple
  - Addition des harmoniques d'ordre impaire avec $A_k = \frac{1}{k}$
  - $f_0 = 200$
  - 2 harmoniques



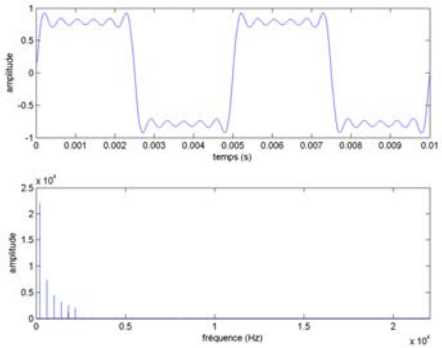
La synthèse additive

- Premier exemple
  - Addition des harmoniques d'ordre impaire avec $A_k = \frac{1}{k}$
  - $f_0 = 200$
  - 3 harmoniques



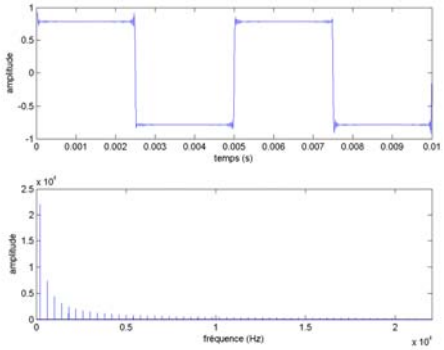
La synthèse additive

- Premier exemple
  - Addition des harmoniques d'ordre impaire avec $A_k = \frac{1}{k}$
  - $f_0 = 200$
  - 6 harmoniques



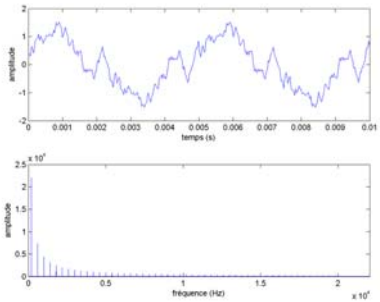
La synthèse additive

- Premier exemple
  - Addition des harmoniques d'ordre impaire avec $A_k = \frac{1}{k}$
  - $f_0 = 200$
  - 50 harmoniques



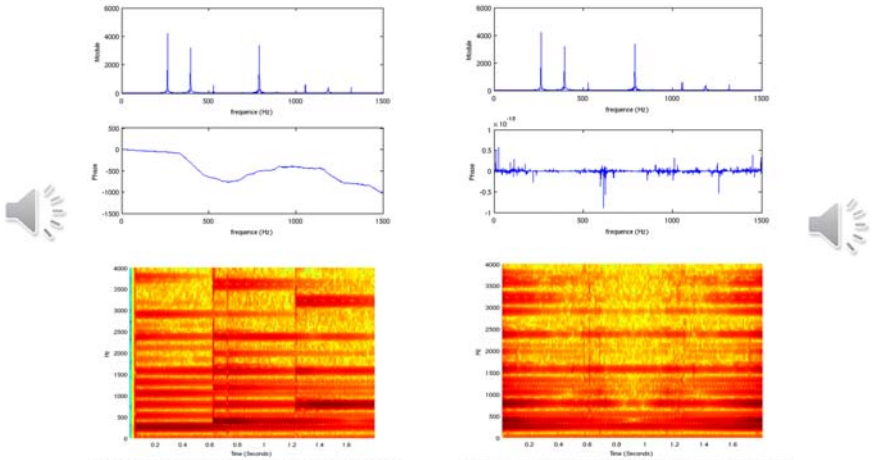
La synthèse additive

- Effet de la phase
  - Addition des harmoniques d'ordre impaire avec $A_k = \frac{1}{k}$
  - $f_0 = 200$
  - Phase aléatoire sur chaque harmonique
  - 50 harmoniques



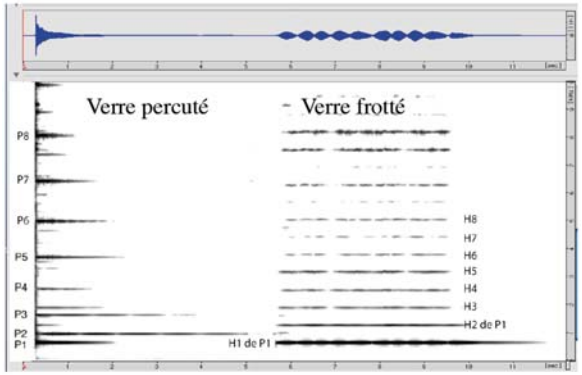
La synthèse additive

- Petite parenthèse sur la phase



La synthèse additive

- Harmoniques vs. Partiels
  - ⇔ Son apériodique vs Son périodique
  - ⇔ Impulsion vs Entretien



La synthèse additive

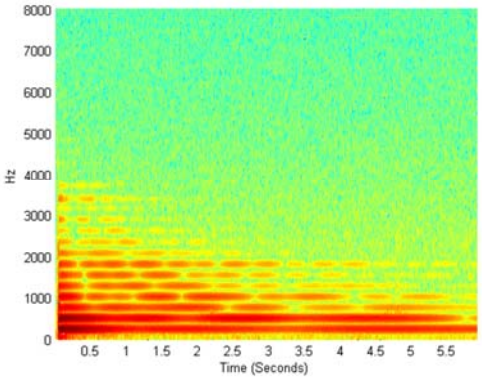
- Modèle généralisé
  - $x(t) = \sum_{k=1}^{N} A_k s_k(t)$
  - $s_k(t) = \sin(2\pi f_k t + \phi_k)$
  - f_k sont des fréquences quelconques
  - 3N paramètres : $\begin{cases} A_k \\ f_k \\ \phi_k \end{cases}$

La synthèse additive

- Analyse
  - Il suffit de calculer les coefficients de la série de Fourier !

La synthèse additive

- Synthèse d'une onde variant au cours du temps « localement périodique »



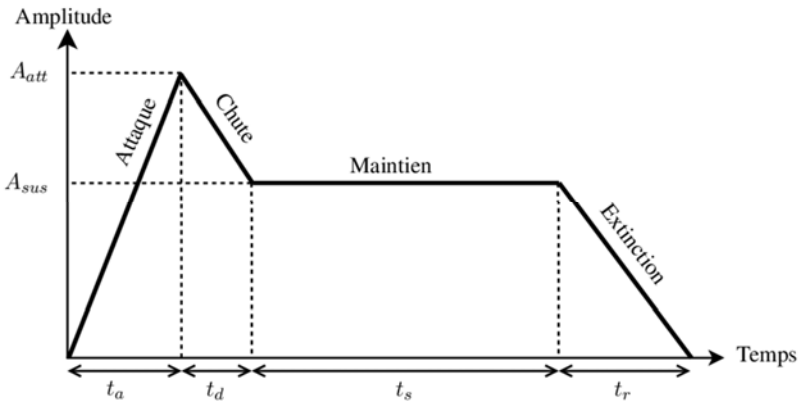
La synthèse additive

- Synthèse d'une onde variant au cours du temps
  - $x(t) = \sum_{k=1}^{N} A_k(t) s_k(t)$
  - $s_k(t) = \sin(2\pi f_k(t) t + \phi_k)$

Attention : différent de la FM/AM, les fonctions $A_k(t)$ et $f_k(t)$ varient lentement !

La synthèse additive

- Enveloppe d'amplitude ADSR

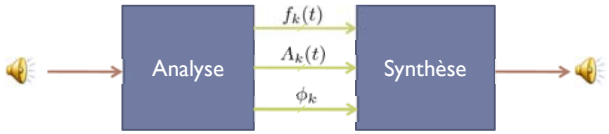


La synthèse additive

- Contrôle de la synthèse additive
  - Un évènement
    - = N partiels (souvent autour de 24 pour les sons instrumentaux)
    - = N x 2 enveloppes (amplitudes et fréquences)
  - Une partition d'orchestre
    - = 10 000 évènements
    - = 20 000 x N enveloppes (480 000 si N = 24)
  - Où trouver tous ces paramètres ?
    - Rentrés à la main par le compositeur
    - Logiciel de composition (semi)automatique
    - Données d'analyse d'un autre domaine que la musique
    - Analyse de sons musicaux pour la re-synthèse (avec modifications)

La synthèse additive

- Analyse-resynthèse d'un son variant dans le temps



- Grand nombre de méthodes d'analyse
  - Approximation polynomiale (ou splines) de l'enveloppe
  - Modélisation exponentielle de l'enveloppe (sons percussifs)
  - Suivi de fréquence des partiels (filtrage adaptatif…)
  - …
  - cf biblio + travaux d'Anssi Kalpuri
    - *Signal Processing Methods for the Automatic Transcription of Music (PhD)*

Synthèse additive

- Extensions de la synthèse additive
  - Musique spectrale : Gondwana de Tristan Murail

Tristan Murail _ Gondwana (1_2).flv

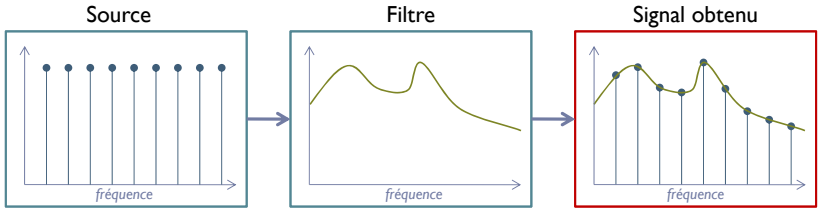
  - Speaking Piano

Speaking Piano.flv

- Historique de la synthèse sonore et de l'informatique musicale
- Synthèse additive
- Synthèse soustractive
- TD / TP dirigés Matlab
- Synthèse par modulation
- Synthèse par modèles physiques
- Panorama des autres méthodes de synthèse
- TD / TP dirigés Matlab

La synthèse soustractive

- Paradigme source-filtre



La synthèse soustractive

- Rappels sur le filtrage numérique
  - Domaine temporel

Soient $e[n]$ un signal de source et $h[n]$ une réponse impulsionnelle de filtre. Le filtrage linéaire de e par h donne le signal $s[n]$ et s'écrit, à l'aide de l'opérateur de convolution $\star$:

$$s[n] = (e \star h)[n] = \sum_{k=-\infty}^{+\infty} h[k]e[n-k] = \sum_{k=-\infty}^{+\infty} h[n-k]e[n]$$

  - Domaine fréquentiel

On note $E(f)$, $H(f)$ et $S(f)$ les TFD correspondantes. On note de plus $E(z)$, $H(z)$ et $S(z)$ les transformées en z. On a alors les relations suivantes :

$$\begin{aligned} S(z) &= H(z).E(z) \\ S(f) &= H(f).E(f) \end{aligned}$$

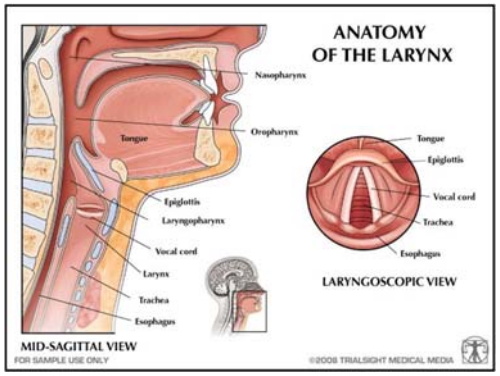
La synthèse soustractive

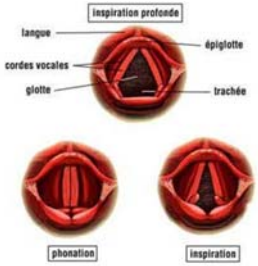
- Modèle de synthèse soustractive
  - Une source
  - Un filtre

La source et le filtre peuvent être de toute sorte (à condition que leur convolution soit stable). Suivant les applications, on pourra utiliser des modèles paramétriques de source et de filtre plus ou moins complexes. L'exemple de synthèse soustractive le plus utilisé est sans doute le modèle AR de la parole.

Synthèse soustractive

- Modélisation AR de la parole

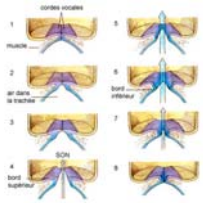


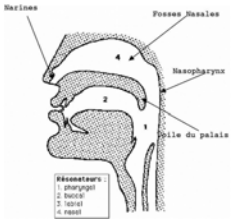


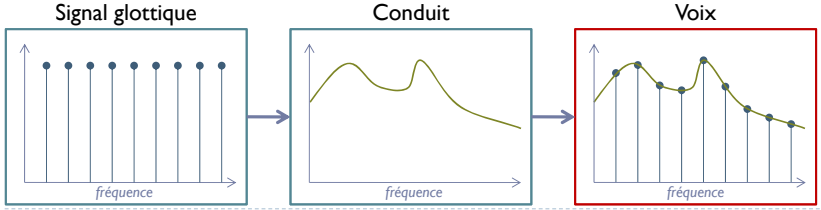
- *http://catalogue.ircam.fr/sites/Voix/index.html*

Synthèse soustractive

- Modélisation AR de la parole
  - Hypothèse sous-jacente : indépendance glotte-conduit

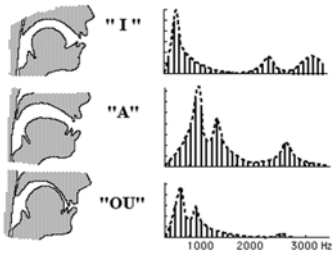






Synthèse soustractive

- Modélisation AR de la parole

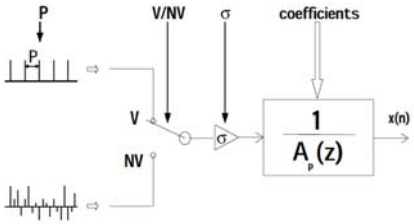


- Le filtre correspondant au conduit vocal est assimilé à un filtre AR, c-à-d à un filtre dont la TZ s'écrit

$$H(z) = \frac{1}{1 + \sum_{k=1}^{P} a_k z^{-k}}$$

Synthèse soustractive

- Modélisation AR de la parole : Analyse
  - Décision voisé/non-voisé
  - Détection de pitch dans le cas voisé
  - Détermination de la puissance
  - Détermination des coefficients du filtre
- Modélisation AR de la parole : Synthèse



$$X(z) = \sigma \frac{1}{1 + \sum_{k=1}^{P} a_k z^{-k}} U(z)$$

Synthèse soustractive

- Modélisation AR de la parole : Analyse

$$X(z) = \sigma \frac{1}{1 + \sum_{k=1}^{P} a_k z^{-k}} U(z)$$

$$X(z) + \sum_{k=1}^{P} a_k z^{-k} X(z) = \sigma U(z)$$

$$x[n] + \sum_{k=1}^{P} a_k x[n-k] = \sigma u[n]$$

$$x[n] = \sigma u[n] - \sum_{k=1}^{P} a_k x[n-k]$$

Synthèse soustractive

- Modélisation AR de la parole : Analyse
  - Estimation des paramètres
    - Equation de Yule-Walker

$$\boxed{R_P \alpha_P = C_P} \Rightarrow \boxed{\alpha_P = {R_P}^{-1} C_P}$$

$$R_P = \begin{pmatrix} c_x[0] & c_x[1] & \cdots & c_x[p-1] \\ c_x[1] & c_x[0] & \cdots & c_x[p-] \\ & & \ddots & \\ c_x[p-1] & c_x[p-2] & \cdots & c_x[0] \end{pmatrix}$$

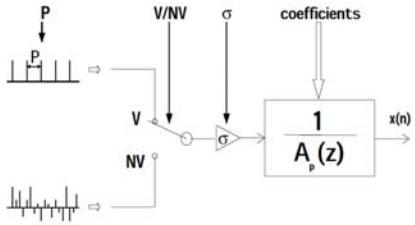
$$\alpha_P = [a_1, \cdots, a_P]$$

$$C_P = \left[c_x[1], \cdots, c_x[P]\right]^T$$

- Estimateur de l'autocovariance : $c_x(k) = \frac{1}{N} \sum_{j=1}^{N-k} x[j+k]x[j]$

Synthèse soustractive

- Modélisation AR de la parole : Analyse
  - Décision voisé/non-voisé
  - Détection de pitch dans le cas voisé
  - Détermination de la puissance
  - Détermination des coefficients du filtre
- Modélisation AR de la parole : Synthèse



$$X(z) = \sigma \frac{1}{1 + \sum_{k=1}^{P} a_k z^{-k}} U(z)$$

Synthèse soustractive

- Modélisation AR de la parole
  - est utilisée (avec quelques améliorations) dans la téléphonie mobile
  - permet de faire de l'extraction de filtre vocal
  - …
- Jonathan Harvey : *Speakings*

et-l-ircam-fit-parler-l-orchestre_M39047.flv

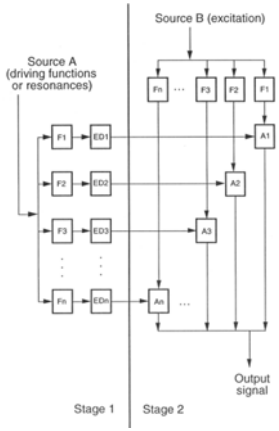
Synthèse soustractive

- Vocoder et synthèse croisée

Vocoder_Moog.flv

Synthèse soustractive

- Vocoder : système d'analyse/resynthèse soustractive



Synthèse soustractive

- Vocodeur de phase :

  Analyse = Transformée de Fourier Court Terme (TFCT)

$$x[n] \xrightarrow{\text{TFCT}} X(n,f) = \sum_{m=-\infty}^{+\infty} \left(w[n-m]x[m]\right) e^{-2i\pi f m}$$

Echantillonnage spectral

$w[n]$ est la fenêtre d'analyse

$$X\left[n, \frac{k}{N}\right] = \sum_{m=-\infty}^{+\infty} \left(w[n-m]x[m]\right) e^{-2i\pi \frac{km}{N}}$$

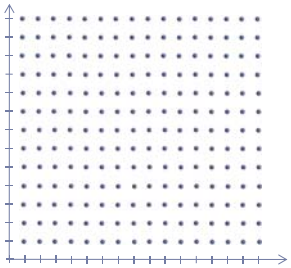
$\frac{k}{N}$ est une fréquence réduite.

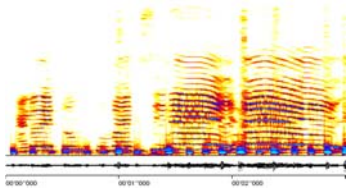
- Il est possible de sous-échantillonner en temporel (dépend de la largeur de bande de la fenêtre d'analyse)

Synthèse soustractive

- Vocodeur de phase :
  Analyse = Transformée de Fourier Court Terme (TFCT)

$$x[n] \longrightarrow X\left[n, \frac{k}{N}\right]$$





Synthèse soustractive

- Vocodeur de phase : Synthèse

$$\begin{aligned} y[n] &= \sum_{k=0}^{N-1} X\left[n, \frac{k}{N}\right] e^{i2\pi\frac{kn}{N}} \\ &\ldots \\ &= N \sum_{r=-\infty}^{+\infty} w[rN]x[n+rN] \end{aligned}$$

  - On retrouve $x[n]$ (à un facteur près) si $w[rN] = 0 \ \forall r \neq 0$

Synthèse soustractive

- Synthèse croisée à l'aide du vocodeur de phase

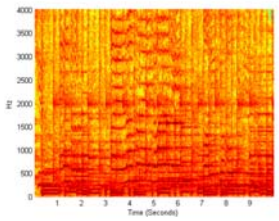
$$x[n] \rightarrow X\left[n, \frac{k}{N}\right]$$

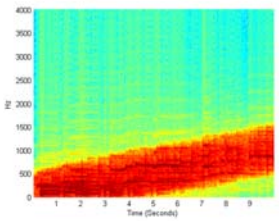
$$y[n] \rightarrow Y\left[n, \frac{k}{N}\right]$$

  - On peut donc faire du filtrage qui varie dans le temps ou même « mélanger fréquentiellement » deux signaux. C'est ce qu'on appelle la « synthèse croisée ».

Synthèse soustractive

- Synthèse croisée à l'aide du vocodeur de phase
  - Utilisations possibles :
    - filtrage (équilibrage fréquentiel) qui varie dans le temps





    - Vocoder Moog !

- Historique de la synthèse sonore et de l'informatique musicale
- Synthèse additive
- Synthèse soustractive
- TD / TP dirigés Matlab
- Synthèse par modulation
- Synthèse par modèles physiques
- Panorama des autres méthodes de synthèse
- TD / TP dirigés Matlab

Méthodes de synthèse sonore et applications en informatique musicale

Thomas Hézard
Institut de Recherche et de Coordination Acoustique/Musique
Equipe Analyse/Synthèse de Sons
Equipe Acoustique Instrumentale

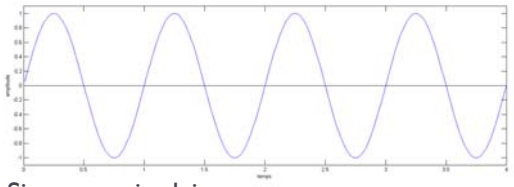
thomas.hezard@ircam.fr

- Historique de la synthèse sonore et de l'informatique musicale
- Synthèse additive
- Synthèse soustractive
- TD / TP dirigés Matlab
- Synthèse par modulation
- Synthèse par modèles physiques
- Panorama des autres méthodes de synthèse
- TD / TP dirigés Matlab

- Historique de la synthèse sonore et de l'informatique musicale
- Synthèse additive
- Synthèse soustractive
- TD / TP dirigés Matlab
- Synthèse par modulation
- Synthèse par modèles physiques
- Panorama des autres méthodes de synthèse
- TD / TP dirigés Matlab

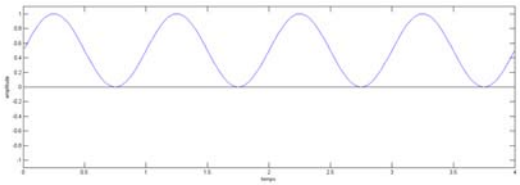
Synthèse par modulation

- Vocabulaire
  - Signaux bipolaires



$$s(t) = \sin(2\pi f t)$$

  - Signaux unipolaires



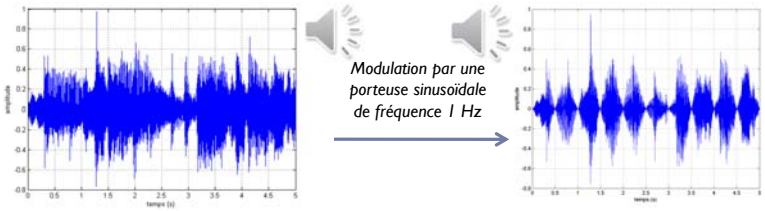
$$s(t) = \frac{1 + \sin(2\pi f t)}{2}$$

Synthèse par modulation

- Modulation en anneaux
  - Multiplication d'un signal porteur par un signal modulant.
    - $s(t) = p(t) \times m(t)$
  - Souvent, le signal modulant est une sinusoïde
    - $s(t) = p(t) \times \sin(2\pi f_0 t)$

Synthèse par modulation

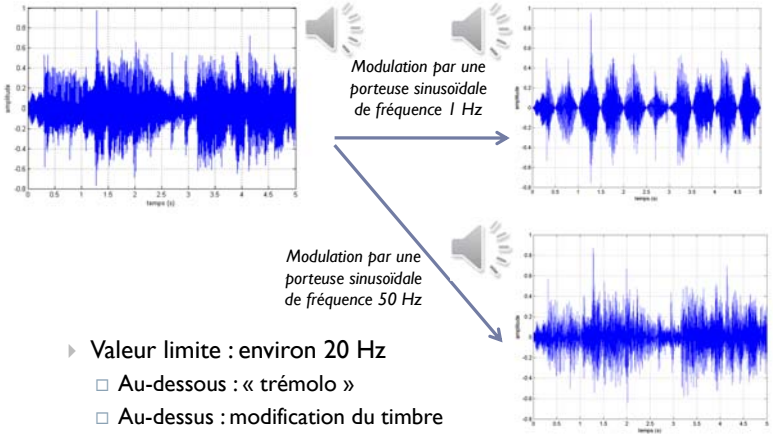
- Influence de la fréquence de la porteuse sinusoïdale



Remarque : Le trémolo perçu a une fréquence de 2 Hz car le signal modulant est une sinudoïde pure donc bipolaire

Synthèse par modulation

- Influence de la fréquence de la porteuse sinusoïdale

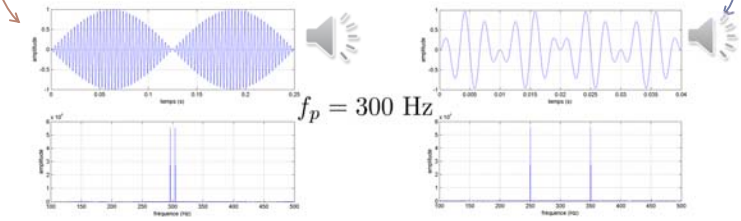


- Valeur limite : environ 20 Hz
  - Au-dessous : « trémolo »
  - Au-dessus : modification du timbre

Synthèse par modulation

- Modulation en anneaux de deux sinusoïdes

$$
\begin{aligned}
s(t) &= \boxed{\sin(2\pi f_p t) \times \sin(2\pi f_m t)} \\
&= \frac{1}{2}\Big[\cos\big(2\pi(f_p - f_m)t\big) - \cos\big(2\pi(f_p + f_m)t\big)\Big] \\
&= \boxed{\frac{1}{2}\left[\sin\left(2\pi(f_p - f_m) + \frac{\pi}{2}\right) + \sin\left(2\pi(f_p + f_m)t - \frac{\pi}{2}\right)\right]}
\end{aligned}
$$



$f_p = 300$ Hz

$f_m = 4$ Hz : effet « trémolo »

$f_m = 50$ Hz : modification du timbre

Synthèse par modulation

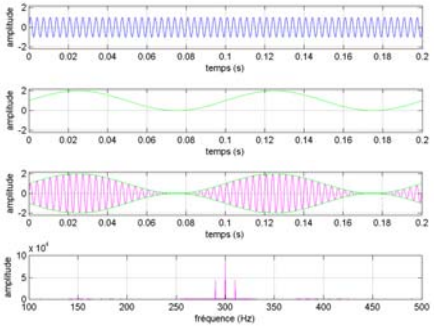
- Modulation d'amplitude : Synthèse AM
  - Multiplication d'une sinusoïde porteuse par une sinusoïde modulante unipolaire

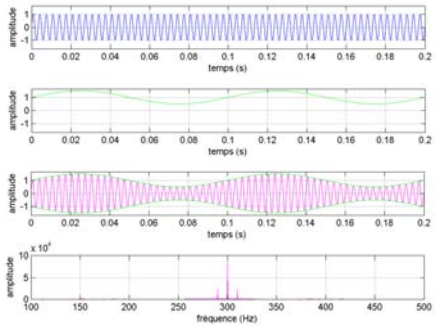
$$\begin{aligned} s(t) &= \sin(2\pi f_p t) \times (1 + \alpha \sin(2\pi f_m t)) \\ &= \sin(2\pi f_p t) + \frac{\alpha}{2}\left[\cos\left(2\pi(f_p - f_m)t\right) - \cos\left(2\pi(f_p + f_m)t\right)\right] \\ &= \sin(2\pi f_p t) + \frac{\alpha}{2}\left[\sin\left(2\pi(f_p - f_m) + \frac{\pi}{2}\right) + \sin\left(2\pi(f_p + f_m)t - \frac{\pi}{2}\right)\right] \end{aligned}$$

- 3 paramètres $\begin{cases} f_p \text{ fréquence porteuse} \\ f_m \text{ fréquence modulante} \\ \alpha \text{ indice de modulation} \end{cases}$

Synthèse par modulation

- $\begin{cases} f_p &= 300 \text{ Hz} \\ f_m &= 10 \text{ Hz} \\ \alpha &= 1 \end{cases}$
- $\begin{cases} f_p &= 300 \text{ Hz} \\ f_m &= 10 \text{ Hz} \\ \alpha &= 0.5 \end{cases}$

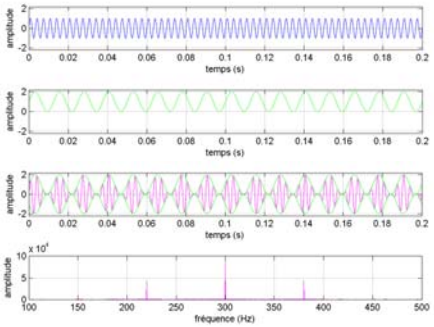


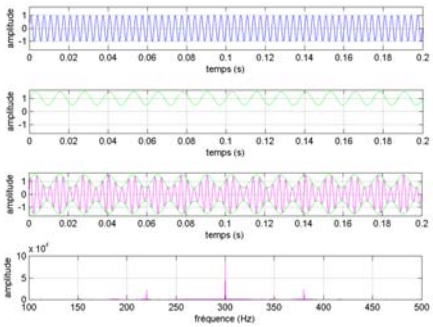


$f_m < 20$ Hz : **effet « trémolo »**

Synthèse par modulation

- $\begin{cases} f_p &= 300 \text{ Hz} \\ f_m &= 80 \text{ Hz} \\ \alpha &= 1 \end{cases}$
- $\begin{cases} f_p &= 300 \text{ Hz} \\ f_m &= 80 \text{ Hz} \\ \alpha &= 0.5 \end{cases}$





$f_m > 20$ Hz : **modification de timbre**

Synthèse par modulation

- Modulation de fréquence : Synthèse FM
  - Composition de deux sinusoïdes

$$\begin{aligned} s(t) &= \sin\left(2\pi f_p t + I_m \sin(2\pi f_m t)\right) \\ &= \sin\left(\phi(t)\right) \end{aligned}$$

  - Fréquence instantanée

$$\begin{aligned} \omega(t) = 2\pi f(t) &= \frac{d\phi}{dt}(t) \\ &= 2\pi f_p + I_m 2\pi f_m \cos(2\pi f_m t) \end{aligned}$$

$$f(t) = f_p + \delta_f \cos(2\pi f_m t) \text{ avec } \delta_f = I_m f_m$$

  - Modèle
    - $s(t) = \sin\left(2\pi f_p t + \frac{\delta_f}{f_m}\sin(2\pi f_m t)\right)$ 3 paramètres : $\begin{cases} f_p \\ f_m \\ \delta_f \text{ (ou } I_m) \end{cases}$

Synthèse par modulation

- Décomposition fréquentielle

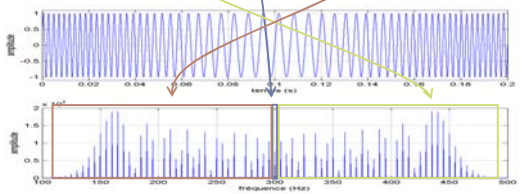
$$
\begin{aligned}
s(t) &= \sin\left(2\pi f_p t + I_m \sin(2\pi f_m t)\right) \\
s(t) &= \boxed{J_0(I_m)\sin(2\pi f_p t)} \\
&\quad + \sum_{k=1}^{+\infty} J_k(I_m)\left[\boxed{\sin\left(2\pi(f_p + k f_m)t\right)} + (-1)^k \boxed{\sin\left(2\pi(f_p - k f_m)t\right)}\right]
\end{aligned}
$$

$$
J_n(x) = \sum_{k=0}^{+\infty} \frac{(-1)^k (x/2)^{n+2k}}{k!\Gamma(n+k+1)}
$$

fonctions de Bessel

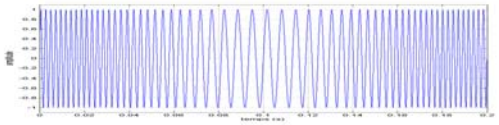
$$
\begin{aligned}
\Gamma(z) &= \int_0^{+\infty} t^{z-1} e^{-t} dt \\
\Gamma(n+1) &= n! \ \forall n \in \mathbb{N}
\end{aligned}
$$

$$
\begin{cases}
f_p &= 300 \text{ Hz} \\
f_m &= 5 \text{ Hz} \\
\delta_f &= 150 \text{ Hz}
\end{cases}
$$



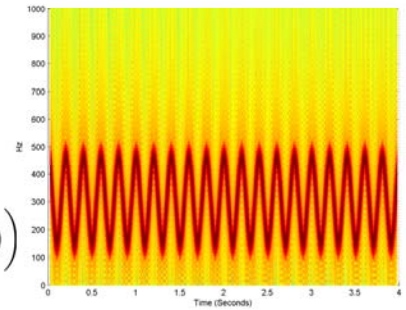
Synthèse par modulation

- Fréquence modulante faible (< 20 Hz) : vibrato



$$
\begin{cases}
f_p &= 300 \text{ Hz} \\
f_m &= 5 \text{ Hz} \\
\delta_f &= 150 \text{ Hz}
\end{cases}
$$

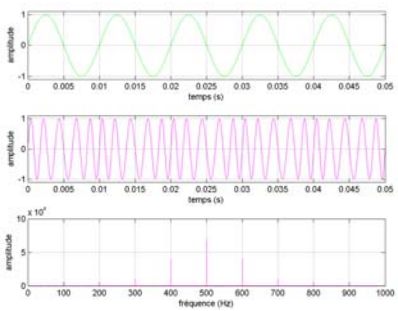
$$
s(t) = \sin\left(2\pi f_p t + \frac{\delta_f}{f_m}\sin(2\pi f_m t)\right)
$$

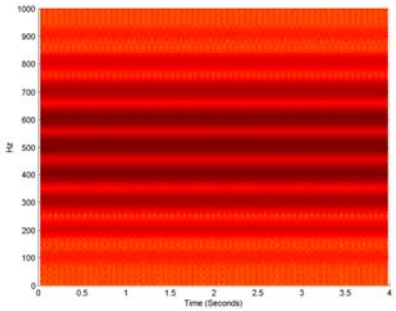


Synthèse par modulation

- Fréquence modulante élevée (> 20 Hz) : son complexe

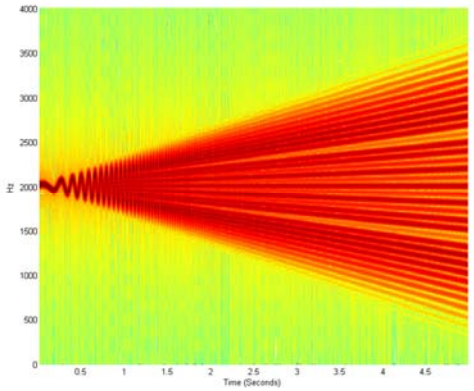
$$
\begin{cases}
f_p &= 500 \text{ Hz} \\
f_m &= 100 \text{ Hz} \\
\delta_f &= 100 \text{ Hz}
\end{cases}
$$





Synthèse par modulation

- Passage du vibrato à la modification de timbre



$$
\begin{cases}
f_p &= 2000 \text{ Hz} \\
f_m &= [1, 50] \text{ Hz} \\
I_m &= 10 \text{ Hz}
\end{cases}
$$

Synthèse par modulation

- Paramètres variant dans le temps, sons complexes

$$s(t) = \sin\left(2\pi f_p(t)t + I_m(t)\sin\left(2\pi f_m(t)t\right)\right) \quad \begin{cases} f_p(t) = a_p + b_p\ t \\ f_m(t) = a_m + b_m\ t \\ I_m(t) = a_i + b_i\ t \end{cases}$$

Synthèse par modulation

- Scripts Matlab
  - modAM.m

$$s(t) = \sin(2\pi f_p t) \times (1 + \alpha \sin(2\pi f_m t))$$

$$\begin{cases} f_p \\ f_m \\ \alpha \end{cases}$$ fonctions affines

```
% Paramètres **********
Fe = 44100;          % Fréquence d'échantillonnage (Hz)
Dursig = 4;          % Durée du signal (s)
FportDeb = 500;      % Fréquence porteuse à t = 0
FportFin = 500;      % Fréquence porteuse à t = Dursig
FmodDeb = 10;        % Fréquence de modulation à t = 0
FmodFin = 10;        % Fréquence de modulation à t = Dursig
alphaDeb = 1;        % Indice de modulation à t = 0
alphaFin = 1;        % Indice de modulation à t = Dursig
% ********************
```

  - modFM.m

$$s(t) = \sin\left(2\pi f_p t + I_m \sin(2\pi f_m t)\right)$$

$$\begin{cases} f_p \\ f_m \\ I_m \end{cases}$$ fonctions affines

```
% Paramètres **********
Fe = 44100;          % Fréquence d'échantillonnage (Hz)
Dursig = 4;          % Durée du signal (s)
FportDeb = 500;      % Fréquence porteuse à t = 0
FportFin = 500;      % Fréquence porteuse à t = Dursig
FmodDeb = 10;        % Fréquence de modulation à t = 0
FmodFin = 10;        % Fréquence de modulation à t = Dursig
IndMDeb = 10;        % Indice de modulation à t = 0
IndMFin = 10;        % Indice de modulation à t = Dursig
% ********************
```


Synthèse par modulation

- Démonstration Pure Data

- Historique de la synthèse sonore et de l'informatique musicale
- Synthèse additive
- Synthèse soustractive
- TD / TP dirigés Matlab
- Synthèse par modulation
- Synthèse par modèles physiques
- Panorama des autres méthodes de synthèse
- TD / TP dirigés Matlab

Synthèse par modèles physiques

- Modèles de signaux vs. Modèles physiques
  - Conséquences vs. Causes
  - Contrôlabilité vs. Précision
  - Efficacité vs. Coût de calcul
  - Calcul des paramètres vs. Inversion du modèle

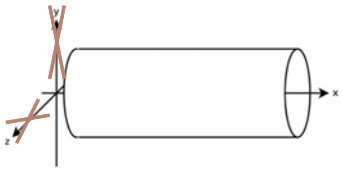
- Modèles mathématiques de la mécanique et de l'acoustique de la production sonore dans les instruments de musique

Synthèse par modèles physiques

- Plusieurs approches
  - Résolution numérique des équations mécaniques et acoustiques
    - Juliette Chabassier : Modélisation et simulation numérique d'un piano de concert http://www-rocq.inria.fr/who/Juliette.Chabassier/Accueil.html
  - Eléments finis
  - Synthèse modale
    - Modalys (IRCAM) http://forumnet.ircam.fr/701.html
  - Modèle virtuel – Cordis-Anima (Claude Cadoz, ACROE)
    - http://www-acroe.imag.fr/ACROE/recherche/cordis-anima/cordis-anima.html
  - ***Modèles par guides d'ondes***
    - ***Julius Orion Smith III*** ***https://ccrma.stanford.edu/~jos/***

Synthèse par modèles physiques

- Propagation du son dans un tube acoustique idéal
  - Tube axisymétrique sans pertes
  - Ondes planes



- Notations
  - Surface $S(x)$
  - Pression $p(x,t),\ P(x,\omega)$
  - Débit $u(x,t),\ U(x,\omega)$

Célérité du son c

Masse volumique ρ

Synthèse par modèles physiques

- Equation des ondes
  - Conservation de la quantité de mouvement 1D

$$S(x)\frac{\partial p(x,t)}{\partial x} = -\rho\frac{\partial u(x,t)}{\partial t}$$

  - Conservation de la masse 1D

$$\rho\frac{\partial u(x,t)}{\partial x} = -\frac{S(x)}{c^2}\frac{\partial p(x,t)}{\partial t}$$

  - Equation sur le débit

$$\frac{\partial}{\partial x}\left[\frac{1}{S(x)}\frac{\partial u(x,t)}{\partial x}\right] = \frac{1}{S(x)c^2}\frac{\partial^2 u(x,t)}{\partial t^2} \xrightarrow{\text{tube droit}} \frac{\partial^2 u(x,t)}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 u(x,t)}{\partial t^2}$$

  - Equation sur la pression

$$\frac{\partial^2 p(x,t)}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 p(x,t)}{\partial t^2}$$

Synthèse par modèles physiques

- Résolution de l'équation 1D : solution de d'Alembert

$$\begin{aligned}
\frac{\partial^2 p(x,t)}{\partial t^2} &= c^2 \frac{\partial^2 p(x,t)}{\partial x^2} \\
\frac{\partial^2 p(x,t)}{\partial t^2} - c^2 \frac{\partial^2 p(x,t)}{\partial x^2} &= 0 \\
\left(\frac{\partial p(x,t)}{\partial t} - c\frac{\partial p(x,t)}{\partial x}\right)\left(\frac{\partial p(x,t)}{\partial t} + c\frac{\partial p(x,t)}{\partial x}\right) &= 0
\end{aligned}$$

$$p(x,t) = p^+\left(t + \frac{x}{c}\right) + p^-\left(t - \frac{x}{c}\right)$$

- On résout de même

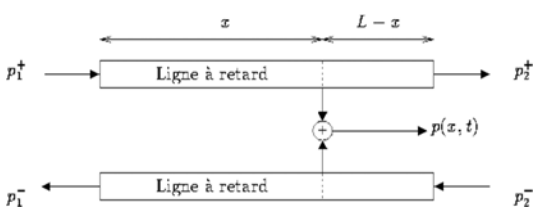
$$u(x,t) = u^+\left(t + \frac{x}{c}\right) + u^-\left(t - \frac{x}{c}\right)$$

Synthèse par modèles physiques

- Impédance caractéristique
  - On définit l'impédance caractéristique $Z(x) = \dfrac{\rho c}{S(x)}$
  - On obtient alors $\begin{aligned} p^+(t + x/c) &= Z(x)\, u^+(t + x/c) \\ p^-(t - x/c) &= -\, Z(x)\, u^-(t - x/c) \end{aligned}$

Synthèse par modèles physiques

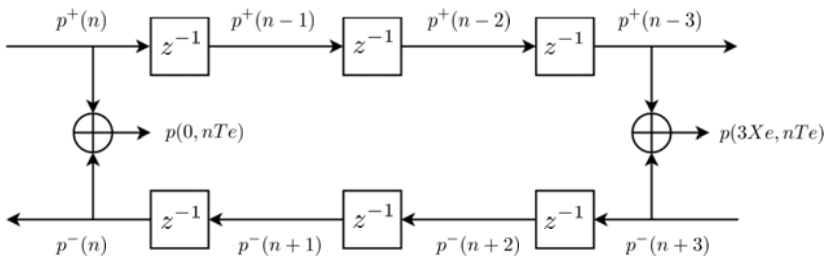
- Modélisation guide d'ondes



Synthèse par modèles physiques

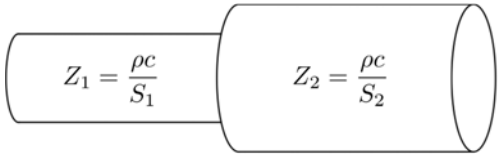
- Discrétisation
  - En espace $x \to x_m = mX_e$
  - En temps $t \to t_n = nTe$

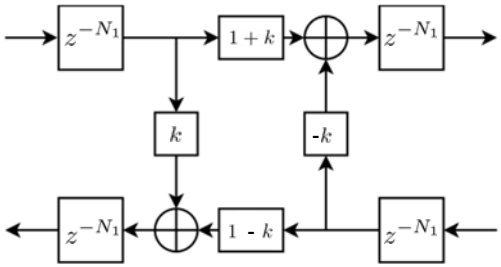
$$p(t_n, x_m) = p^+(n-m) + p^-(n+m)$$



Synthèse par modèles physiques

- Jonction de deux cylindres de rayons différents

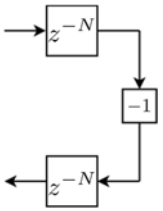




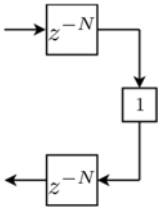
$$k = \frac{Z_2 - Z_1}{Z_2 + Z_1}$$

Synthèse par modèles physiques

- Tube ouvert, tube fermé
  - Extrémité ouverte : $p^+ = -p \quad \Leftrightarrow$

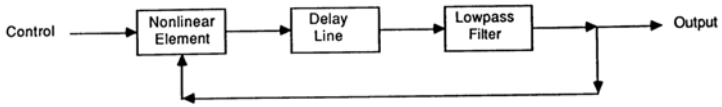


  - Extrémité fermée : $p^+ = p^- \quad \Leftrightarrow$

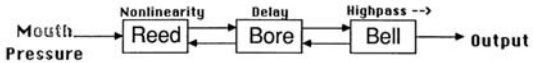


Synthèse par modèles physiques

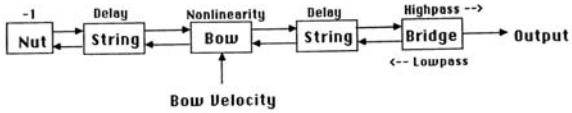
- Instrument complet



- Clarinette

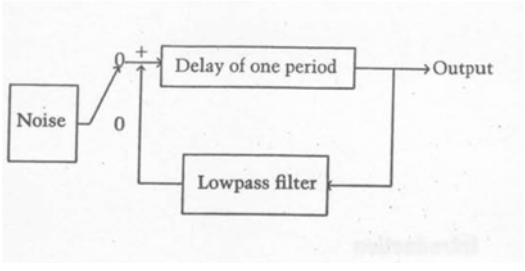


- Violon



Synthèse par modèles physiques

- Une implémentation des guides d'ondes pour les cordes : Algorithme de Karplus-Strong



- Historique de la synthèse sonore et de l'informatique musicale
- Synthèse additive
- Synthèse soustractive
- TD / TP dirigés Matlab
- Synthèse par modulation
- Synthèse par modèles physiques
- Panorama des autres méthodes de synthèse
- TD / TP dirigés Matlab

Autres méthodes de synthèse

- Echantillonnage et tables d'onde
  - Changement de la vitesse de lecture
  - Bouclage
  - Transposition
- Tables d'ondes multiples
  - Fondu enchaîné
  - Empilement

Autres méthodes de synthèse

- Lecture d'échantillons
  - Très utilisée en MAO
  - Combinée avec la synthèse soustractive
  - Inconvénient : très lourd (stockage)
- Exemple d'applications : le Prophet VS (Sequential Circuit)

:t;Vector Synthe

Autres méthodes de synthèse sonore

- Filtrage non-linéaire
  - Distorsion non linéaire (waveshaping)
  - Distorsion de phase
- Synthèse analogique virtuelle
  - Orosys / Two Notes http://www.two-notes.com/

Autres méthodes de synthèse

- Synthèse granulaire (Gabor 1947, Xenakis 1971)
  - Extension de la synthèse additive
  - Grains sonores de l'ordre quelques ms
  - Contrôle
    - Densité des grains
    - Hauteur
    - Longueur
    - Enveloppe
  - Contrôle semi-automatique
  - FOF est une sorte de synthèse granulaire

Autres méthodes de synthèse

- Synthèse granulaire
  - Concret PH, I. Xenakis (1958, Exposition Internationale de Bruxelles)

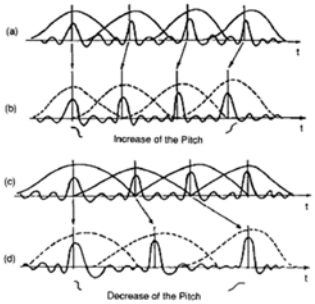
ncret PH.flv (Ligne de c

  - Contrôle de synthèse granulaire à la Wiimote

_h264_aac.mp4 (Ligr

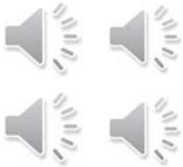
Autres méthodes de synthèse

- PSOLA = granulaire adapté à la voix ?
  - Concept de PSOLA



Autres méthodes de synthèse

- PSOLA
  - Traitement
    - Modification de la durée
    - Modification de la hauteur
  - Synthèse
    - MBROLA : Thierry Dutoit
    - Acapela-tv : http://www.acapela.tv



Autres méthodes de synthèse

- Synthèse graphique
  - Synthèse TFCT à partir d'une image
  - Dessiner un spectrogramme
  - Exemple : Mycène Alpha (I. Xenakis, 1978) crée avec l'UPIC

ae Alpha.flv (Ligne de comn

  - Logiciel libre : Coagula

Autres méthodes de synthèse

- Mélanger tout ça : à vous de jouer !

- Historique de la synthèse sonore et de l'informatique musicale
- Synthèse additive
- Synthèse soustractive
- TD / TP dirigés Matlab
- Synthèse par modulation
- Synthèse par modèles physiques
- Panorama des autres méthodes de synthèse
- TD / TP dirigés Matlab