

I. LA NATURE DU SON MUSICAL

I.1. Introduction

Depuis que le disque compact est devenu la norme pour la diffusion de la musique dite de « haute fidélité », il est admis par tout le monde que tout son musical peut être codé numériquement sans perte d'information du point de vue de la perception auditive¹.

La synthèse sonore sur ordinateur permet la production de tous types de sons. Les champs d'investigation qu'elle offre sont illimités et, face à toutes ces possibilités, le compositeur doit faire des choix. Les nombreuses expérimentations effectuées au hasard afin d'explorer l'immense domaine des sons synthétiques ont rarement donné des résultats satisfaisants. Pour produire les sons qu'il imagine, le compositeur doit disposer de points de repère. Ces points de repère, il peut les trouver par la théorie et la pratique, en développant ses connaissances en acoustique.

I.2. L'acoustique physique

I.2.1. Historique

Les théories fondamentales de l'acoustique ont été énoncées à partir du XVII^e siècle, mais il a fallu attendre l'ère de la mécanisation puis l'avènement de l'informatique pour obtenir des vérifications scientifiques et permettre de nouvelles applications de ces théories.

L'acoustique est la science du son. C'est une discipline scientifique relevant du domaine de la physique qui se divise en trois secteurs : la production du son, sa propagation et sa réception ou ses effets. La notion de son n'est pas limitée aux phénomènes aériens responsables de la sensation auditive. Elle concerne également les phénomènes basés sur des principes physiques analogues : études portant sur les infrasons et les ultrasons, trop graves ou trop aigus pour être perçus par l'oreille humaine ou études portant sur la transmission des sons dans d'autres milieux que l'air.

L'acoustique, à la différence de l'optique ou des ondes radioélectriques, ne concerne que les mouvements ondulatoires mécaniques.

L'acoustique est liée à l'étude des rayonnements mécaniques générés par des causes naturelles et par l'activité humaine auxquels sont associées les notions de sensation auditive, de voix humaine, de communication par le canal sonore et toute une variété d'incidences psychologiques. La parole, la musique, l'enregistrement et la reproduction des sons, la téléphonie, l'amplification, l'audiologie, l'acoustique architecturale ou le contrôle acoustique sont intimement liés à la sensation auditive.

¹ La norme de codage numérique du CD est à 44,1 kHz pour le taux d'échantillonnage et à 16 bits pour la résolution d'amplitude. Cela confère au son numérique une bande passante allant jusqu'à 22 kHz et une dynamique, ou rapport signal/bruit, de 96 dB ($16 \cdot 20 \cdot \log(2)$).

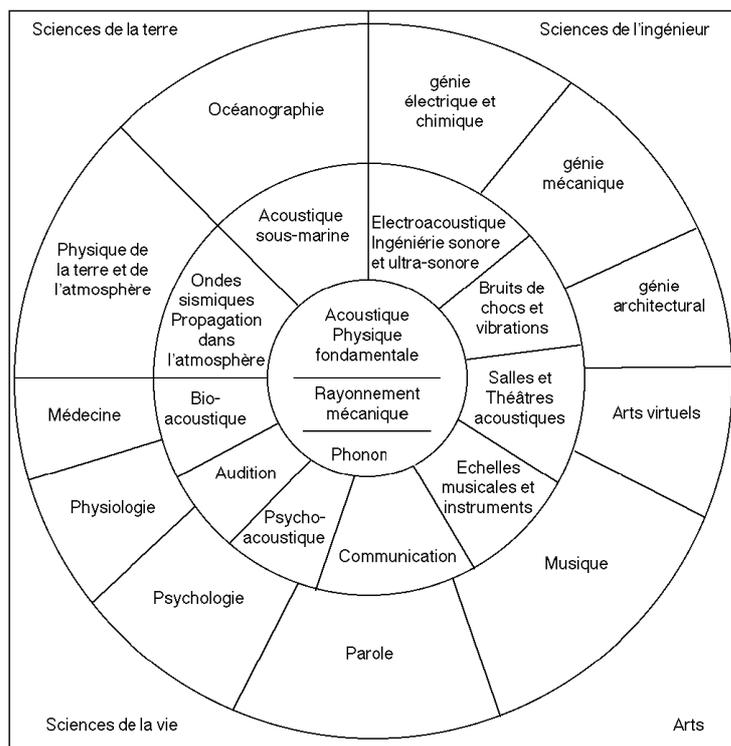


figure I-1 : disciplines associées à l'acoustique. Le premier anneau à partir du centre contient les subdivisions traditionnelles de l'acoustique et l'anneau extérieur les noms des champs artistiques et techniques qu'elle couvre (d'après [Bruneau-1983, p. 41]).

I.2.1.1. La production du son

Les recherches en acoustique ont souvent porté sur l'étude des sons musicaux. Ces recherches existaient déjà chez les philosophes grecs, ainsi au VI^e siècle avant J.—C., Pythagore avait déjà mis en évidence la relation de hauteur d'une octave existant entre les notes émises par deux cordes tendues quand l'une était deux fois plus longue que l'autre.

Galilée² est l'un des premiers à avoir déterminé les principes généraux de l'acoustique. Il a décrit les lois du mouvement pendulaire ainsi que les phénomènes de vibrations par sympathie. Il a également décrit la relation entre la hauteur d'un son produit et la longueur de la corde en vibration en énonçant que la hauteur correspondait à un nombre donné de vibrations par unité de temps.

C'est à son contemporain, le moine franciscain Marin Mersenne que l'on doit le traité « *Harmonicorum Liber* » [Mersenne-1636] qui contient l'ensemble des connaissances en acoustique de son époque. Cet ouvrage publié en 1636 donne des informations précises sur la mesure de la fréquence de vibration des cordes et la première détermination directe de la fréquence d'un son musical.

Un peu plus tard Robert Hooke a établi une correspondance entre la fréquence et la hauteur d'un son en utilisant un dispositif mécanique basé sur des roues dentées.

L'intérêt suscité par cette nouvelle discipline amena Joseph Sauveur en 1701 à créer le terme d'« *acoustique* » pour désigner la science de l'étude des sons. On lui doit également les termes de « fréquence fondamentale » et de « son harmonique ». Il a indiqué le premier que le timbre d'un son est déterminé par la superposition de ses harmoniques.

² Galilée (1564-1642), Marin Mersenne (1588-1648), Robert Hooke (1635-1703), Joseph Sauveur (1653-1716), Brook Taylor (1685-1731), Daniel Bernouilli (1700-1782), Jean d'Alembert (1717-1783), Léonhard Euler (1707-1783), Joseph Fourier (1768-1830), Louis de Lagrange (1736-1813), E. F. F. Chladni (1756-1824), G. Robert Kirchhoff (1824-1887), S. Denis Poisson (1781-1840), Clebsch (1833-1872), Hermann von Helmholtz (1821-1894), J. William Strutt Rayleigh (1842-1919).

Les découvertes des savants de l'époque moderne sont à l'origine de nos connaissances en acoustique. Ainsi, Brook Taylor a établi les équations des courbes de vibration des cordes. Ses recherches furent complétées par celles de Daniel Bernouilli, de Jean d'Alembert et de Léonhard Euler. L'introduction des dérivées partielles a conduit à la représentation de l'équation du mouvement sous sa forme actuelle. Daniel Bernouilli montra également qu'une corde en vibration produit simultanément une multitude d'oscillations élémentaires dont la somme donne le son résultant, c'est le « principe de superposition ».

Les travaux du mathématicien Joseph Fourier apportèrent à l'acoustique un théorème fondamental sur lequel sont encore construits la plupart des outils d'analyse du son (cf. Annexe I).

Louis de Lagrange résolut de façon complète et analytique l'équation de la vibration d'une corde. Cette équation, assimilant la corde à un ensemble de particules identiques et réduisant la vibration à une somme de vibrations élémentaires, a permis le calcul des fréquences harmoniques émises par une corde tendue. Louis de Lagrange s'intéressa également aux sons produits par les instruments à vent pour lesquels il établit approximativement les fréquences des harmoniques.

Les études des vibrations des corps solides ne purent se développer qu'avec la mise au point de la théorie mathématique des vibrations élastiques entreprise par Robert Hooke entre 1660 et 1676. Les équations de la vibration des barres furent publiées en 1744 par Léonhard Euler et en 1751 par Daniel Bernouilli.

Les mécanismes de vibration de plaques solides élastiques furent découverts par E. F. F. Chladni, l'un des plus illustres acousticiens expérimentaux. C'est à lui que l'on doit les célèbres figures de vibrations des plaques recouvertes de sable fin mettant en évidence les positions des lignes nodales. En 1850, G. Robert Kirchhoff avança une théorie mathématique décrivant les vibrations de ces plaques.

Le problème des membranes souples fut résolu par S. Denis Poisson à l'exception de celui des membranes circulaires dont les équations furent données par Clebsch.

Toutes ces recherches furent réalisées en confrontant des études expérimentales et théoriques. L'étude des instruments de musique s'est généralement développée en marge des recherches en acoustique et ce n'est que très tardivement que les deux domaines ont commencé à s'influencer de façon notable.

Au milieu du XIX^e siècle, Hermann von Helmholtz [Helmholtz-1868/1990] et J. William Strutt Rayleigh [Rayleigh-1945] établirent les fondements théoriques de l'acoustique moderne. J. William Strutt Rayleigh montra comment simuler des phénomènes acoustiques en utilisant par exemple des circuits électriques pour simuler les vibrations contraintes dans les résonateurs acoustiques. C'est lui qui proposa l'utilisation de cornes coniques pour la transmission et la réception de messages de signalisation. Il mit aussi en évidence les effets de masquage d'une sphère sur une radiation ainsi que l'effet binaural pour la perception du son. A cette période, la caractérisation du son repose sur sa phase stationnaire, bien qu'Helmholtz signale déjà le rôle des facteurs temporels dans la définition du son.

Au XX^e siècle, des progrès manifestes ont été effectués grâce à l'utilisation de nouveaux équipements électroniques, acoustiques et optiques puis avec l'apparition des enregistreurs magnétiques, des analyseurs de spectres, des lasers. Stumpf mis ainsi en évidence l'importance des transitoires d'attaque pour la reconnaissance du timbre.

C'est surtout avec l'avènement de l'informatique que les chercheurs ont pu commencer à étudier les instruments de musique en réalisant des observations fines sur leur comportement acoustique ainsi qu'en effectuant des modélisations de

leur fonctionnement grâce à la synthèse sonore sur ordinateur. Les travaux de Max Mathews [Mathews-1969] et John Pierce [Pierce-1984], puis ceux de Jean-Claude Risset [Risset-1986], John Chowning, David Wessel [Wessel-1979] ou Stephen MacAdams [McAdams-1986] ont permis d'aborder le problème du timbre dans toute sa complexité, en le reliant à notre perception. Leurs travaux seront développés dans le deuxième chapitre de cet ouvrage.

I.2.1.2. La propagation du son

La notion d'onde peut être définie comme une perturbation oscillatoire qui se propage à partir d'une source.

C'est Robert Boyle³ qui a admis le premier que l'air était un milieu de transmission du son, en observant les diminutions d'intensité d'un son émis dans une atmosphère progressivement raréfiée.

La vitesse de propagation du son dans l'air fut mesurée successivement par Gassendi - 477 m/s -, par Marin Mersenne - 450 m/s - puis par Borelli et Viviani - 350 m/s -. La valeur actuellement admise est de 331,36 + 0.08 m/s à une température de 0°C et à une pression de 760 mm de mercure.

En 1740, l'italien Branconi montra expérimentalement que la vitesse du son augmente avec la température.

En 1808, le physicien français Jean-Baptiste Biot démontra que la vitesse de propagation du son dans le fer est bien plus élevée que dans l'air. Charles Sturm et J. D. Colladon estimèrent la vitesse du son dans l'eau à 1435 m/s à une température de 8°C.

La vitesse de transmission du son dans un milieu homogène et à température constante est donnée par la formule de Laplace :

$$v = \sqrt{\frac{PC}{\rho c}}$$

P correspond à la pression de l'air, ρ est la masse spécifique du milieu de propagation, C est la chaleur spécifique à pression constante et c est la chaleur spécifique à volume constant.

milieu de propagation	température	vitesse du son
air	0°C	331 m/s
air	15°C	340 m/s
eau	13°C	1428 m/s
fer	20°C	4887 m/s
bois de sapin	non spécifiée	6000 m/s

tableau I-1: vitesse de propagation du son selon le milieu.

Les premières théories sur la propagation des ondes sonores dans l'air furent émises par Isaac Newton à partir de 1687. Des améliorations de sa théorie furent réalisées par Pierre S. de Laplace vers 1816.

À la fin du XVIII^e siècle, les lois de la propagation des ondes dans un tube furent établies, ce qui conduisit à la détermination des harmoniques correspondants.

C'est S. Denis Poisson qui en 1820 indiqua comment se propage une onde dans un milieu fluide à trois dimensions. Il mit au point la théorie de la propagation des ondes en développant le cas des ondes stationnaires observées dans un tube ouvert ou fermé. Vers 1860, Hermann von Helmholtz apporta également une grande contribution à l'étude de ces phénomènes.

³ Robert Boyle (1627-1691), Pierre Gassend (1592-1655), Borelli (1608-1679), Viviani (1622-1703), Charles Sturm (1803-1855), J. D. Colladon (1802-1893), Isaac Newton (1642-1727), Pierre S. de Laplace (1749-1827).

Les problèmes de réflexion et de réfraction du son furent résolus vers 1838 par George Green. Cela permit d'établir les correspondances et les oppositions entre les phénomènes optiques et acoustiques. Ces recherches conduisirent aux applications actuelles de l'acoustique en géophysique pour l'étude des ondes sismiques et de la structure des différentes zones de l'écorce terrestre.

I.2.1.3. La réception du son

Jusqu'à la fin du XIX^e siècle, la plupart des recherches en acoustique étaient effectuées en utilisant l'oreille comme récepteur du son. En 1830, Félix Savart donna comme limites de l'audition humaine les valeurs de 8 vibrations par seconde vers le grave et de 24000 vibrations par seconde vers l'aigu. Actuellement on les situe respectivement à 20 Hz et 20000 Hz.

Le seuil d'audibilité en intensité fut étudié par Toepler, Botzmann et J. William Strutt Rayleigh.

En 1843, Georg S. Ohm indiqua que la qualité d'un son musical est due à la superposition de sons purs de différentes fréquences que l'oreille est capable de percevoir. Hermann von Helmholtz le suivit dans cette voie et donna la première théorie du fonctionnement de l'oreille, appelée « théorie de la résonance ».

Les premières considérations concernant l'acoustique des salles intervinrent en 1853 lors de la parution des articles de J. B. Upham définissant les phénomènes de réverbération et de réflexion du son à la surface des parois d'une pièce. En 1856, Joseph Henry se mit à étudier l'acoustique des auditoriums. La fondation de l'acoustique architecturale est attribuée à C. Sabine en 1900.

Parmi les tentatives d'amplification du son, on peut citer l'utilisation des trompes qui sont connues depuis l'antiquité⁴. En 1650, Kircher proposa l'utilisation de trompes paraboliques pour l'aide à l'audition. René Laënnec inventa en 1819 le stéthoscope. Le terme « microphone » fut inventé par Charles Wheatstone en 1827 pour désigner un instrument similaire au stéthoscope.

Comme pour la plupart des phénomènes étudiés en physique, les chercheurs en acoustique ont tenté de mettre au point des techniques permettant de donner une représentation visuelle des phénomènes.

En 1858, John LeConte observa les variations d'une flamme sous l'influence du son. John Tyndall développa le même principe pour l'observation des sons à hautes fréquences et de leur réflexion, réfraction et diffraction. Le microphone électrique supplanta ce système.

Divers dispositifs furent employés pour la visualisation des phénomènes sonores dont le « phonautographe » mis au point par Leon Scott. Cet instrument, précurseur du phonographe, permettait de tracer une courbe décrivant l'onde sonore à l'aide d'un stylet.

La possibilité offerte par l'enregistrement du son a révolutionné les sciences du son avec l'apparition du phonographe d'Edison en 1875. Le son peut ainsi être fixé sur un support, il n'est plus un phénomène acoustique fugitif, il est reproductible à loisir. L'année suivante, c'est Bell qui invente le son électrique avec le téléphone en 1876 qui conduira au traitement du signal sonore et aux transmissions à distance.

L'électronique a apporté de nouveaux instruments de mesure. Pour connaître les partiels présents dans un son, on peut capter ce son à l'aide d'un microphone puis analyser le signal électrique ainsi produit grâce à des filtres électroniques. Un filtre peut être décrit par une courbe qui montre comment le rapport entre l'amplitude de sortie et l'amplitude d'entrée varie en fonction de la fréquence. L'analyseur de spectre est un appareil utilisant un ensemble de filtres. Il permet de

⁴ Ce système était utilisé dans les théâtres.

visualiser sur un écran la courbe donnant les intensités d'un ensemble de fréquences. Cette courbe est appelée **spectre de raies**.

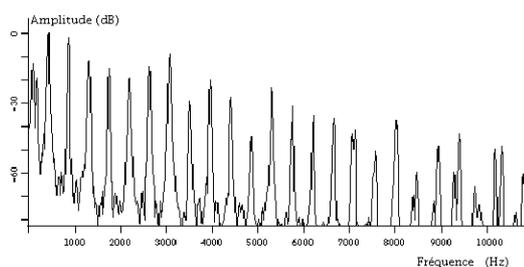


figure I-2 : spectre de raies d'un son de guitare.

L'ordinateur a également permis de produire des spectres de raies sur des sons numériques. Les variations d'intensité des différents partiels sont mis en évidence sur des représentations en trois dimensions qui incluent une dimension temporelle.

Le **sonographe** est certainement l'outil qui permet la meilleure représentation du son en trois dimensions — **fréquence**, **amplitude** et **temps**. Le sonogramme manque de précision en ce qui concerne les valeurs d'amplitude des partiels rendues par les niveaux de gris ou par l'épaisseur du trait, mais cette représentation sur deux axes — temps et fréquence — permet une lecture assez instinctive, proche de la perception. Elle comporte des analogies avec la partition musicale traditionnelle. Nous avons employé à de nombreuses reprises le sonogramme dans cet ouvrage afin de mieux décrire les sons cités dans le texte.

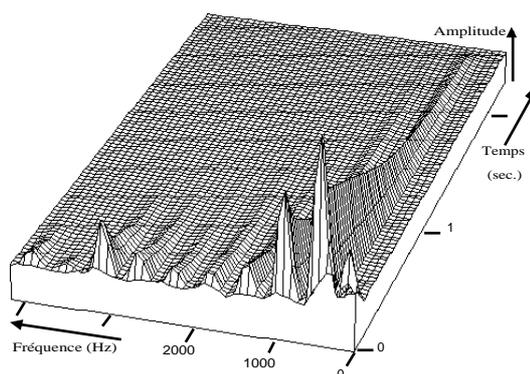


figure I-3 : spectre à trois dimensions d'un son de guitare.

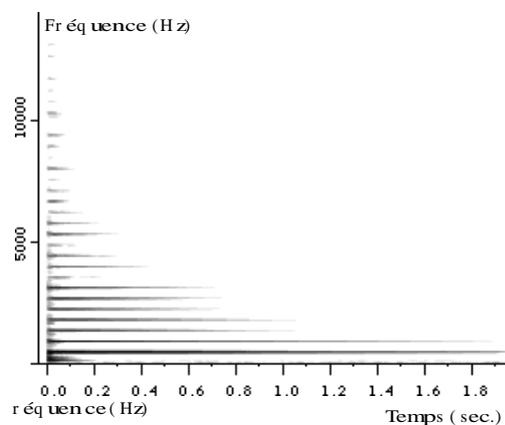


figure I-4 : sonogramme d'un son de guitare (exemple sonore n°2).

I.2.2. Les mécanismes mis au cause pour la production du son

Un des dispositifs les plus simples de production de son est le diapason, formé de deux barres reliées à leur base. Soumis à une excitation, il délivre une fréquence pure et stable, après s'être stabilisé en quelques instants.

Le spectre initial du diapason comporte quelques partiels transitoires de fréquences élevées qui disparaissent rapidement en ne laissant qu'une fréquence partielle à 440 Hz : le « La » du diapason.⁵

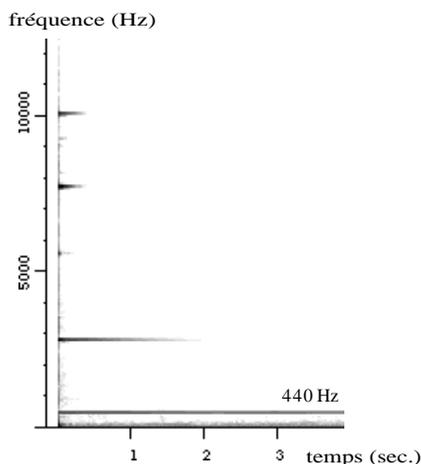


figure I-5 : sonagramme d'un son de diapason (exemple sonore n°1).

Un diapason est un quadripôle. Ses deux barres se déplacent simultanément vers l'extérieur puis vers l'intérieur en comprimant puis détendant l'air dans deux zones situées dans l'axe des branches du diapason. Deux autres zones, situées dans un axe orthogonal à l'axe des branches du diapason, subissent des compressions et dépressions inverses à celles des deux premières zones.

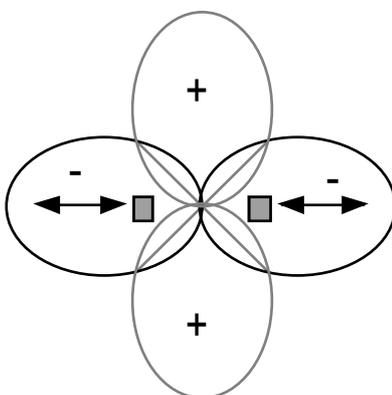


figure I-6 : fronts d'ondes du diapason.

L'oscillation du diapason s'atténue progressivement avec le temps par rayonnement et par pertes thermiques. Si on mesure l'intensité de la vibration en faisant tourner le diapason, on rencontre des zones d'intensité nulle à l'intersection des zones positives et négatives représentées sur la figure I-6.

⁵ Arthur Benade note que lorsque le diapason est frappé en un point situé entre 1/2 et 1/4 vers le bas des branches, il donne un son clair de hauteur très nette. Par contre, si le diapason est frappé à l'extrémité d'une de ses branches, on entend simultanément un autre partiel de fréquence environ 3000 Hz [Benade-1976, p. 38].

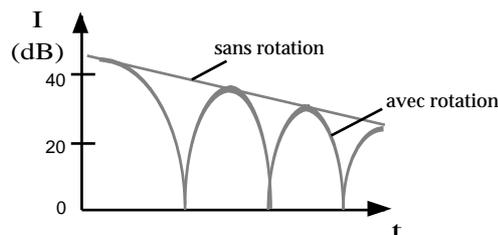


figure I-7 : variation de l'intensité du son émis par un diapason en fonction du temps.

I.2.2.1. L'oscillateur harmonique

Un **oscillateur**, système vibrant avec un mouvement cyclique régulier, peut être réalisé par trois types de systèmes qui comportent des analogies très fortes. Cela permet d'utiliser les uns pour simuler les autres et facilite la compréhension des lois auxquels ils sont soumis. Il s'agit des circuits électriques, mécaniques et acoustiques.

Nous commencerons par étudier les systèmes oscillants ne comportant qu'un degré de liberté, pour lesquels un seul paramètre permet de caractériser à n'importe quel instant la configuration du système.

A) Les systèmes mécaniques

Le système masse-ressort

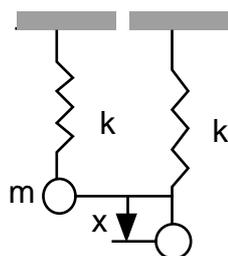


figure I-8 : déplacement d'un ressort soumis à l'action d'une masse m.

Le ressort est caractérisé par son coefficient de raideur k . Le solide est défini par sa masse m et est censé se déplacer sans frottement.

Dans l'hypothèse où la masse est ponctuelle et où le ressort est parfaitement élastique, sans masse et sans hystérésis, la force de rappel est constamment proportionnelle à l'élongation. Dans ce cas, l'équation dynamique du système prend la forme :

$$m \frac{d^2 x}{dt^2} + kx = 0$$

La solution de cette équation peut s'écrire :

$$x = x_0 \sin(\omega_0 t + \varphi) \quad \text{avec } \omega_0 = \sqrt{\frac{k}{m}}$$

La position verticale de la masse m par rapport au ressort au repos est X_0 telle que :

$$mg = kX_0$$

donc la pulsation de la vibration ω_0 vaut :

$$\omega_0 = \sqrt{\frac{k}{m}} = \sqrt{\frac{g}{X_0}}$$

Le pendule

Un pendule correspond à un corps de masse m suspendu à une corde de longueur l et de masse négligeable. La force de rappel du pendule, écarté de sa position d'origine d'un angle θ , a une direction tangente à la trajectoire et a pour module :

$$f_r = mg \cdot l \cdot \sin\theta$$

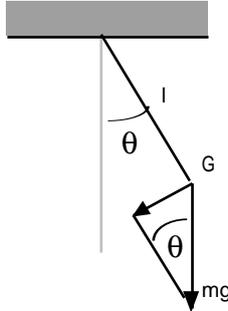


figure I-9 : déplacement d'un pendule écarté de sa position d'origine.

L'équation de la dynamique, appliquée aux moments des forces s'exerçant au centre de gravité G , est de la forme:

$$J \frac{d^2\theta}{dt^2} + mg \cdot l \cdot \sin\theta = 0 \quad J \text{ étant le moment d'inertie} \quad (1)$$

La fonction sinus peut se développer en une série de Taylor :

$$\sin\theta = \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \dots$$

La relation (1) se simplifie dans le cas où l'on ne considère que des mouvements de faible amplitude, pour lesquels $\sin\theta$ est équivalent à θ :

$$\frac{d^2\theta}{dt^2} + \left(\frac{mg \cdot l}{J}\right)\theta = 0$$

la solution peut s'écrire :

$$\theta = \theta_0 \sin(\omega_0 t + \varphi)$$

La pulsation de la vibration ω_0 vaut :

$$\omega_0 = \sqrt{\frac{m \cdot g \cdot l}{J}}$$

Dans le cas simple où la masse est concentrée sur son centre de gravité G , le moment d'inertie se réduit à $J = ml^2$ et la pulsation propre de ce pendule devient :

$$\omega_0 = \sqrt{\frac{g}{l}}$$

Dans le cas de mouvements de plus grande amplitude, les termes $\theta^3/3! + \theta^5/5! \dots$ introduisent un comportement non linéaire.

Pour les instruments de musique, comme la contrebasse ou le piano, les cordes en vibration présentent un comportement linéaire pour de faibles dynamiques de jeu et un régime non linéaire pour les fortes dynamiques.

B) Les systèmes acoustiques

Le résonateur d'Helmholtz

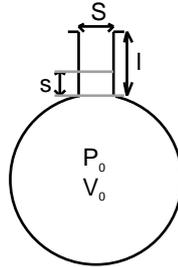


figure I-10 : résonateur d'Helmholtz.

Ce résonateur se comporte comme le système masse-ressort. La masse est celle du fluide contenu dans le goulot et la compression de l'air contenu dans le résonateur correspond au ressort.

La force de rappel due à la compression ou à la dépression de l'air est donnée par l'équation de variation adiabatique d'état du fluide. Cette équation est appelée équation de Laplace et s'exprime sous la forme $PV^\gamma = \text{constante}$, dans laquelle γ est le rapport des chaleurs massiques du fluide à pression constante et à volume constant.

L'équation fondamentale de la dynamique appliquée au mouvement de la masse du fluide contenu dans le goulot ($m = \rho_0 l S$) s'écrit :

$$\rho_0 l S \cdot \frac{d^2 s}{dt^2} + \left(\frac{\gamma P_0 S^2}{V_0} \right) \cdot s = 0 \quad \rho_0 \text{ est la masse volumique du fluide}$$

dont la solution peut s'écrire :

$$s = s_0 \sin(\omega_0 t + \varphi)$$

Le mouvement est donc une vibration de pulsation ω_0

$$\omega_0 = \sqrt{\frac{\gamma P_0 S}{\rho_0 l V_0}}$$

C) Les systèmes électriques

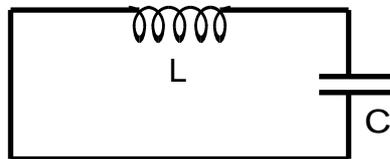


figure I-11 : circuit électrique comportant une self-inductance et une capacité.

Lorsqu'on met en contact une capacité chargée avec une inductance non résistive, le circuit ainsi formé entre dans un processus d'échange alternatif.

L'équation électrique du système s'écrit :

$$\frac{d^2 i}{dt^2} + \frac{i}{LC} = 0 \Rightarrow i = i_0 \sin(\omega_0 t + \varphi) \quad \text{avec} \quad \omega_0 = \frac{1}{\sqrt{LC}}$$

i est l'intensité du courant parcourant le circuit, L est l'inductance et C la capacité.

D) L'oscillateur harmonique en régime libre amorti

Les systèmes que nous venons de décrire — le système masse-ressort, le pendule, le résonateur et le circuit LC — correspondent à des oscillateurs harmoniques dont l'équation du mouvement peut s'écrire :

$$\psi(t) = A \cos(\omega t + \varphi)$$

La fréquence de l'oscillation est $f = \frac{2\Pi}{\omega}$ si t est exprimé en secondes.

Tous ces systèmes comportent une force de rappel et une force d'inertie.

La notion de système libre oscillant sans perte est une notion exclusivement théorique. En pratique, tous les systèmes dissipent de l'énergie et ne peuvent pas, de ce fait, être décrits par des fonctions sinusoïdales.

Les équations deviennent alors souvent :

$$\psi(t) = A e^{-t/2\tau} \cdot \cos(\omega t + \varphi)$$

dans lesquelles τ représente le facteur d'amortissement.

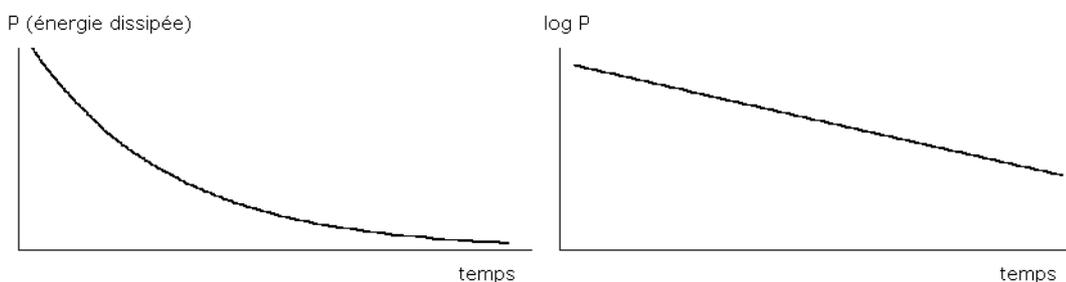


figure I-12 : décroissance de l'énergie dissipée par une vibration amortie.

Les enveloppes d'amplitude d'un son de marimba ou d'un son de guitare sont des exemples caractéristiques de décroissances exponentielles.

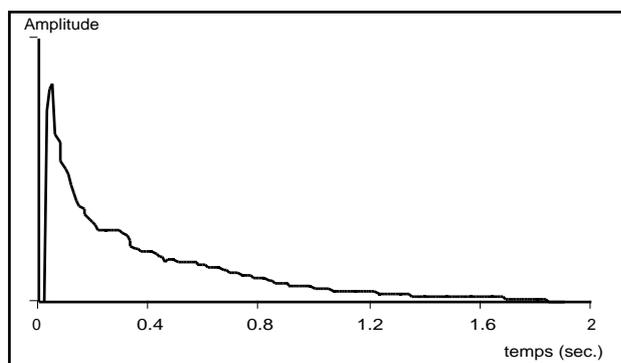


figure I-13: enveloppe d'amplitude d'une note de guitare (exemple sonore n°2).

Les deux types d'amortissements rencontrés le plus fréquemment sont l'amortissement solide et l'amortissement visqueux.

Amortissement solide

Dans le cas de l'amortissement solide, correspondant par exemple au frottement d'une masse mobile sur une surface solide fixe, la force de frottement est une constante égale à f_s , indépendante de la vitesse du mouvement.

L'équation des forces appliquées à la masse s'écrit :

$$m \frac{d^2x}{dt^2} + kx \pm f_s = 0$$

On obtient une solution donnant un **amortissement linéaire** du signal en fonction du temps :

$$x = A \cdot (-\alpha t) \cdot \sin(\omega t + \varphi)$$

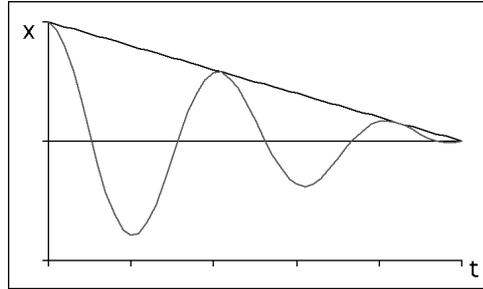


figure I-14 : amortissement linéaire d'une oscillation.

Amortissement visqueux

Dans le cas d'un amortissement dû à un frottement dans l'air, la force de frottement F_f devient proportionnelle à la vitesse de déplacement de la masse mobile.

$$F_f = c_f \frac{dx}{dt}$$

c_f est le **coefficient de frottement** dû à la présence du fluide.

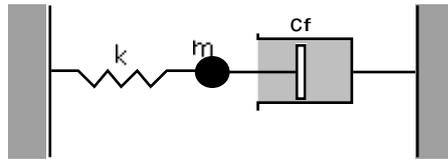


figure I-15 : système mécanique en régime libre comportant un ressort, une masse et un dispositif d'amortissement.

L'équation de la dynamique peut s'écrire :

$$m \frac{d^2x}{dt^2} + F_f \frac{dx}{dt} + kx = 0$$

On obtient une solution donnant un **amortissement exponentiel** du signal en fonction du temps :

$$x = A \cdot e^{-\lambda t} \cdot \sin(\omega t + \varphi)$$

λ est le **coefficient d'amortissement** du système et ω est sa **pseudo-pulsation**.

$$\lambda = \frac{c_f}{2m}$$

$$\omega = \sqrt{\omega_0^2 - \lambda^2}$$

Pour la gamme des fréquences audibles, λ est généralement très faible devant ω et on peut considérer dans de nombreux cas que la pseudo-fréquence de l'oscillation est égale à sa fréquence propre.

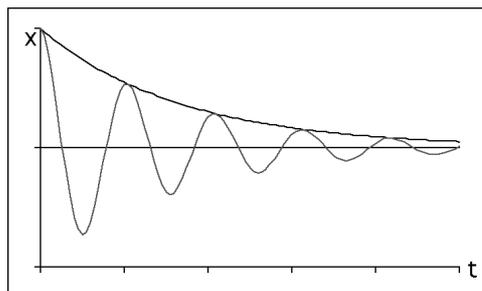


figure I-16 : amortissement exponentiel d'une oscillation.

Ce type d'amortissement est le plus couramment rencontré dans le cas des sons percussifs, soumis à une excitation puis qui résonnent librement.

E) Récapitulatif

Les analogies formelles existant entre les quatre systèmes oscillants que nous venons de décrire sont présentées dans le tableau suivant.

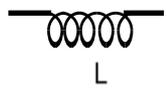
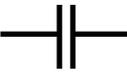
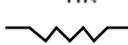
 L	 masse	 moment d'inertie	 masse acoustique	mouvement
 R	 résistance mécanique	 résistance mécanique de rotation	 résistance acoustique	amortissement
 C	 compliance mécanique	 compliance mécanique de rotation	 compliance acoustique	force de rappel
électricité	mécanique	rotation	acoustique	

tableau I-2 : analogies formelles existant entre divers types de systèmes oscillants.

F) L'oscillateur harmonique en régime forcé

Les systèmes vibrants sont parfois soumis à une force extérieure permanente — instruments à vent, cordes frottées.

Lorsque cette force est périodique, on observe généralement, après une **période transitoire** d'adaptation du système à l'excitation, un mouvement régulier et stable qui constitue le **régime permanent**. Le **résonateur** mécanique soumis à cette excitation harmonique oscille alors à la même fréquence que l'**excitateur** et fonctionne en **régime forcé**.

Le mouvement peut être décrit par l'équation :

$$m \frac{d^2x}{dt^2} + F_f \frac{dx}{dt} + kx = F_m \cos(\Omega t + \varphi) \quad (2)$$

F_m est l'amplitude maximum de la force et Ω sa pulsation.

Dans ces systèmes, l'**excitateur** finit par imposer la **fréquence** tandis que le **résonateur** impose l'**amplitude**.

La **fréquence réduite** de l'oscillation est notée α , l'**amplitude réduite** est notée Δ et l'**amortissement réduit** est noté λ_0 :

$$\alpha = \frac{f_e}{f_r} ; \Delta = \frac{A_r \alpha}{A_{r0}} ; \lambda_0 = \frac{\lambda_r}{\omega}$$

avec f_e la fréquence de l'excitation, f_r la fréquence propre du résonateur, $A_{r\alpha}$ l'amplitude du résonateur à la fréquence réduite α , A_{r0} l'amplitude du résonateur pour $\alpha=0$, λ_r le coefficient d'amortissement du résonateur et ω sa pulsation propre.

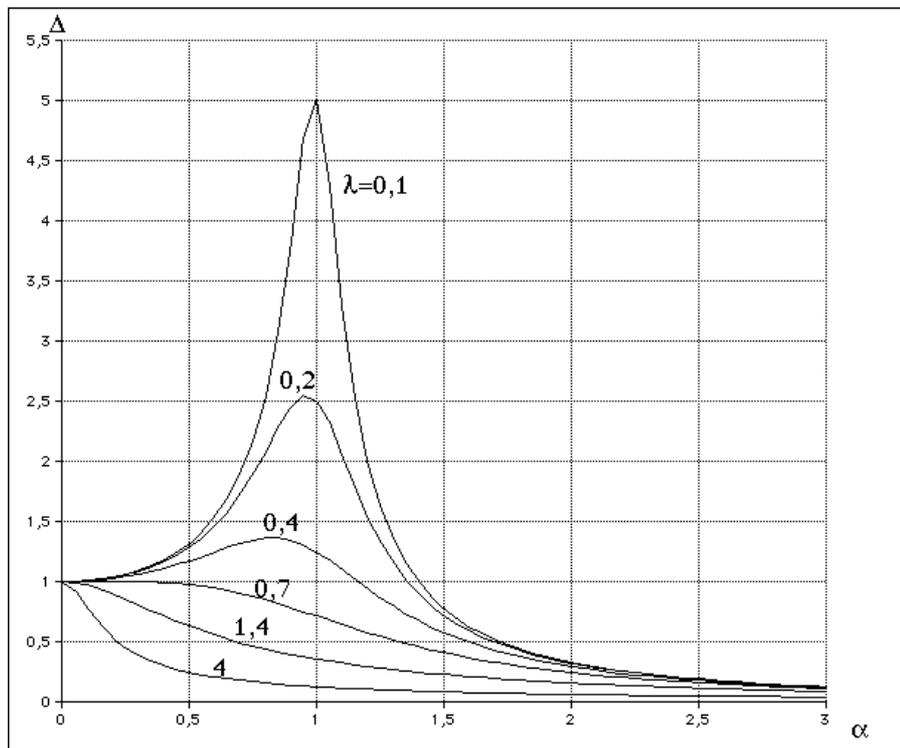


figure I-17 : valeur de l'amplitude relative d'un mouvement oscillant en fonction de la fréquence réduite de l'excitation (d'après [Jouhaneau-1994]).

Dans la figure précédente, lorsque la fréquence de l'excitation est grande par rapport à la fréquence propre du résonateur, l'amplitude du déplacement est faible. L'amplitude est maximale lorsque les deux fréquences sont égales.

La solution de l'équation (2) est de la forme :

$$x = x_0 \cdot e^{-\lambda t} \cos(\omega t + \mu) + X_m \cos(\Omega t + \Phi)$$

avec

$$X_m = \frac{F_m}{\sqrt{(F_f \cdot \Omega)^2 + (m \cdot \Omega^2 - k^2)^2}}$$

Dans la figure suivante, on voit comment l'excitateur impose la fréquence au résonateur.

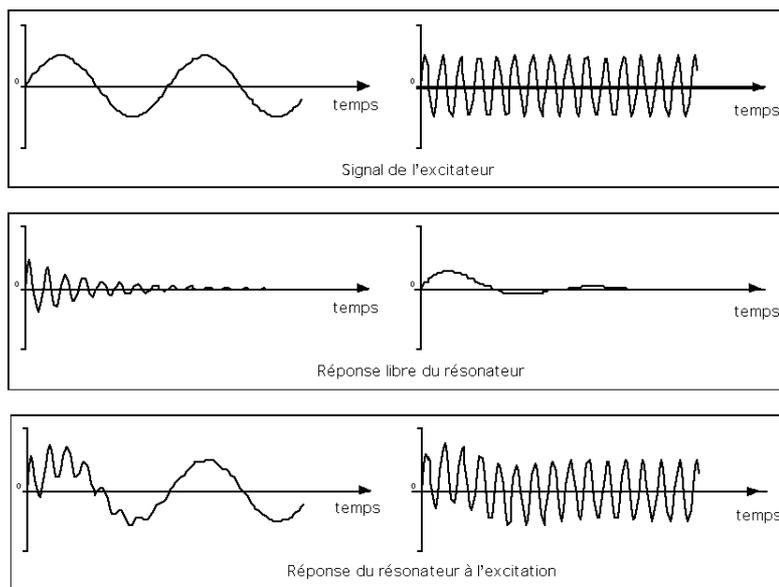


figure I-18 : réponse d'un résonateur à une excitation périodique. A gauche, la fréquence propre du résonateur est supérieure à celle de l'excitateur, à droite c'est l'inverse.

I.2.2.2. Modèles vibratoires

Jusqu'à présent, nous n'avons présenté que des systèmes élémentaires ne comportant qu'un seul degré de liberté. Les instruments de musique sont formés d'éléments pouvant vibrer de façon beaucoup plus complexe. Ceci dit, l'instrument peut être décomposé comme une somme d'éléments vibrant de façon élémentaires couplés entre eux.

On rencontre avec les instruments de musique plusieurs modèles vibratoires que l'on peut classer en cinq catégories. La **corde** représente une extension du modèle « masse-ressort » lorsque le nombre de masses reliées entre elles devient très important. Le deuxième type correspond à la **membrane** que l'on peut assimiler à une corde à deux dimensions : on passe de la ligne à la surface. Dans ces deux cas, c'est la tension qui détermine la fréquence de la vibration. On trouve ensuite des systèmes rigides comme les **barres**, dont l'épaisseur et la rigidité assurent une force de rappel, ou les **plaques**. Enfin, le dernier type correspond aux instruments pour lesquels c'est la **colonne d'air** qui est directement mise en vibration.

A) Systèmes à plusieurs degrés de liberté

Les systèmes qui nécessitent deux coordonnées ou plus pour décrire leur mouvement sont à plusieurs degrés de liberté.

Par exemple, dans le cas de deux masses reliées par des ressorts, la nature des oscillations va dépendre d'un coefficient de couplage χ .

$$\chi = \frac{K}{\sqrt{(k_1 - K)(k_2 + K)}}$$

Suivant les dispositifs, les couplages peuvent être réalisés par un ressort, par frottement, par inertie ou par une combinaison des trois.

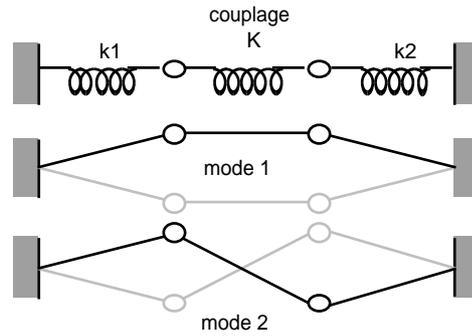


figure I-19 : couplage de deux masses par l'intermédiaire d'un ressort et modes de vibration transversale associés.

Dans le cas général, les équations du mouvement peuvent être écrites à partir de la connaissance des énergies mises en jeu et du travail fourni au système.

On peut utiliser la fonction $L = E_c - E_p$, différence de l'énergie cinétique et de l'énergie potentielle. Si le vecteur $r_i = r_i(q_1, q_2, q_3 \dots q_n)$ décrit le déplacement de chaque élément i d'un système à n degrés de liberté, on peut calculer les énergies :

$$E_c = \sum_i \frac{1}{2} m_i \dot{r}_i^2$$

$$E_p = \sum_i \frac{1}{2} k_i r_i^2$$

d'où l'équation de Lagrange pour un système conservatif :

$$(d/dt) \delta L / \delta \dot{q}_i - \delta L / \delta q_i = 0$$

B) Vibrations transversales des cordes

Une masse attachée par deux ressorts à des supports immobiles peut vibrer selon différents axes.

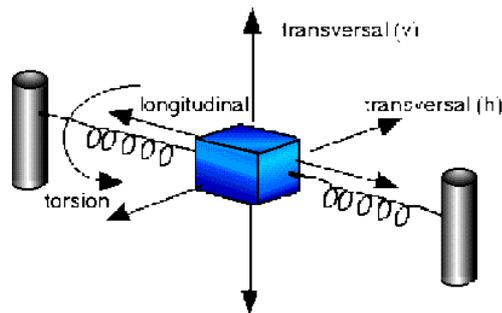


figure I-20 : principaux axes de vibration d'une masse reliée par deux ressorts à un support.

En multipliant les masses, on peut modéliser une corde⁶.

Pour une corde tendue, la vitesse v des ondes est donnée par l'expression :

$$v = \sqrt{\frac{T}{M}}$$

T est la tension de la corde et M sa masse par unité de longueur.

⁶ C'est en utilisant ce type de méthode que les chercheurs du centre de l'ACROE à Grenoble ont construit leur programme de synthèse par modélisation d'objets physiques à partir de composants élémentaires " Cordis-Anima " (cf. [Cadoz, Luciani, Florens-1990]).

A chaque vibration d'une corde tendue, une onde transversale parcourt la corde deux fois, une fois dans chaque sens. Pour une corde de longueur L , la fréquence fondamentale de la vibration qui en résulte peut s'écrire :

$$f = \frac{v}{2L} = \frac{\sqrt{T/M}}{2L}$$

C) Vibrations longitudinales des cordes

La relation entre la force longitudinale appliquée sur une barre de matériau homogène et sa déformation est donnée par la loi de Hooke :

$$f = -S \cdot E \cdot \frac{\delta s}{\delta x}$$

S est la section du matériau, E le module d'élasticité ou module de Young et s le déplacement d'une portion de la corde.

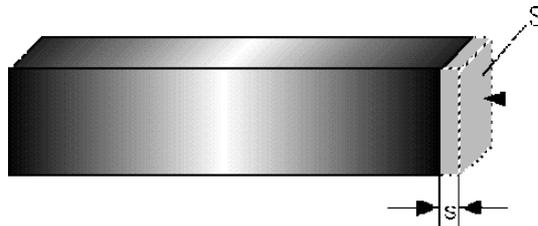


figure I-21 : compression longitudinale d'une barre à section rectangulaire

D) Vibrations de torsion des cordes

Les vibrations de torsion sont régies par des formules similaires à celles utilisées pour les vibrations longitudinales. Si l'on tord une barre avec un angle θ , l'équation du mouvement s'écrit :

$$\frac{\delta^2 \theta}{\delta t^2} = c^2 \frac{\delta^2 \theta}{\delta x^2} \text{ avec } c^2 = G/\rho$$

G est le module d'élasticité ou coefficient de rigidité transversale et ρ la masse volumique de la barre.

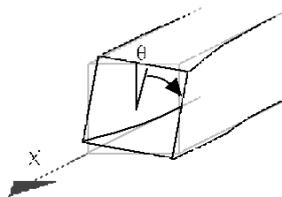


figure I-22 : torsion d'une barre de section rectangulaire.

E) Vibrations d'une corde pincée

Les vibrations d'une corde sont essentiellement transversales. Les différents modes de vibration d'une corde sont en relation harmonique :

$$f_n = n \cdot f_1 = n \cdot \frac{v}{2L}$$

Le mouvement de la corde est une combinaison de ces divers modes de vibration. Les amplitudes relatives de chaque mode définissent le spectre de la vibration et dépendent du mode d'excitation de la corde et en particulier de la position de l'excitation sur la corde.

En effet, si la corde est pincée en son centre, les harmoniques pairs sont inexistantes et l'amplitude des harmoniques impairs est inversement proportionnelle au carré de leur rang harmonique. Si l'amplitude de la

fondamentale est égale à 1, le troisième harmonique aura une amplitude de $1/9$, le cinquième harmonique une amplitude de $1/25$. Si par contre la corde est pincée vers une extrémité, par exemple à $1/5$ de sa longueur, les harmoniques de rang 5, 10 ... disparaissent mais toutes les autres harmoniques sont présents. Leurs amplitudes relatives décroissent moins vite.

Plus la corde est pincée près de l'extrémité, plus les harmoniques élevés sont intenses.

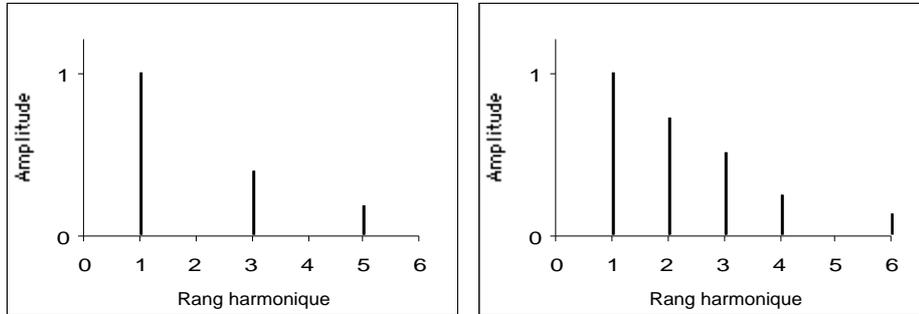


figure I-23 : spectre d'un son de corde pincée, à gauche jouée en son centre, à droite jouée à $1/5$ de sa longueur, amplitudes logarithmiques (d'après [Rossing—1990, pp. 189-190]).

F) Vibrations des membranes et des plaques

Alors qu'une corde est un système à une dimension, les membranes et les plaques sont des systèmes vibrants à deux dimensions. Au sens large, ce sont les peaux des tomes de percussions, les lames des percussions à claviers, les cloches ou même les éléments d'une caisse de résonance d'un violon.

Une membrane peut être représentée par un réseau de cordes orthogonales répondant aux mêmes conditions de raideur et de dissipation que la corde unidimensionnelle. Elle est caractérisée par ρ_s , sa masse par unité de surface (kg/m^2) et par τ , sa tension par unité de longueur (N/m).

Dans le cas d'une membrane rectangulaire fixée sur son pourtour, on peut connaître les fréquences propres du système pour différentes valeurs entières de m et n divisant respectivement la longueur et la largeur de la membrane :

$$f_{mn} = \frac{c}{2} \sqrt{\left(\frac{m}{X}\right)^2 + \left(\frac{n}{Y}\right)^2} \quad \text{avec } c = \sqrt{\frac{\tau}{\rho_s}}$$

X et Y sont les dimensions des côtés de la membrane.

Pour une membrane circulaire, les équations deviennent plus complexes. Le figure suivante donne les principaux modes de vibration d'une membrane circulaire idéale. Dans le cas de membranes réelles, les fréquences observées peuvent être différentes des fréquences théoriques. Cela est dû aux phénomènes de compression de l'air, à la résistance à la courbure et à la résistance au cisaillement.

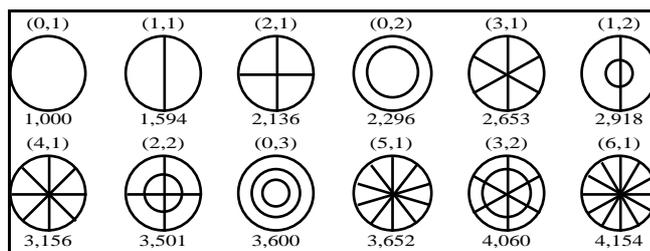


figure I-24 : principaux modes de vibration d'une membrane circulaire idéale. Pour obtenir les fréquences réelles, il faut multiplier les coefficients par $2,4/2\pi r \sqrt{\tau/\rho}$, r = rayon de la membrane, ρ , sa masse surfacique, τ , sa tension par unité de surface [Fletcher & Rossing-1998, p. 75].

Une plaque peut être interprétée comme une barre à deux dimensions ou comme une membrane épaisse. Elle peut transmettre des ondes de compression, de cisaillement et de torsion. Les modes et fréquences de vibration vont dépendre des conditions de fixation de la plaque : libre, fixée ou simplement posée. Nous citerons simplement les cas des plaques circulaires et rectangulaires.

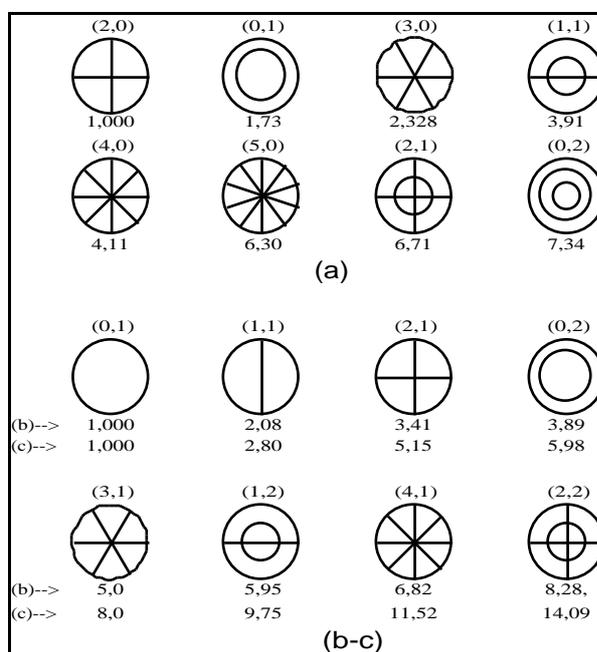


figure I-25 : principaux modes de vibration d'une plaque circulaire idéale. (a) : plaque à bords libres, (b) : plaque fixée, (c) : plaque posée [Fletcher & Rossing-1998, p. 79].

Pour les plaques circulaires, lorsque n et m augmentent, les fréquences peuvent être obtenues avec la loi de Chladni :

$$f = \alpha \cdot (m + 2n)^p$$

Pour des plaques plates, $p = 2$. Pour des cymbales par exemple, p est généralement inférieur à 2.⁷

Pour des plaques rectangulaires posées, si L_x et L_y sont les dimensions de la plaque et n et m des entiers supérieurs ou égaux à zéro,

$$f_{mn} = \alpha \cdot \left[\left(\frac{m+1}{L_x} \right)^2 + \left(\frac{n+1}{L_y} \right)^2 \right]$$

Pour des plaques libres ou fixées, les équations donnant les fréquences des différents modes de vibration sont complexes et dépendent des rapports de largeurs et de longueur des plaques.

⁷ Pour des excitations intenses, le comportement des cymbales ou des tamtam devient non linéaire et chaotique.

Mode	$L_x/L_y=1$	1,5	2	2,5	3	∞
(0,0)	1	0,75	0,68	0,66	0,64	0,62
(0,1)	2,04	1,88	1,82	1,79	1,78	1,72
(1,0)	2,04	1,16	1,16	0,88		
(1,1)	3,01	2,27	2,02	1,91	1,86	1,72

tableau I-3 : fréquences de vibrations relatives d'une plaque rectangulaire à extrémités fixés (d'après [Fletcher et Rossing-1998, p. 87]).

G) Vibrations des tubes et des tuyaux

Les instruments à vent se distinguent par la forme de leur résonateur et la nature de la anche utilisée. Les résonateurs sont généralement coniques ou cylindriques. Les vibrations sont longitudinales.

On distingue deux sortes de tuyaux sonores : les tuyaux à anche qui sont toujours ouverts à l'autre extrémité et les tuyaux à embouchure de flûte qui peuvent être ouverts ou fermés à l'autre extrémité. Il se produit toujours un nœud à une extrémité fermée et un ventre à une extrémité ouverte. Dans ce dernier cas, la pression de l'air à l'extrémité ouverte est égale en permanence à la pression atmosphérique

Un tuyau ouvert aux deux extrémités a une fréquence de vibration propre F inversement proportionnelle à la longueur L du tuyau.

$$F = c/2L \text{ ou } \lambda = 2L$$

λ est la longueur d'onde et c la célérité du son

Un tuyau fermé à une extrémité donne un son d'une octave plus bas qu'un tuyau ouvert de même longueur.

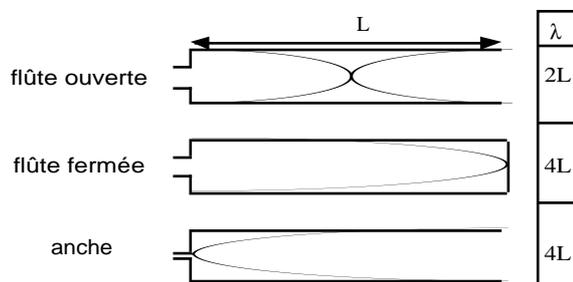


figure I-26 : répartition des nœuds et des ventres dans les tuyaux cylindriques.

L'impédance⁸ d'un tube cylindrique varie avec la fréquence et favorise les harmoniques impairs.

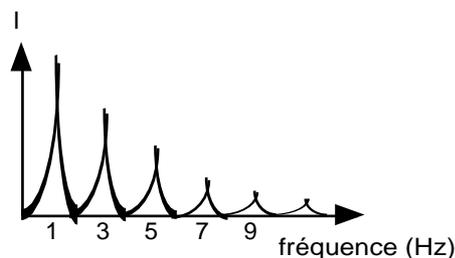


figure I-27 : impédance d'un tuyau cylindrique (d'après [Jouhaneau-1994]).

⁸ L'impédance Z (ohm) est définie par le rapport de la pression sonore sur le débit d'air.

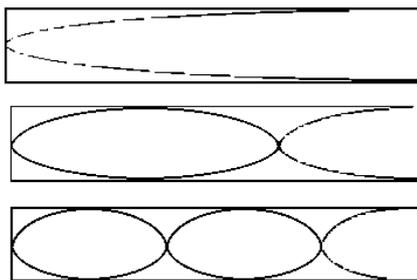


figure I-28 : répartition des nœuds et des ventres des harmoniques dans un tuyau cylindrique ouvert à une extrémité

Les fréquences de résonance du tube sont donc $c/4L$, $3c/4L$, $5c/4L$, etc.. Seules les harmoniques impaires sont présents.

Pour un tube conique par contre, tous les harmoniques sont présents car la propagation du son est de type sphérique. Enfin, dans le cas d'un cône évasé (cor anglais), les partiels ne sont plus harmoniques ($1f$, $1,8f$, $2,7f$, $3,6f$ [Rossing-1990, p. 244]).

I.3. Les techniques d'analyse des sons

Dans les paragraphes précédents, nous avons étudié la production du son, en indiquant que l'ordinateur peut servir à modéliser les différentes parties d'un système vibrant, comme celles d'un instrument de musique, pour réaliser la synthèse de son. Ces techniques sont complexes et commencent tout juste à fournir des programmes de synthèse par modèles physiques utilisables pour la réalisation de pièces musicales.

Un des apports de l'informatique réside dans les possibilités offertes dans le domaine de l'analyse du son ou plus généralement du traitement du signal. L'ordinateur est remarquable à la fois par la diversité et par la précision des analyses qu'il permet de réaliser et pour les possibilités qu'il offre de réutiliser les données fournies par l'analyse afin d'établir des modèles ou des règles de synthèse.

John Chowning indique très justement dans un article paru en 1993 ce qu'il considère comme « les faits importants qui découlent de l'utilisation de l'informatique en musique » : pour lui, trois points sont essentiels : le fait que l'utilisation des ordinateurs révèle des caractéristiques du son dont l'importance peut être confirmée par l'écoute mais que l'écoute ne permet pas d'étudier ; le fait que la synthèse, à travers la modification systématique des paramètres des sons synthétisés, permette de souligner les importances relatives de ces paramètres par rapport à la perception, en utilisant les techniques de la psycho-acoustique et pas seulement une description purement acoustique ; enfin le fait qu'en supprimant des détails non décelés par l'écoute on puisse procéder à une réduction des données nécessaires à la synthèse et rendre les ordinateurs plus efficaces sans pour autant altérer la perception [Chowning-1993, p. 5].

Nous allons donc étudier le signal sonore, et non plus la source émettrice, pour indiquer quels sont les paramètres qui permettent de caractériser le son et, surtout, quels sont leurs rôles vis-à-vis de la perception. Ces informations nous permettront de mieux comprendre les principes des différents outils de synthèse ainsi que la démarche des différents compositeurs présentés dans ce mémoire.

I.3.1. Les grandeurs physiques du son

I.3.1.1. La fréquence

Les premières mesures de la fréquence de vibration d'un corps oscillant ont été réalisées en observant des cordes de très grandes dimensions. En divisant progressivement la longueur d'une corde par deux, on double à chaque fois la fréquence de la vibration. On peut ainsi passer du domaine du rythme, dans lequel on peut compter les oscillations, au domaine des hauteurs et établir la relation entre la vitesse de vibration et la hauteur du son.

Au XIX^e siècle, pour déterminer avec précision la hauteur d'un son, les chercheurs ont utilisé une méthode comparative qui consistait à confronter la hauteur d'un son avec celle d'une sirène de fréquence connue.

Helmholtz a utilisé des résonateurs en verre pour mettre en évidence les fréquences partielles, en particulier les harmoniques, d'un son. Il en a déduit des théories sur la tonalité et sur les consonances. Il a également indiqué les relations entre les systèmes d'accord des instruments selon les différents tempéraments.

Actuellement, les instruments de la musique occidentale sont accordés selon un système tempéré ne permettant pas d'obtenir de quinte juste. En effet, si pour accorder un instrument couvrant la gamme chromatique on ajuste les fréquences des notes en progressant de quinte juste en quinte juste, on parvient, après avoir parcouru l'ensemble des douze notes, à un écart de fréquence entre la note de départ et la note d'arrivée égal à 1,36% de la fréquence originale. Cela correspond à un intervalle compris entre 1/8^e et 1/9^e de ton.

rapport à Do	note
1,0000	Do
1,5000	Sol
1,1250	Ré
1,6875	La
1,2656	Mi
1,8984	Si
1,4238	Fa#
1,0679	Do#
1,6018	Sol#
1,2014	Ré#
1,8020	Sib
1,3515	Fa
1,0136	Do

tableau I-4 : rapports de fréquences entre les douze notes de la gamme chromatique en progressant par intervalles de quintes justes (ascendantes ou descendantes)

On ne peut donc pas obtenir d'accord pour lequel toutes les quintes soient justes. Différents systèmes d'accord existent et la solution qui a été adoptée, depuis Jean Sébastien Bach, a été le **tempérament égal**. Le tempérament égal est obtenu par une progression vers le haut de demi-ton en demi-ton en multipliant à chaque fois la fréquence par $2^{1/12}$ soit par 1,05946.

rapport à Do	note
1,0000	Do
1,4983	Sol
1,1225	Ré
1,6818	La
1,2599	Mi
1,8877	Si
1,4142	Fa#
1,0595	Do#
1,5874	Sol#
1,1892	Ré#
1,7818	Sib
1,3348	Fa
1,0000	Do

tableau I-5 : rapports de fréquences entre do et les douze notes de la gamme chromatique dans un tempérament égal

Si l'on prend comme note de référence le « La » du diapason à 440 Hz, la formule précédente permet de donner la fréquence fondamentale de toutes les notes de la gamme chromatique tempérée.

note	MIDI	F (Hz)
Do ₃	60	261.63
Do# ₃	61	277.18
Ré ₃	62	293.66
Mib ₃	63	311.13
Mi ₃	64	329.63
Fa ₃	65	349.23
Fa# ₃	66	369.99
Sol ₃	67	392.00
Sol# ₃	68	415.30
La₃	69	440.00
Sib ₃	70	466.16
Si ₃	71	493.88

tableau I-6 : tableau de correspondance entre les notes, les numéros MIDI et les fréquences pour une octave située autour du diapason

La hauteur d'un son est un paramètre lié à sa perception et provient donc d'une appréciation subjective. Si les musiciens arrivent généralement à déterminer la hauteur d'un son musical, cette détermination peut dans certains cas s'avérer difficile, voire impossible à réaliser.

Dans le cas de sons complexes, par exemple des sons de cloches ou des sons multiphoniques, plusieurs hauteurs distinctes peuvent être perçues. On peut percevoir des **hauteurs** dites « **virtuelles** », correspondant à la perception d'espacements réguliers entre les principaux partiels contenus dans le son ou des **hauteurs** « **spectrales** » dues à une forte intensité de la composante la plus basse du son.⁹

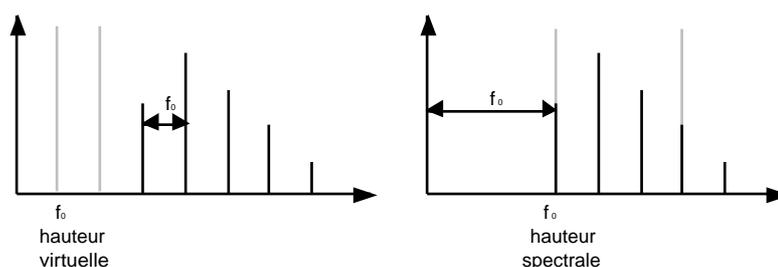


figure I-29 : représentation graphique de la hauteur virtuelle et de la hauteur spectrale à partir d'un même spectre de fréquences.

Suivant les sons considérés, la sélection de l'une ou l'autre des deux hauteurs peut s'avérer difficile. En outre, la détermination de la hauteur d'un son peut être encore compliquée par le fait que souvent les partiels ne sont pas parfaitement harmoniques et que la hauteur perçue varie quand l'intensité du son change ou quand d'autres sons lui sont superposés. La plupart de ces critères d'évaluation sont utilisés dans l'algorithme de recherche des hauteurs virtuelles « virtual-pitch theory » mis au point par Ernst Terhardt et décrit au paragraphe 1.3.4.1 [Terhardt, Stoll, Seewa-1982a].

La détermination de la hauteur est délicate dans le cas de certains sons très aigus ou de sons inharmoniques.

Parfois, le choix de l'octave à l'intérieur de laquelle se situe le son peut s'avérer difficile à réaliser. C'est par exemple le cas du violoncelle présenté à la figure I-36.

D'autre part, la sensation de hauteur est liée à l'intensité du son. Stevens a observé des écarts de fréquence d'environ 10% pour des sons purs de différentes intensités [Leipp-1989, p. 121].

⁹ Parmi ses exemples sur les paradoxes de hauteur, Jean-Claude Risset a produit un son dont la hauteur semble baisser lorsqu'il est joué deux fois plus vite. Le son en question est simplement formé d'un ensemble de partiels qui se suivent à intervalle de 1,1 octave. Lorsqu'on multiplie la vitesse de lecture du son par deux, tous les partiels sont transposés d'une octave et à la place de chaque partiel, on trouve un nouveau partiel situé 0,1 octave plus bas, excepté le partiel le plus grave dont on ne remarque pas la disparition (exemple sonore n°10).

Enfin, la hauteur d'un son est rarement stable au cours du temps. Lorsqu'un son est soumis à un vibrato, les écarts de hauteur peuvent dépasser un ton.

I.3.1.2. L'amplitude et l'intensité du son

Si l'on considère un signal variant au cours du temps selon la loi $A=u(t)$, on peut définir son amplitude moyenne A_0 pendant le temps T par :

$$A_0 = \frac{1}{T} \int_0^T u dt$$

Physiquement, l'intensité I d'un son en un point quelconque est liée à l'amplitude du mouvement vibratoire. I correspond à l'énergie produite par ce mouvement par unité de temps et de surface et se mesure en watt/m².

$$I = 2\Pi^2 A^2 \rho F^2 v$$

A est l'amplitude du signal au point considéré, assimilée à une pression, F est la fréquence du signal, ρ est la masse spécifique du milieu et v est la vitesse de propagation du signal.

La puissance acoustique émise par une source ponctuelle s'étale sur la surface d'une sphère dont l'aire est exprimée par :

$$S = 4\Pi r^2$$

où r désigne la distance du point à la source.

Si la puissance originale émise par la source est W (en watts), alors l'intensité I , ou puissance par unité de surface est :

$$I = \frac{W}{4\Pi r^2}$$

L'intensité du son est donc inversement proportionnelle au carré de la distance lorsque l'on fait abstraction des phénomènes d'absorption du son dans l'air.

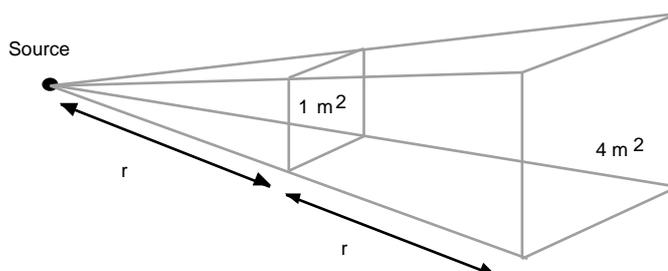


figure I-30 : variation de la surface rayonnée à partir d'une source sonore en fonction de la distance

L'intensité perçue, appelée communément le « volume d'un son », est proportionnelle au logarithme de l'intensité physique. En multipliant l'intensité physique du son par dix, l'intensité perçue est à peu près multipliée par deux. L'audition humaine peut percevoir une gamme d'intensités très étendue, en effet, la pression de l'air causée par le son le plus bruyant est de plusieurs billions de fois plus élevée que celle causée par le son le plus bas que l'on puisse entendre.

Pour mesurer les variations entre une intensité de référence I_0 et une intensité sonore I_1 , on utilise une grandeur appelée le gain, ΔI , exprimée en décibels (dB). On l'exprime par la formule suivante :

$$\Delta I = 10 \log\left(\frac{I_1}{I_0}\right) = 20 \log\left(\frac{A_1}{A_0}\right)$$

L'intensité sonore de référence, I_0 , est généralement égale à l'intensité sonore au seuil d'audibilité à 1000 Hz :

$$I_0 = 10^{-12} \text{ W/m}^2$$

Le niveau subjectif perçu n'est pas le même pour une intensité donnée selon les fréquences. La sensibilité de l'ouïe varie avec la hauteur du son. La mesure du **niveau sonore perçu** s'effectue avec une unité appelée le *phone*. Pour un son de 1000 Hz, l'échelle des phones coïncide avec l'échelle des intensités des sons en dB dont on situe l'origine au seuil d'audibilité. Pour l'ensemble des fréquences, l'échelle des phones est donnée par un diagramme expérimental établi par Fletcher et Munson.

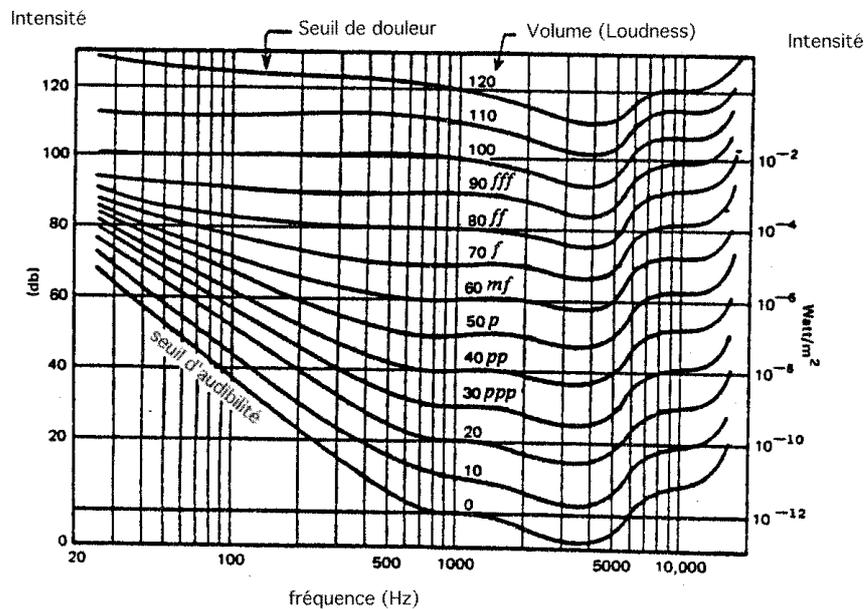


figure I-31 : diagramme de Fletcher et Munson [Moore-1990, p. 23]

Lorsque deux sons d'intensités égales se superposent, l'amplitude résultante est doublée et l'intensité est multipliée par quatre. Par contre, la variation d'intensité mesurée en dB sera de $10 \cdot \log 4 = 6,0206$ soit d'environ +6dB, quelles que soient les intensités de ces sons.

I.3.2. Les représentations de l'amplitude du son par rapport au temps

Un son numérique est constitué par une liste de nombres. La norme utilisée pour le disque compact correspond à un taux d'échantillonnage de 44100 échantillons par seconde pour une résolution d'amplitude de 16 bits.

Pour la plupart des applications informatiques utilisant du son numérique, celui-ci peut être représenté sous une forme graphique. La forme la plus élémentaire de ces représentations est une courbe ayant pour abscisse le temps et pour ordonnée l'amplitude du signal.

A travers trois exemples, une note de guitare, une note tenue de contrebasse et un fragment de voix parlée, nous allons indiquer quelles informations peut nous apporter la représentation de l'amplitude du son par rapport au temps.

I.3.2.1. L'enveloppe d'amplitude

En utilisant pour l'axe du temps une échelle d'affichage des données de l'ordre de la seconde, la représentation graphique donne une idée de la variation d'amplitude globale du son en fonction du temps. Les points affichés sur le graphe correspondent à des moyennes des valeurs d'amplitude du son. La courbe obtenue est appelée **l'enveloppe d'amplitude** du son.

Les enveloppes d'amplitude des sons de guitare et de contrebasse et le rythme de l'articulation de la voix parlée, permettent facilement de différencier les trois sons.

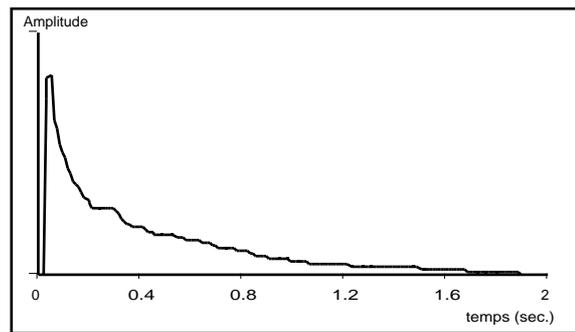


figure I-32 : enveloppe d'amplitude d'un son de guitare, corde nylon, note La3(ex. sonore 2).

L'attaque du son de la guitare est très rapide. Le son décroît ensuite de façon exponentielle par amortissement de la vibration. On observe une enveloppe de type « percussion/résonance ».

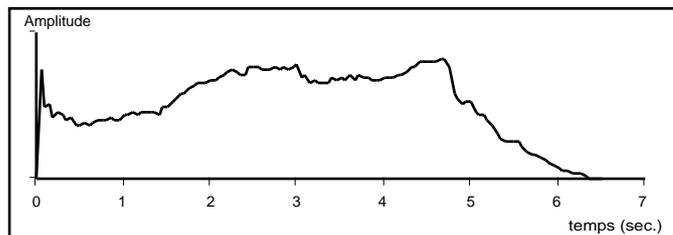


figure I-33 : enveloppe d'amplitude d'un son de violoncelle tenu joué à l'archet, note Do1(ex. son. 3).

Le violoncelle joué à l'archet présente un niveau d'amplitude assez stable. Sur les deux dernières secondes l'archet ne touche plus la corde et le son s'amortit progressivement.

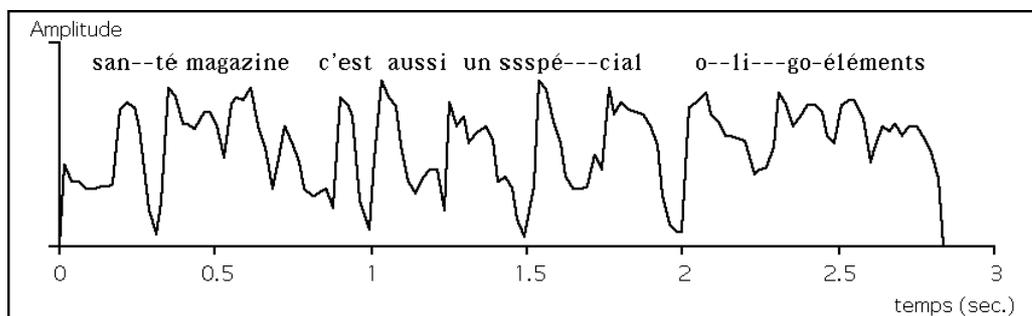


figure I-34 enveloppe d'amplitude d'un extrait de voix parlée (exemple sonore n°4).

Dans le cas de la voix parlée, l'enveloppe du son est plus complexe que dans les cas précédents. Il ne s'agit plus d'une note, mais d'une phrase dont certains détails de l'articulation peuvent être repérés sur le graphique lorsque l'on étudie les variations d'amplitude.

I.3.2.2. La forme d'onde

Lorsque l'on utilise des unités situées autour de la milliseconde pour l'axe du temps, on obtient un tracé qui représente la forme de la vibration. On visualise l'onde sonore et on parle de « forme d'onde ». Dans le cas d'un son périodique, la dénomination « forme d'onde » peut être attribuée à la portion de la forme de l'onde limitée à une seule période du signal.

La forme de l'onde contient toute l'information qui permet de définir le timbre d'un son. Cependant, cette représentation graphique n'est pas interprétable telle qu'elle. En effet, un même instrument peut présenter des formes d'onde extrêmement variées, même à l'intérieur d'un seul son. Ceci est souvent dû à des déphasages entre les différents partiels qui composent le son. De plus, des formes d'onde assez proches correspondent parfois à des timbres différents.

Néanmoins, quelques informations peuvent être extraites de ce mode de représentation du son bien qu'elles restent rudimentaires et approximatives la plupart du temps.

Dans le cas d'un son périodique, il est souvent possible de repérer les cycles de la vibration et d'en déduire la période.

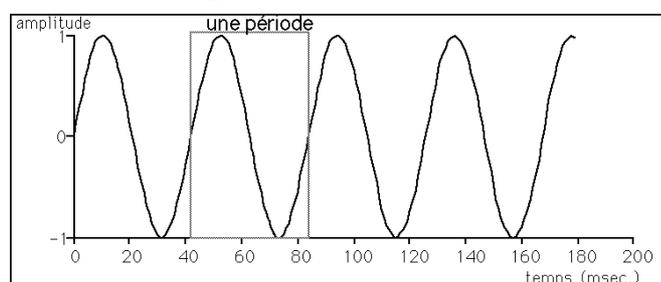


figure I-35 : détection graphique de la période d'un son sinusoïdal

Le degré de complexité de l'onde, lisse ou perturbée, peut également donner des informations sur le contenu harmonique du son. Une forme d'onde perturbée et comportant de violentes ruptures de pente correspond souvent à un timbre riche mais ces informations sont généralement peu explicites.

(a)

(b)

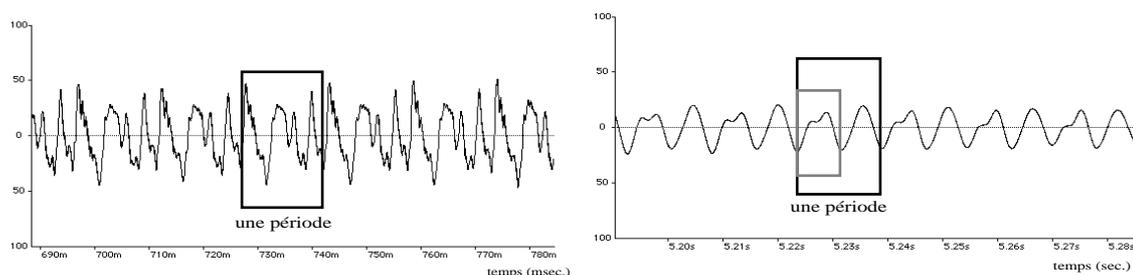


figure I-36 : formes d'onde d'un son de violoncelle :

(a) dans la partie entretenue du son, (b) dans la partie résonante du son (note Do1 : fréquence 65,4 Hz, période = 15,3 msec.) (exemple sonore n°3).

Sur les courbes de la figure I-36, prélevées sur le même son de violoncelle en deux points différents, on observe deux formes d'onde nettement différenciées.

Dans la partie entretenue du son, la forme d'onde est perturbée et la période est difficile à déceler.

Dans la partie résonante du son, la forme d'onde est plus lisse, ce qui indique un son plus pur, dépourvu d'harmoniques de rang élevé. Visuellement, on peut hésiter entre la note Do1 et Do2. A l'audition, la même hésitation persiste.

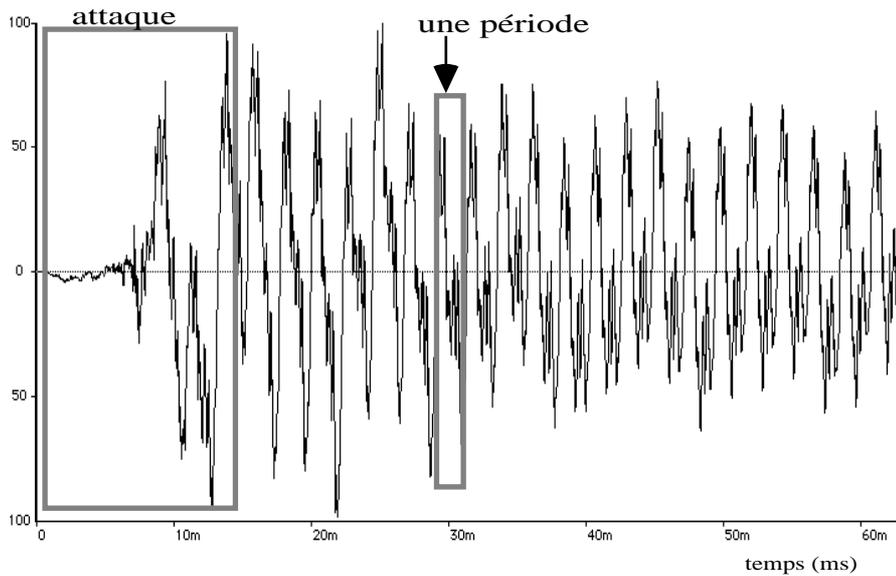


figure I-37 : forme d'onde d'un son de guitare (exemple sonore n°2).

La figure I-37 représente les soixante premières millisecondes d'un son de guitare. Dans les quinze premières millisecondes, on observe une phase complexe et transitoire à la suite de laquelle s'établit un régime périodique qui n'est pas entièrement stable. L'intensité et la forme de la vibration varient légèrement d'une période à l'autre.

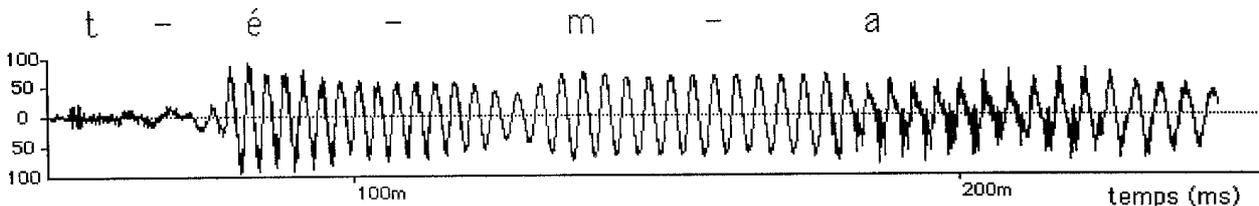


figure I-38 : voix parlée « ..té ma.. » (exemple sonore n°4).

Dans le cas de la voix parlée, l'analyse du graphe permet de séparer chaque lettre des deux syllabes. Le « t » est une consonne plosive, bruitée, sans organisation périodique ; les deux voyelles sont nettement périodiques et présentent des formes d'onde assez anguleuses ; le « m » est un son périodique assez proche d'un son sinusoïdal.

La représentation de la forme d'onde permet de délimiter des fragments de sons avec une grande précision.

I.3.3. Les représentations spectrales du son

I.3.3.1. Les transformées de Fourier

Au début du XIX^e siècle, Fourier découvrit qu'il était possible sous certaines conditions de décrire une fonction arbitraire comme la somme de fonctions trigonométriques. Cela aboutit à la formulation du principe de la transformée de Fourier : une fonction périodique, ou vibration, de période T peut être transformée en une série de fonctions sinusoïdales de fréquences $f, 2f, 3f, 4f$ etc., avec $f=1/T$. Ces fréquences sont appelées les harmoniques de la vibration. Le calcul permet de leur associer deux valeurs : une amplitude A_j et une phase Φ_j . Le premier harmonique est appelé la fondamentale.

C'est Ohm qui, en 1843, proposa d'appliquer la théorie de Fourier au signal sonore. Il est le premier scientifique à avoir indiqué que tout son harmonique peut se décomposer en une somme de sons purs de fréquences multiples de la fondamentale.

Les transformées de Fourier (cf. Annexe I) fournissent un ensemble de coefficients d'amplitude et de phase qui, une fois attribués à un ensemble de signaux sinusoïdaux de différentes fréquences, permettent de reconstituer le son original. Elles sont au son ce que le prisme est à la lumière.

Les transformées de Fourier sont le point de départ de la plupart des techniques actuelles d'analyse et de transformation des sons. Elles présentent un ensemble de propriétés mathématiques assimilables aux propriétés physiques du signal. Elles permettent en particulier d'établir une correspondance entre le **domaine temporel**, relatif aux formes d'onde du signal et aux valeurs d'amplitude des échantillons, et le **domaine fréquentiel**, relatif aux amplitudes et phases des différentes fréquences partielles d'un son.

Cette correspondance est essentielle car l'oreille humaine entend les sons dans le domaine fréquentiel alors que ceux-ci sont stockés en mémoire sur l'ordinateur dans le domaine temporel.

Le théorème de Fourier s'applique à un signal continu, supposé infini dans le temps. Les sons numériques ont une durée limitée et leur signal $x(t)$ est une fonction discrète du temps. On utilise alors les transformées de Fourier discrètes (DFT pour « Discret Fourier Transform »).

Théoriquement, les transformées de Fourier discrètes sont destinées à l'étude de sons harmoniques. La portion de son à analyser doit être exactement égale à une période du son. La DFT calcule le spectre de l'onde en donnant la liste des fréquences harmoniques, leurs amplitudes et leurs phases.

La partie de son sur laquelle porte la DFT constitue ce que l'on appelle une « fenêtre ».

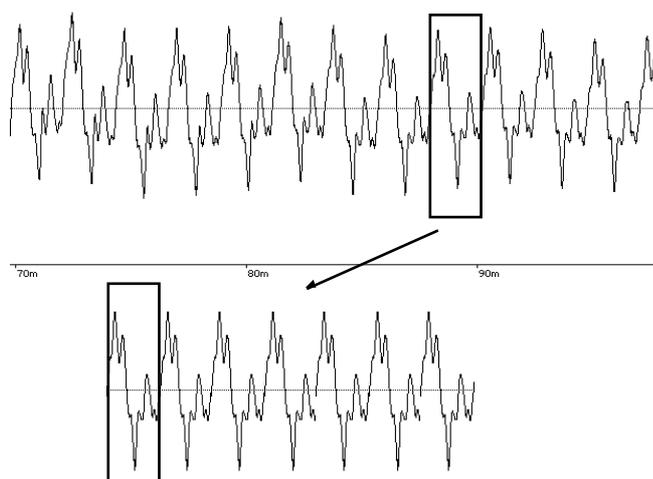


figure I-39 : extraction d'une fenêtre de son pour l'analyse en DFT et reconstitution du signal d'après les résultats de l'analyse.

Cette décomposition du signal en une somme de sinusoïdes est unique. Il n'existe aucun autre ensemble de sinusoïdes qui puisse recomposer le signal original. Par contre, il existe d'autres décompositions possibles du signal en une somme de signaux élémentaires non sinusoïdaux.

L'informatique, par le biais de la DFT, nous permet de mesurer avec exactitude l'énergie contenue dans un son selon différentes fréquences. Ceci permet de

poursuivre les travaux engagés par Helmholtz qui, avec ses résonateurs, tentait de mesurer subjectivement l'énergie contenue dans un son à une fréquence donnée.

Si la taille de la fenêtre ne coïncide pas avec la période, les résultats de la DFT sont faussés.

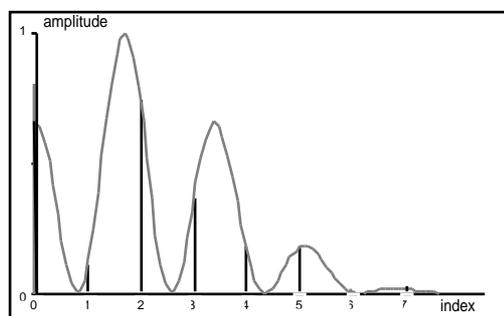


figure I-40 : valeurs obtenues pour une DFT (en noir) dont la fréquence fondamentale ne coïncide pas avec la fréquence fondamentale du son.

Le spectre réel du son est donné par la courbe en pointillés.

Pour étudier un son harmonique, on choisit donc une fréquence d'analyse égale à la fréquence fondamentale du son à analyser de façon à déterminer l'amplitude de chacun de ses harmoniques. Pour un son quelconque, l'analyse va introduire des artefacts en assimilant la portion du son à analyser à sa période. Ces artefacts peuvent être atténués en multipliant le signal de la portion de son à analyser par une fonction qui en atténue les extrémités.

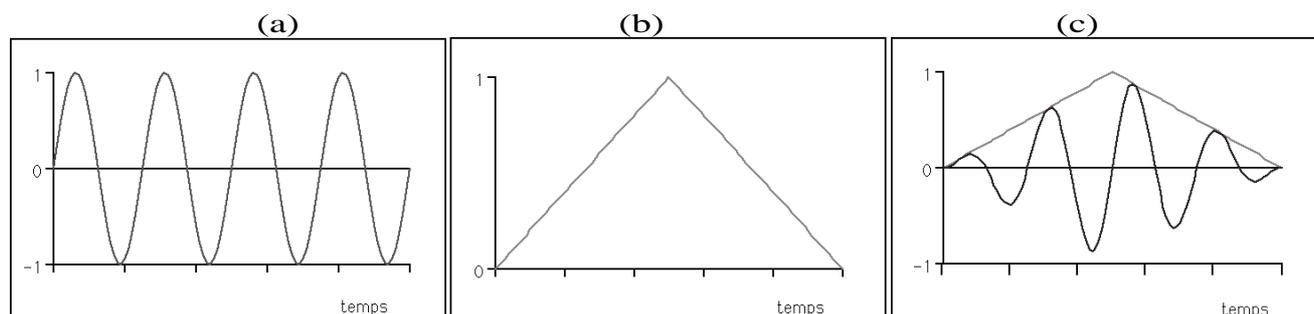


figure I-41 : (c) : multiplication d'une partie du signal à analyser (a) par une fonction triangulaire (b).

Dans le cas de sons quelconques, on peut effectuer des DFT sur des fenêtres de grandes dimensions. On obtient alors une liste de fréquences à forte résolution qui indique quelles sont les fréquences des sons partiels. Les valeurs des amplitudes et des phases que l'on obtient pour ces sons partiels correspondent aux moyennes des amplitudes et des phases observées sur la durée de la fenêtre. Le fait que les valeurs des amplitudes et des phases soient données par des moyennes ne présente aucun problème dans le cas d'un son stable mais il se trouve que les sons naturels ne sont jamais stables. En effet, les paramètres du son varient de façon complexe, en particulier dans les phases transitoires, lors de l'attaque ou de la chute d'une note et lors du passage d'une note vers une autre. Pour un son tenu, on peut observer des changements dus au vibrato, au trémolo et à des microvariations d'amplitude et de fréquence. De même, lorsque l'amplitude d'un son augmente, le timbre est généralement modifié, le plus souvent avec une accentuation des partiels aigus.

Pour repérer des variations rapides des paramètres du son, il faut donc limiter les tailles des fenêtres d'analyse. Il en résulte un antagonisme entre la résolution en fréquence que l'on améliore en augmentant la taille de la fenêtre et la résolution temporelle que l'on améliore en diminuant la taille de la fenêtre. Une bonne analyse est un compromis réussi entre ces deux contraintes.

Lorsque la taille de la fenêtre d'analyse augmente, le calcul de la DFT requiert un nombre d'opérations progressant suivant le carré de N . Le temps de calcul peut devenir prohibitif pour des fenêtres de grande taille. Il existe une technique d'analyse par transformée de Fourier qui permet de calculer la DFT beaucoup plus rapidement : la FFT ou « Fast Fourier Transform ».

La FFT est une DFT qui a été optimisée pour le calcul informatique. Le nombre d'opérations à réaliser pour la FFT croît proportionnellement à $N \log N$ au lieu de N^2 pour la DFT. L'optimisation porte sur un regroupement des données qui permet d'éviter la répétition de certaines opérations. La FFT impose comme contrainte de choisir $N = 2^x$, x étant un nombre entier. Pour analyser une portion de son de taille inférieure à une fenêtre choisie, on complète cette portion de son par des valeurs nulles pour obtenir une fenêtre complète.

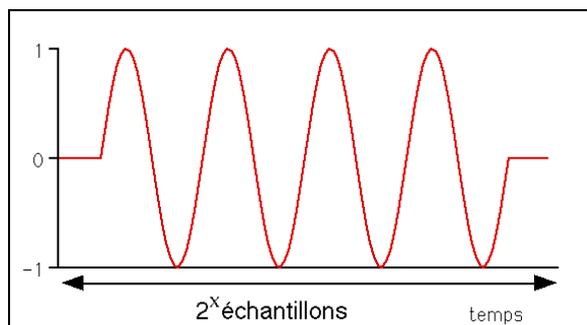


figure I-42 : complément du nombre d'échantillons à 2^x par ajout de valeurs nulles.

I.3.3.2. Les caractéristiques spectrales des sons

Certaines formes d'onde donnent des spectres caractéristiques. Par exemple, une **série d'impulsions** également espacées donne un spectre dont les amplitudes sont les mêmes pour toutes les fréquences harmoniques des partiels. Une forme d'onde en **dents de scie** produit un spectre dont les amplitudes diminuent de façon inversement proportionnelle à la fréquence, en présentant une décroissance d'environ 6 dB par octave. La forme d'**onde carrée** présente également une décroissance de 6 dB par octave mais les harmoniques pairs sont absents. Une forme d'**onde triangulaire** donne un spectre dont les amplitudes diminuent de façon inversement proportionnelle au carré de la fréquence, en présentant une décroissance d'environ 12 dB par octave.

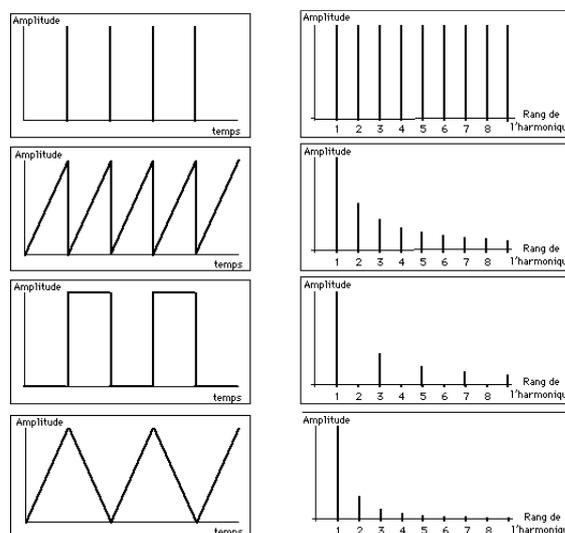


figure I-43 : exemples de signaux harmoniques et de leurs FFT associées

Les sons instrumentaux ne sont jamais totalement harmoniques. On peut toutefois les classer en plusieurs types.

Le premier type est celui des **sons à hauteur fixe et à amplitude stable**. C'est le cas de la plupart des sons entretenus produits par les instruments à cordes ou à vent. Des FFT réalisées en plusieurs endroits dans un même son vont donner approximativement les mêmes spectres : les formes d'onde évoluent lentement. Nous qualifierons ces sons de « **quasi-périodiques** ».

Le second type est celui des **sons à hauteur fixe et à intensité et timbre variables**. On peut les séparer en deux catégories : dans la première, nous avons placé les **sons percussifs et résonants**, comme ceux produits par la plupart des instruments de percussion et par les instruments à cordes pincées ou frappées ; dans la deuxième, nous trouvons les **sons entretenus** variant en timbre et en intensité comme les sons émis par les cuivres pour lesquels les variations de dynamique sont accompagnées d'importantes variations de timbre. Dans ce cas nous parlerons de sons « **pseudo-périodiques** ».

Le troisième type comprend les **sons inharmoniques, et polyphoniques** qui peuvent être stables ou instables. Il peut s'agir de sons de cloches, de sons multiphoniques d'instruments à vent et, de façon plus générale, de polyphonies instrumentales ou orchestrales.

Enfin, le dernier type est celui des **sons à hauteurs indéterminées**, souvent qualifiés de bruits. Ceux-ci peuvent couvrir tout le spectre des fréquences ou n'en couvrir qu'une fraction délimitée. Ils peuvent être stables comme le bruit blanc, agités comme le bruit du vent lors d'une tempête, percussifs comme un coup de feu ou une impulsion.

Suivant la nature du son que l'on cherche à analyser, la DFT peut être abordée de différentes façons.

Pour un son dont on connaît la hauteur, on peut effectuer une DFT dans la partie stable du son en choisissant une taille de fenêtre égale à la période du son. La DFT fournit alors les amplitudes et phases respectives des différents harmoniques de ces sons. Le tableau I-6 en est un exemple, il correspond à l'analyse d'un son de violoncelle.

rang	fréquence	amplitude	phase
1	110.000	0.061153	90.00
2	220.000	0.009381	-150.96
3	330.000	0.043478	130.79
4	440.000	0.009761	108.54
5	550.000	0.055024	32.44
6	660.000	0.077452	-120.60
7	770.000	0.033741	44.40
8	880.000	0.032959	43.54
9	990.000	0.013719	142.95
10	1100.000	0.012247	-158.04
11	1210.000	0.007972	-8.70
12	1320.000	0.008154	-29.80
...			
77	8470.000	0.000021	-159.60

tableau I-7 : résultat de l'analyse en DFT d'un son de violoncelle (note La1)

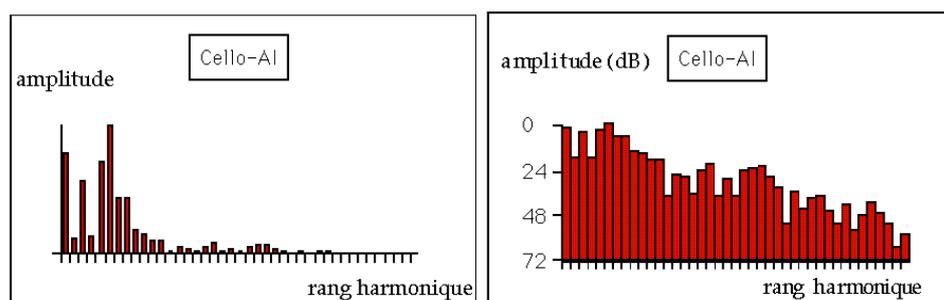


figure I-44 : spectre d'une analyse en DFT d'un son de violoncelle (note La1)
A gauche amplitudes linéaires, à droite, amplitudes en dB.

Suivant les instruments, des partiels indépendants de la série harmonique peuvent être présents et jouer un rôle important dans le timbre de l'instrument. Si l'on cherche à connaître l'ensemble des partiels qui composent un son, il faut effectuer une FFT portant sur une fenêtre de taille importante, par exemple égale à une dizaine de fois la période, pour un son de hauteur déterminée.

Sur la figure I-45 certaines raies de fréquence apparaissent au-dessus des autres, ce sont des pics. Le premier pic (raie numéro 6) donne une fréquence d'environ 64 Hz et correspond à la fondamentale du son. Les pics suivants sont tous séparés par le même écart : ce sont les harmoniques.

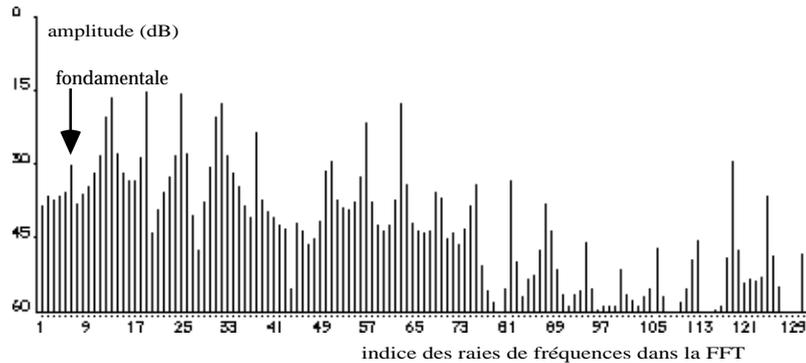


figure I-45 : spectre d'une note de violoncelle pour une taille de fenêtre de FFT de 4096 échantillons. Sur l'axe des fréquences, une graduation = 10,7 Hz

En changeant d'échelle, on obtient un contour moins précis :

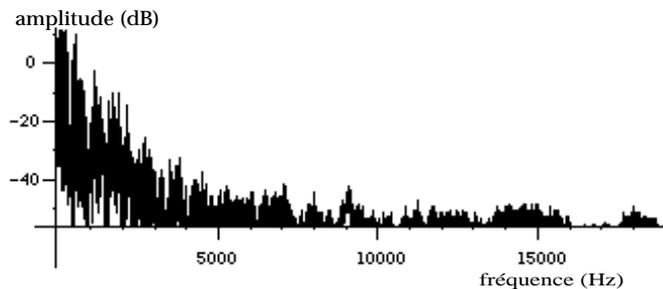


figure I-46 : spectre d'une note de violoncelle pour une taille de fenêtre de FFT de 4096 échantillons.

En traçant une courbe qui relie les différents sommets des pics entre eux on obtient le graphe de l'enveloppe spectrale.

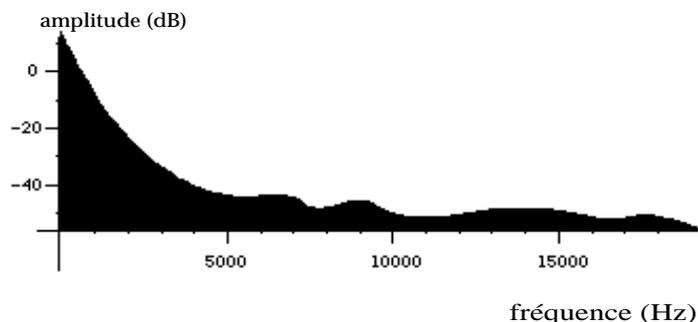


figure I-47 : enveloppe spectrale d'une note de violoncelle

Il est également possible de réaliser une FFT d'une taille égale à la durée totale d'un son. Dans ce cas, en utilisant les données obtenues, on peut effectuer une resynthèse exacte du son mais la quantité de données fournies par cette FFT est tellement importante que celles-ci sont inexploitable pour une étude du son.

Si, comme nous l'avons indiqué, la FFT permet de passer d'une représentation temporelle du son à une représentation spectrale beaucoup plus proche de la perception, quelques cas particuliers peuvent poser un problème conceptuel.

Lorsque l'on analyse une simple impulsion - un « dirac » - on obtient un spectre formé de la somme de toutes les fréquences audibles à des amplitudes égales.

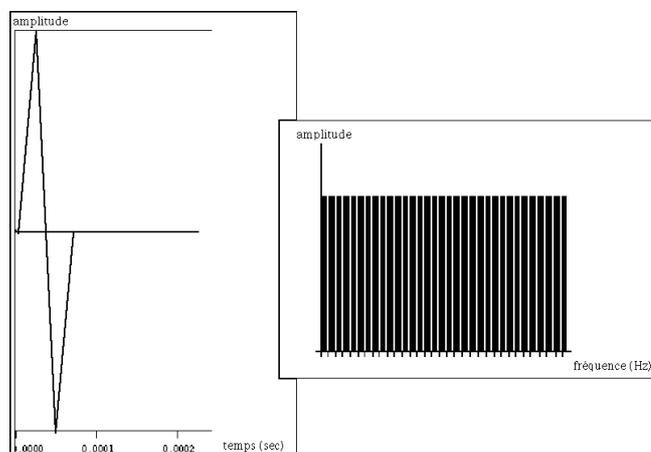


figure I-48 : à gauche un dirac, à droite, son spectre associé

Il peut apparaître comme un non-sens d'imaginer qu'une impulsion d'une durée infiniment courte puisse contenir toutes les fréquences à la fois. Aussi doit-on prendre garde aux conclusions à tirer des analyses.

L'analyse d'un bruit blanc, son obtenu en prenant comme signal une valeur aléatoire pour chaque échantillon, donne, comme celle du dirac, un spectre contenant toutes les fréquences audibles pour des amplitudes égales. Seules les différences de phases différencient le spectre d'un bruit blanc de celui d'un dirac. A l'audition, les deux sons ne présentent pourtant aucune analogie.

Par ailleurs, certains sons harmoniques peuvent donner un spectre dans lequel la fondamentale est absente. Ceci ne modifie pas la perception de la hauteur du son qui est plus liée aux écarts de fréquences entre les partiels qu'à la présence de la fondamentale elle-même. Dans l'exemple ci-dessous, un son synthétique a été réalisé en additionnant les harmoniques de rangs 4 à 14 d'une fondamentale située à 200 Hz. Malgré l'absence de la fondamentale et de ses premiers harmoniques, la hauteur perçue est bien de 200 Hz. Sur le schéma de l'onde, la période observée est de 5 ms. ce qui confirme le fait que la fréquence fondamentale de ce son est bien de 200 Hz.

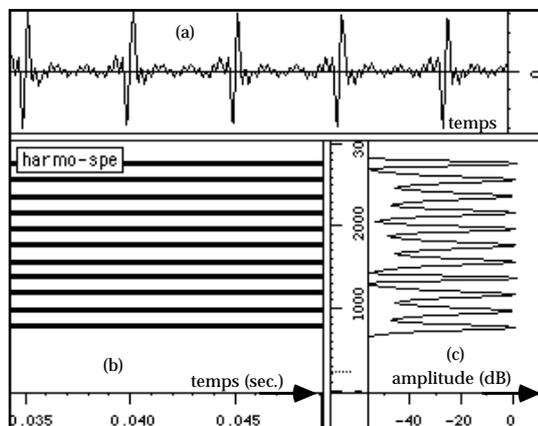


figure I-49 : représentation d'un son harmonique de synthèse, (a) : forme de l'onde, (b) : sonagramme, (c) : spectre.

I.3.3.3. La recherche de pics

Lorsque la FFT a été réalisée sur de grandes tailles de fenêtres, il est utile d'essayer de réduire la quantité d'informations qu'elle délivre, pour ne conserver que l'information pertinente. Etudions par exemple les résultats fournis par une FFT réalisée sur un son pur de 100 Hz. Ceux-ci sont représentés par les graphes de la figure I-50 en fonction des différentes tailles de fenêtres utilisées.

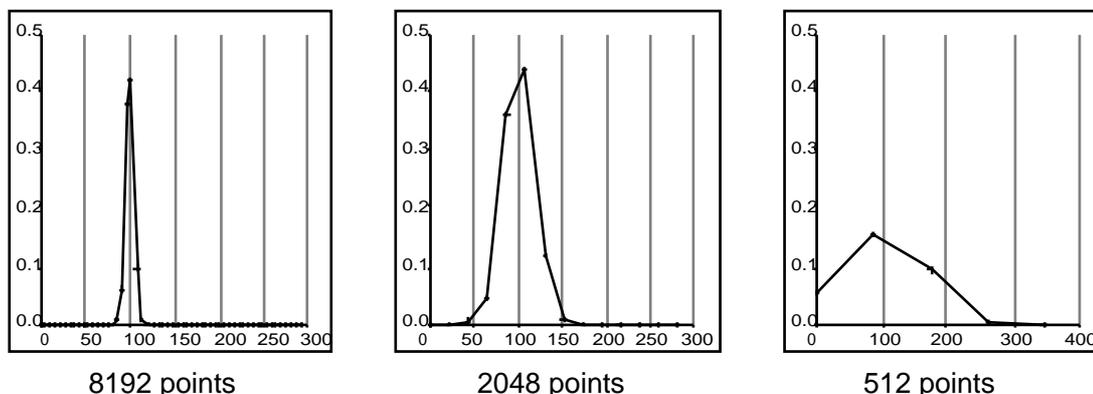


figure I-50 : FFT d'un sinus à 100 Hz selon différentes tailles de fenêtres.

Dans le cas d'une taille de fenêtre de 512 points — ou échantillons —, à la place de la valeur 100 Hz correspondant à la réalité, les valeurs les plus proches calculées par la FFT sont de 86.13 Hz et 172.27 Hz.

Pour réduire l'information contenue dans la FFT et pour améliorer la résolution fréquentielle de l'analyse, il est possible d'effectuer une recherche de pics.

Ceux-ci peuvent être extraits de la FFT par plusieurs algorithmes. Un algorithme simple consiste à chercher les maximums locaux du spectre puis à interpoler les pics ainsi détectés en fonction des points voisins en utilisant la méthode de Lagrange qui repose sur la recherche d'un polynôme du troisième degré passant par ces points (cf Annexe II).

Les figures qui suivent représentent les pics obtenus par l'analyse d'un son de cymbalum, quasi-harmonique, et par l'analyse d'un son de cloche, inharmonique. Un quadrillage vertical met en évidence la fréquence fondamentale du son et ses multiples.

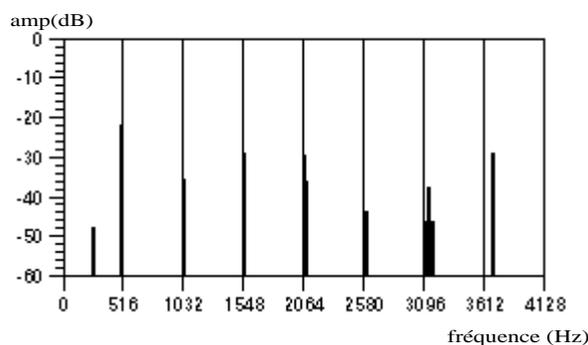


figure I-51 : son de cymbalum, note Do4
(analyse par modèles de résonance)

Le son du cymbalum apparaît comme un son fortement harmonique puisque presque tous les pics sont placés sur les divisions de l'axe des fréquences. On peut noter une légère inharmonicité pour les harmoniques élevés qui s'écartent progressivement des rapports rigoureusement harmoniques. On peut également constater pour les harmoniques de rang 4, 5 et 6, la coexistence de partiels très proches dont les interactions peuvent créer des phénomènes vibratoires

complexes dans le son, notamment des battements dont les fréquences sont égales aux fréquences différentielles de ces partiels.

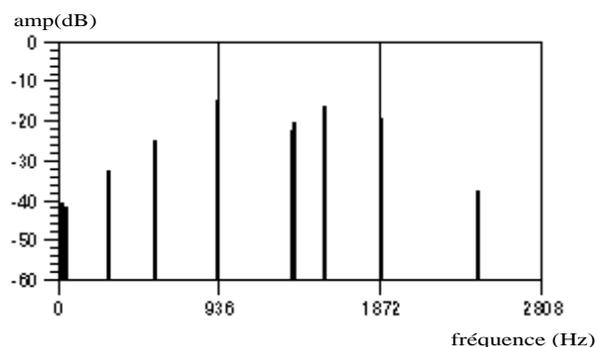


figure I-52 : son de cloche tubulaire note Sib4
(analyse par modèles de résonance)

Le son de la cloche est un son inharmonique. Ses partiels principaux ne possèdent pas de diviseur commun. La note perçue est le Sib4 (936 Hz),

I.3.3.4. Le sonagramme

Contrairement à ce qu'on pensait au début du XX^e siècle, le spectre du son à un instant donné ne fournit qu'une information très partielle du timbre. Les analyses décrites précédemment sont des analyses statiques qui ne prennent pas en compte l'évolution des paramètres du son dans le temps. Pour étudier l'évolution du son dans le temps, il est nécessaire de réaliser un ensemble d'analyses à différents endroits dans le son. Dans le cas de sons quasi-périodiques ou pseudo-périodiques, on doit choisir une taille de fenêtre 5 à 10 fois supérieure à la période présumée du son. On reproduit ensuite l'analyse à intervalles de temps réguliers, en progressant par exemple du quart de la taille de la fenêtre d'analyse à chaque fois. De cette façon, on peut pouvoir suivre les fluctuations de fréquence et d'amplitude des composantes partielles du son.

Cette technique d'analyse dite « FFT à fenêtre glissante » est très efficace car elle permet, si elle est associée à une technique de synthèse, de reproduire très fidèlement tout type de son.

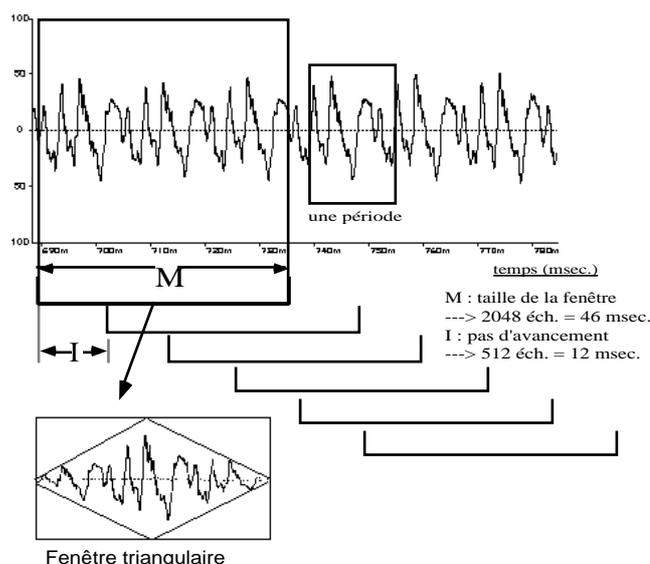


figure I-53 : répartition des zones d'analyse pour une FFT à fenêtre glissante

Le sonagramme est une représentation graphique des données de la FFT à fenêtre glissante très utile pour l'étude des sons. Emile Leipp le décrit ainsi :

« Le sonagramme est le document de choix pour l'acoustique musicale parce que cette image correspond tout à fait à l'image mentale que nous suggère un son et tout *ce qu'on entend, on le voit sur le document*. Grâce à ce document, tout son, aussi compliqué soit-il, *peut être décrit aussi bien dans le langage des physiciens que dans le langage des musiciens...* Les sonagrammes rendent compte de la réalité sonore... la précision des documents est comparable à celle de l'audition... Le sonographe est vraiment l'appareil du spécialiste en acoustique musicale lui permettant d'étudier non des artefacts de laboratoire mais la réalité musicale enregistrée sur bande ou sur disques. » [Leipp-1989, p. 89, 90, 92].

Les premiers sonographes, dont Emile Leipp situe l'invention dans les années cinquante, étaient formés d'un ensemble de filtres de fréquences croissantes pouvant produire des étincelles allant brûler une feuille de papier enroulée sur un cylindre en rotation.

Sur les ordinateurs, les sonagrammes ont gagné en précision et en simplicité de mise en œuvre. Ils représentent le temps sur l'axe des abscisses, la fréquence sur l'axe des ordonnées et indiquent l'amplitude des partiels par des tracés plus ou moins intenses.

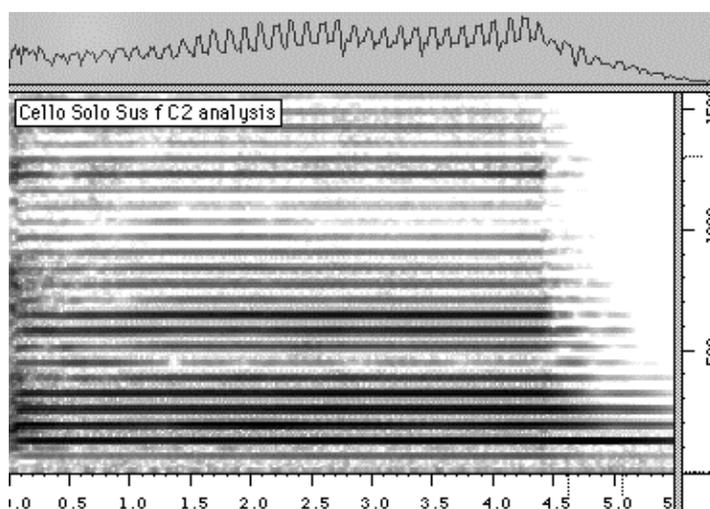


figure I-54 : enveloppe d'amplitude (en haut) et sonagramme (en bas) d'un son de violoncelle, note Do1 (exemple sonore n°3).

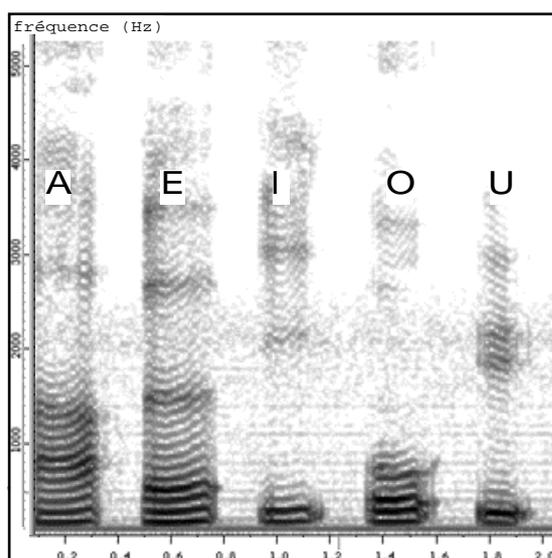


figure I-55 : sonagramme représentant les voyelles « a-e-i-o-u » parlées (exemple sonore n°5).

La figure I-54 représente le sonagramme d'un son de violoncelle. On peut observer des lignes horizontales espacées régulièrement qui correspondent aux

harmoniques. Les plus foncées sont les plus intenses. On remarque que la fondamentale, qui est le tracé le plus proche de l'axe des abscisses est de faible amplitude. On remarque également que certains harmoniques sont plus intenses que d'autres et délimitent des formants, zones de plus forte énergie, comme c'est le cas entre 100 Hz et 400 Hz ou entre 500 Hz et 700 Hz. On distingue très nettement, vers 4,5 secondes, le moment où l'archet n'entraîne plus la corde, qui résonne alors librement. Dans cette zone, les partiels aigus s'atténuent plus rapidement que les partiels graves.

La figure I-55 permet de visualiser les principales zones formantiques des voyelles.

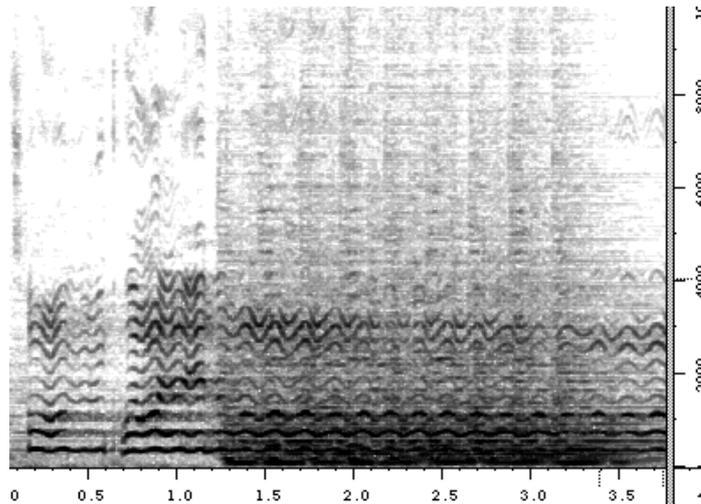


figure I-56 : sonagramme d'un extrait de la musique du film *Farinelli* (exemple sonore n°6).

Dans la figure I-56, le son extrait de la bande sonore du film « *Farinelli* » débute avec la voix chantée. On en suit très facilement les variations grâce à un vibrato très prononcé. On peut facilement trouver la fréquence fondamentale de la note jouée car les harmoniques sont très bien identifiés. Les zones formantiques de la voix apparaissent également assez nettement : entre 500 Hz et 1000 Hz d'une part et autour de 3000 Hz d'autre part. A partir de 1,2 secondes, l'orchestre fait son entrée mais la voix reste identifiable par le dessin de son vibrato. On peut suivre le rythme des attaques des cordes mais leurs hauteurs ne sont pas identifiables.

I.3.3.5. L'analyse par ondelettes

Les analyses basées sur le principe de la DFT à fenêtre glissante ne permettent pas d'analyser de façon satisfaisante des phénomènes ayant lieu simultanément quand les échelles de temps diffèrent. Le problème se pose, par exemple, pour l'analyse de la voix pour laquelle interviennent constamment deux composantes très différentes : les voyelles et les consonnes. La voyelle est un signal long assez bien localisé en fréquence et la consonne est un signal qui délivre une information riche et complexe sur des échelles de temps qui peuvent être très courtes.

Il existe des analyses qui ne privilégient aucune échelle particulière mais qui généralisent à toutes les échelles l'analyse locale des fréquences obtenues par la DFT à fenêtre glissante. C'est le cas de l'analyse par ondelettes, cas particulier des analyses dites à « Q constant ». Le facteur Q provient de la théorie des filtres et peut être assimilé au rapport entre la largeur de bande et la fréquence centrale d'un filtre. L'analyse par ondelettes est une analyse pour laquelle la largeur de bande est proportionnelle à la fréquence tandis que la FFT est une analyse à largeur de bande constante. L'analyse par ondelettes est donc une analyse dont les largeurs de bande exprimées en intervalles musicaux sont constantes (par exemple : analyse par 1/8^e d'octave).

L'utilisation de la transformée par ondelettes pour des applications en acoustique et en musique a été développée par des chercheurs de l'équipe dirigée par Jean-Claude Risset au laboratoire LMA du CNRS de Marseille-Luminy [Arfib, Kronland-Martinet-1993].

La transformée du signal en ondelettes consiste à définir une grille indiquant les découpages de l'axe du temps en fonction des découpages de l'axe des fréquences. Dans chaque cellule de la grille est placée une ondelette définie par sa position, sa durée, sa fréquence et sa phase initiale. La particularité des ondelettes est d'avoir toujours le même nombre de périodes quelle que soit la cellule considérée. La taille d'une ondelette varie donc avec sa fréquence. Il en résulte que les ondelettes de faible fréquence ont une résolution fréquentielle importante tandis que les ondelettes de fréquences élevées ont une résolution temporelle importante.

La transformée en ondelettes permet donc d'obtenir à la fois une grande précision temporelle pour des transitoires à haute fréquence et donne un spectre à haute résolution dans les basses fréquences.

Les ondelettes sont obtenues par multiplication du signal à analyser par des ondelettes de référence de différentes dimensions. On peut assimiler chaque ondelette à la réponse impulsionnelle d'un filtre résonant.

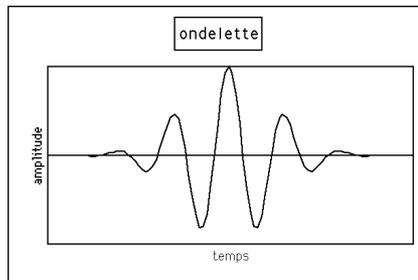


figure I-57 : ondelette de J. Morlet, $\Psi(t) = \cos(5t) \cdot e^{-t^2/2}$

I.3.4. Les techniques d'analyses spécialisées

I.3.4.1. Recherche des hauteurs perçues (algorithmes de Terhardt)

Parmi les techniques permettant de réduire et de simplifier les données contenues dans la FFT, les algorithmes mis au point par Ernest Terhardt, Gerhard Stoll et Manfred Seewa (cf. [Terhardt-1982] et Annexe III) permettent d'extraire les hauteurs significatives pour la perception qui sont présentes à l'intérieur d'un son complexe.

Les algorithmes utilisés par Ernest Terhardt proviennent de diverses études psychoacoustiques. Ils contiennent des traductions mathématiques des effets de masquage que peuvent exercer des partiels de fréquences voisines les uns sur les autres. Par ailleurs, ils permettent de déterminer quels partiels sont significatifs pour la perception et attribuent à ces partiels un coefficient appelé poids perceptif.



figure I-58 : principales étapes du calcul pour l'algorithme de Terhardt

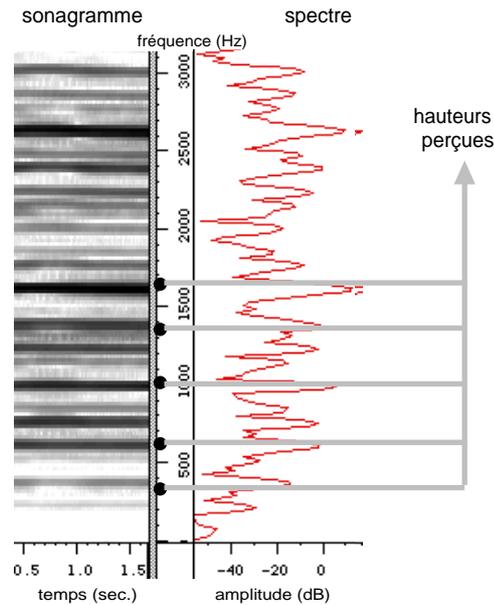


figure I-59 : hauteurs perçues obtenues par l'analyse de Terhardt sur un son multiphonique de saxophone (exemple sonore n°7).

I.3.4.2. La recherche de la fondamentale

Nous avons indiqué au paragraphe 1.3.1.1. que la hauteur d'un son était parfois difficile à déterminer avec précision.

Par ailleurs, les microvariations de la fondamentale d'un son musical sont souvent extrêmement significatives de la qualité du son. Sans l'aide de l'informatique, ces variations sont impossibles à caractériser.

Il existe sur ordinateur des techniques de détection de la fréquence fondamentale qui ont été mises au point pour l'analyse des sons monophoniques.

Ces techniques sont généralement réalisées à partir d'une FFT, suivie d'une recherche de pics. La détection de la fondamentale est effectuée en recherchant la fréquence du plus petit commun diviseur des pics les plus intenses. La fondamentale retenue sera la fréquence dont la somme des amplitudes des harmoniques sera la plus élevée. Des algorithmes auxiliaires, liés à la perception, peuvent être également mis en œuvre. Ceux-ci peuvent prendre en compte les effets de masquage qu'une fréquence peut exercer sur une fréquence voisine.

L'exemple suivant est extrait d'une recherche de fondamentale sur un son de voix chantée asiatique présentant d'importantes variations de hauteur. La première colonne indique, en secondes, les dates des analyses et la seconde colonne indique les variations de la fondamentale exprimée en Hz.

0.030000	231.154327
0.040000	169.978409
0.050000	174.266403
0.060000	174.316345
0.070000	174.316345
0.080000	177.691299
0.090000	176.469315
0.100000	177.164246
0.110000	177.164246
0.120000	177.691299
[...]	
0.200000	188.133743
....	

Une visualisation de ces résultats est donnée dans le graphe ci-dessous.

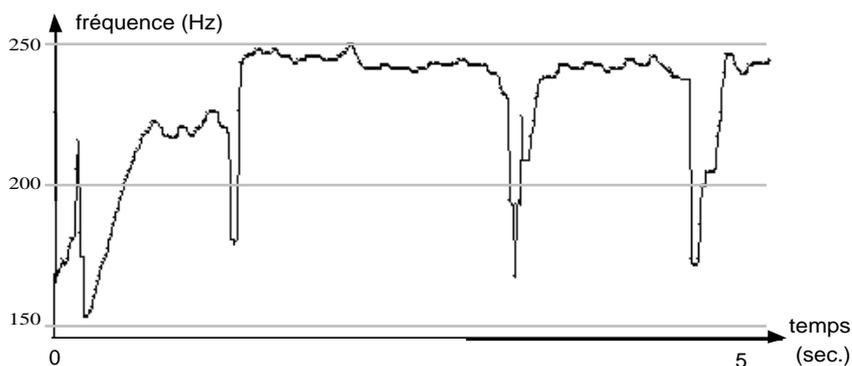


figure I-60 : représentation graphique de la fondamentale d'une voix asiatique(ex. sonore 9a).

I.3.4.3. L'analyse par prédiction linéaire

Une des techniques « générales » les plus performantes pour l'estimation d'enveloppes spectrales est la prédiction linéaire (LPC). Elle présuppose qu'il est possible de construire un filtre dont la réponse en fréquences sera similaire à l'enveloppe spectrale du son. Ce filtre pourra être appliqué à une excitation de forme d'onde arbitraire de façon à lui imposer la même enveloppe spectrale que celle du son analysé.

En réalisant des analyses successives sur le son, on peut construire un ensemble de filtres variant dans le temps.

L'analyse en prédiction linéaire est basée sur les équations des filtres numériques et la « transformée en z » du signal.

La LPC, ou « Linear Predictive Coding », est basée sur le fait que la plupart des signaux acoustiques comportent une information fortement redondante. En analysant le signal sur N échantillons, on peut le remplacer par un ensemble de M filtres et un signal $x(n)$ à filtrer.

Les avantages de la LPC tiennent dans le fait que le signal $x(n)$ peut être un signal très simple. Dans le cas de la voix par exemple, $x(n)$ peut être un train d'impulsions pour des syllabes voisées ou du bruit blanc pour des syllabes non voisées.

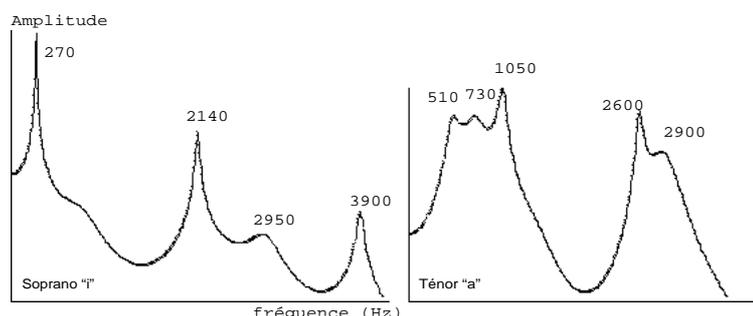


figure I-61 : analyse en LPC de sons de voix chantée, à gauche voix de soprano - voyelle « i », à droite voix de basse – voyelle « a ».

La LPC est la principale technique utilisée pour l'analyse des sons de voix.

I.3.4.4. L'analyse harmonique additive

Plusieurs outils permettent l'analyse et la resynthèse de sons à caractère harmonique. C'est le cas de la technique d'analyse et resynthèse « Additive » mise au point par l'équipe de Xavier Rodet à l'Ircam.

L'analyse Additive est une analyse qui débute par une FFT à fenêtre glissante. Les données fournies par la FFT sont traitées pour la recherche du trajet de la fondamentale. Les harmoniques f_i de cette fondamentale sont ensuite recherchés par sélection du pic maximum dans chaque bande de fréquences comprise entre $(i-c).f$ et $(i+c).f$. f représente la fréquence de la fondamentale, i représente le numéro de l'harmonique, c représente le seuil d'harmonicité de l'analyse et est généralement compris entre 0 et 0,5. Une valeur supérieure à 0,5 peut provoquer un saut au partiel supérieur.

Lorsque le niveau de bruit empêche la détection d'un pic dans une zone de fréquence, l'harmonique correspondant peut être masqué et ne pas apparaître dans les résultats.

Les données de l'analyse peuvent être utilisées pour effectuer une synthèse à l'aide d'un synthétiseur additif afin de produire un fichier-son. Ce dernier peut ensuite être soustrait du son original et fournir par différence la partie bruitée du son original qui n'a pas pu être décelée par l'analyse additive.

Pour des sons quasi-harmoniques, il est possible d'obtenir des resynthèses très proches des sons originaux pour la plupart des instruments de musique. La fraction résiduelle est alors très réduite. On y retrouve principalement quelques transitoires.

Les résultats sont moins bons lorsque le son à analyser comporte des parties bruitées ou présente de brusques variations d'intensité.

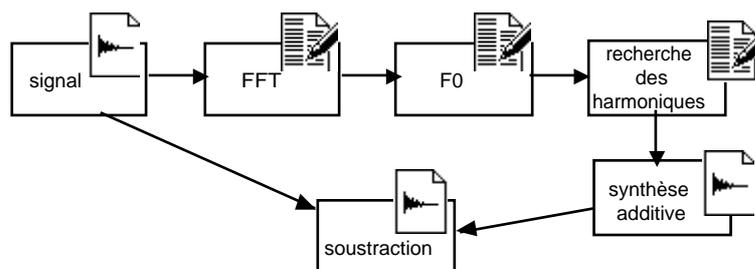


figure I-62 : les différentes étapes de l'analyse/synthèse par la méthode « Additive ».

Les résultats de l'analyse sont présentés au format suivant : chaque analyse élémentaire commence par un en-tête qui indique le nombre d'harmoniques trouvés, puis donne la date de cette analyse en secondes. On trouve ensuite le tableau des données de cette analyse élémentaire avec dans la première colonne, les rangs des harmoniques, dans la deuxième, leurs fréquences (en Hz), dans la troisième, leurs amplitudes linéaires et dans la quatrième, leurs phases.

nb de partiels	date	phases
n° des partiels	fréqs	amps
20	0.020000	
1	320.402222	0.174731
2	638.924133	0.120691
3	957.553162	0.078010
4	1269.705811	0.038462
5	1591.073608	0.066457
6	1911.491211	0.005330
[...]		
18	5771.773926	0.000235
19	6191.083496	0.000636
20	6373.922852	0.001586
20	0.030000	
1	320.402222	0.262096
2	638.924133	0.181036
3	957.553162	0.117015
4	1269.705811	0.057693
5	1591.073608	0.099685

tableau I-8 : format d'affichage des données produites par l'analyse « Additive ».

Le tableau I-8 reproduit un extrait de l'analyse d'un son de voix chantée asiatique, dont la fondamentale a été décrite précédemment. Il est limité aux vingt premiers partiels.

La figure I-63 représente les trajets de fréquence des harmoniques en fonction du temps.

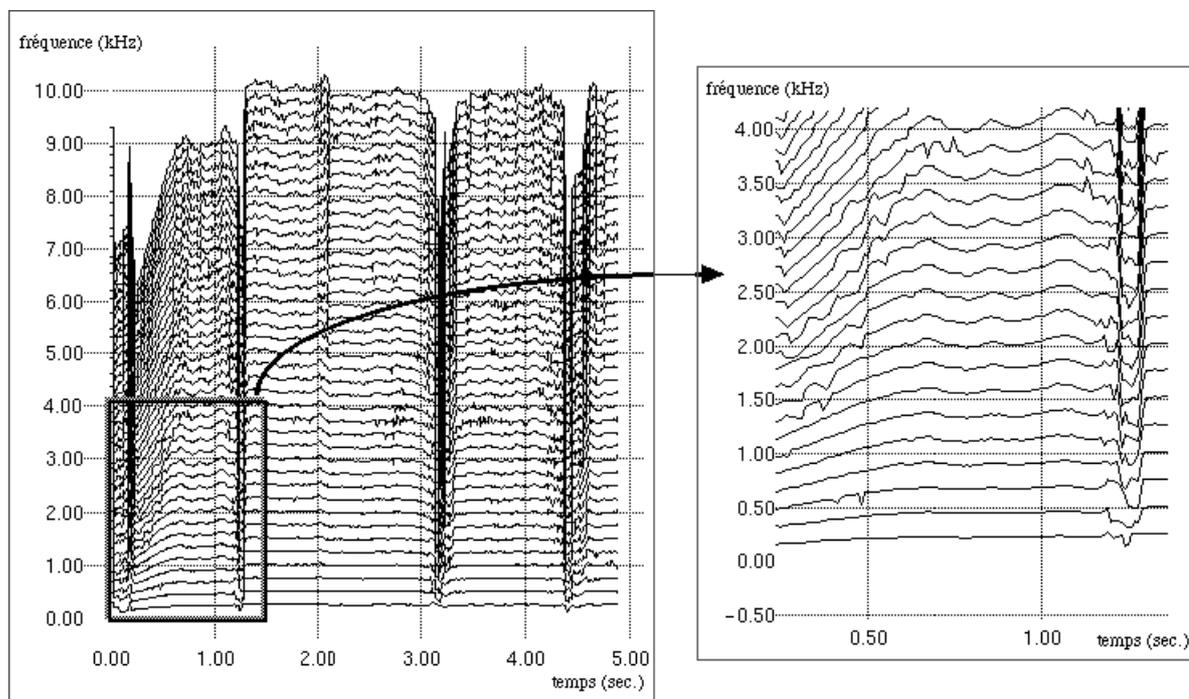


figure I-63 : représentation graphique des trajets des partiels harmoniques calculés par l'analyse Additive sur une voix asiatique (exemple sonore n°9).

L'indice d'inharmonicité, qui est un des paramètres de l'analyse, permet de déceler des composants non harmoniques. Par ailleurs, plus la taille de la fenêtre est petite et, surtout, plus le recouvrement des analyses élémentaires est important et plus l'analyse décèle les variations rapides des composants.

I.3.4.5. L'analyse par « modèles de résonance »

L'analyse par « modèles de résonance » est une technique qui a été mise au point à l'Ircam par l'équipe de Jean-Baptiste Barrière entre 1984 et 1986 [Potard, Baisnée, Barrière-1991]. Elle permet de construire des modèles d'instruments offrant des resynthèses de très bonne qualité malgré une réduction très importante des données.

Le principe de cette modélisation consiste à identifier tous les modes de résonance d'un instrument, de façon à produire un filtre dont la fonction de transfert est caractéristique de cet instrument. Ce filtre est formé par la superposition d'un ensemble de résonateurs élémentaires.

Pour cela, on effectue une FFT sur une fenêtre du son, suivie d'une extraction de pics. En comparant deux FFT de positions décalées dans le son on peut déceler des résonances, pics présents dans les deux analyses et mesurer leur taux d'atténuation d'une analyse sur l'autre.

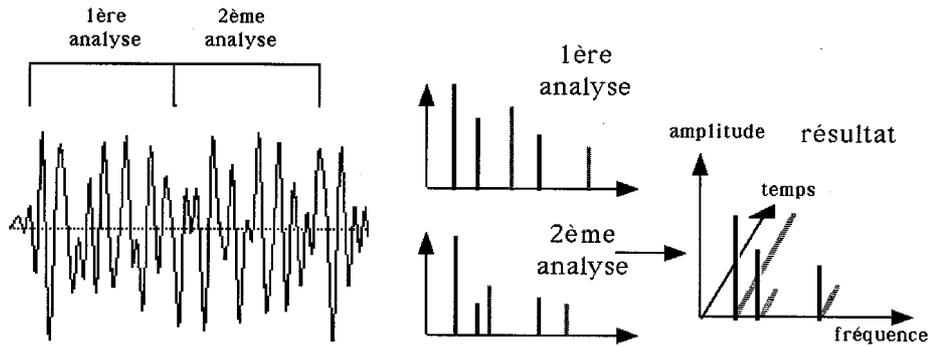


figure I-64 : comparaison de deux FFT suivies de recherches de pics pour trouver des résonances.

L'opération de détection des résonances est répétée pour des tailles de fenêtres croissantes afin d'augmenter progressivement la résolution en fréquence de l'analyse. Les nouvelles résonances trouvées viennent s'ajouter aux résonances précédemment obtenues.

Les premières analyses, réalisées sur des fenêtres de petites tailles permettent de déceler les résonances à amortissement rapide dans le temps tandis que les analyses suivantes permettent de déceler une gamme de résonances de plus en plus étendue, en particulier vers les graves. L'analyse s'effectue par étapes successives en conservant en mémoire les résultats obtenus d'une analyse sur l'autre.

Sur les premières étapes de l'analyse et selon la richesse spectrale du son, on détecte souvent une dizaine de résonances. A la fin de l'analyse, on obtient de une à plusieurs centaines de partiels.

Le résultat de cet ensemble d'analyses fournit une liste de résonances caractérisées par leur fréquence f_i , leur amplitude A_i et leur largeur de bande bw_i .

Une réduction peut être appliquée sur les données de l'analyse. Elle porte sur des critères d'amplitude, de fréquence ou de largeur de bande. Elle permet de supprimer de la liste des résonances celles qui n'ont pas une amplitude suffisante ou qui seraient masquées par des résonances voisines.

Les durées de résonance sont inversement proportionnelles aux largeurs de bande. En effet, l'amplitude d'une vibration en oscillation libre amortie - pour un amortissement visqueux (cf. paragraphe I-2-2-1) - que nous appelons ici résonance, est donnée par la relation suivante :

$$Amp(t) = A_0 \cdot e^{-(bw \cdot t)}$$

La durée de résonance représente le temps mis pour passer de l'amplitude A_0 à l'amplitude A_0/n . Elle est donnée par la formule :

$$t = \log(n)/bw$$

Le tableau suivant donne les résultats obtenus sur un son de marimba de hauteur Do3.

no	freq	ampl	band
1	52.49763	-48.39835	1.685818
2	263.0164	-6.122798	1.957626
3	612.6320	-18.98892	12.18928
4	860.4391	-54.03692	.9343989
5	1048.307	-11.47423	5.908646
6	1242.274	-57.52631	2.08433
7	1316.117	-50.68672	1.993205
8	1526.048	-59.30801	2.301462
9	1604.516	-56.30065	2.968727
.....			

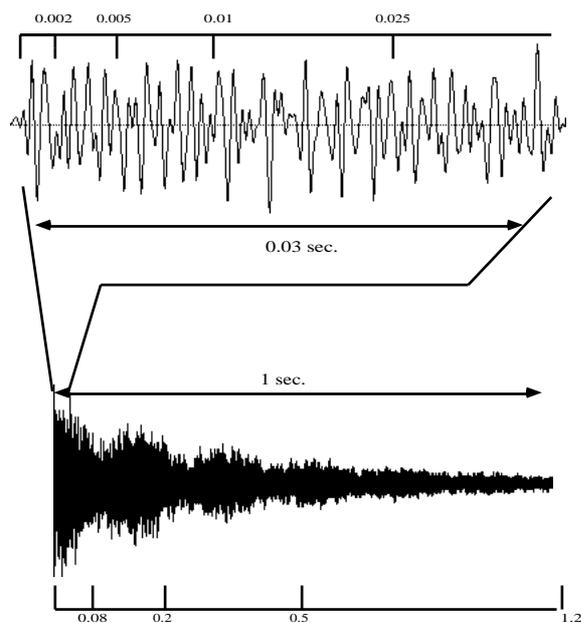


figure I-65 : différentes tailles de fenêtre utilisées pour le découpage du son pour une analyse par « modèles de résonance ».

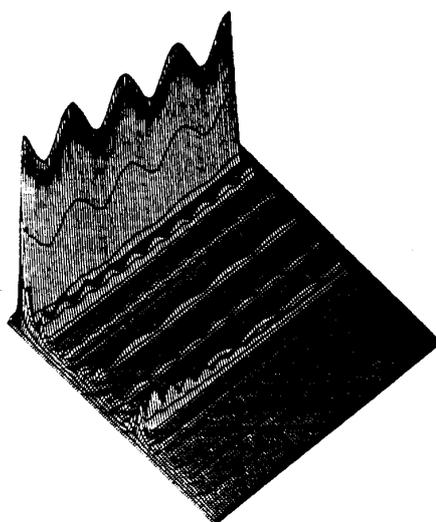


figure I-66 : son de cloche à vache simulé par resynthèse par modèle de résonance [Potard, Baisnée, Barrière-1991, p. 154].

Les données de l'analyse par modèles de résonance étaient à l'origine destinées à être utilisées par le synthétiseur CHANT pour la synthèse de sons percussifs en employant des Fof ou par la synthèse par filtrage.

I.3.4.6. L'analyse modale

A la description des phénomènes sonores, ou description du signal, on peut parfois substituer une description « mécanico-acoustique » des instruments qui produisent le son. Le son est alors considéré comme l'ensemble des vibrations des structures qui composent l'instrument. Ce type d'analyse vise à modéliser de façon exhaustive ces structures ainsi que les interactions de leur assemblage.

Les connaissances en acoustique physique sont actuellement assez précises pour permettre la modélisation de la plus grande partie des instruments de musique par analyse modale.

Une structure pourvue d'un nombre fini de degrés de liberté est caractérisée par un ensemble de fréquences de résonance et de coefficients d'absorption ainsi

que par une matrice de déformées modales. Les colonnes de cette matrice sont formées par les coordonnées des déformées dans la base des coordonnées principales. Cette caractérisation se limite aux comportements linéaires des structures mais des comportements non linéaires peuvent être introduits aux points de jonction entre des sous-structures.

Les techniques d'analyse modale ont connu d'importants développements dans les années 80 et leurs résultats sont directement utilisables dans le cas des structures mécaniques. Ils permettent de caractériser les systèmes acoustiques à partir de leurs propriétés géométriques.

Cette technique d'analyse est utilisée pour la synthèse par modèles physiques dans le programme Modalys mis au point par Jean-Marie Adrien en 1990 [Adrien-1990].

I.4. Principes de production de la voix chantée

Le point commun entre tous les instruments de musique est qu'ils convertissent de l'énergie, fournie sous une forme ou sous une autre, en énergie acoustique, c'est-à-dire en des oscillations régulières de la pression sonore.

Pour les instruments à vent et pour la voix, c'est un flux d'air qui est à l'origine de la production sonore. La conversion du flux d'air en son s'effectue le plus souvent sous la forme d'une mise en oscillation : le flux est interrompu puis reprend sous l'effet d'une valve qui s'ouvre et se referme à intervalles réguliers. Ceci est le cas de la voix chantée. Parfois la production peut s'effectuer de manière différente. Le flux peut être fourni sous une forme laminaire puis être mis en oscillation. Le hautbois et la flûte sont des exemples de ce type.

Un flux d'air pulsé génère un son complexe composé d'une série de partiels harmoniques. Le timbre du son produit dépend du résonateur formé par le corps de l'instrument dont la forme détermine les fréquences des résonance et leurs largeurs de bande.

Il existe des instruments pour lesquels les phénomènes d'émission et de résonance sont indépendants. En ce qui concerne la voix, il n'y a pratiquement aucune relation entre les caractéristiques de l'oscillation émise par les cordes vocales et le résonateur. Dans ce cas, le résonateur joue un rôle très important comme filtre. Lors du passage du son dans le résonateur, les amplitudes relatives de ses partiels sont modifiées. Si un partiel est proche des fréquences caractéristiques du résonateur, il est émis avec une forte amplitude, sinon il est atténué.

Lors de la production du son, les cordes vocales agissent comme un oscillateur, tandis que le larynx, le pharynx et la bouche, parfois complétés par le conduit nasal, agissent comme des filtres.

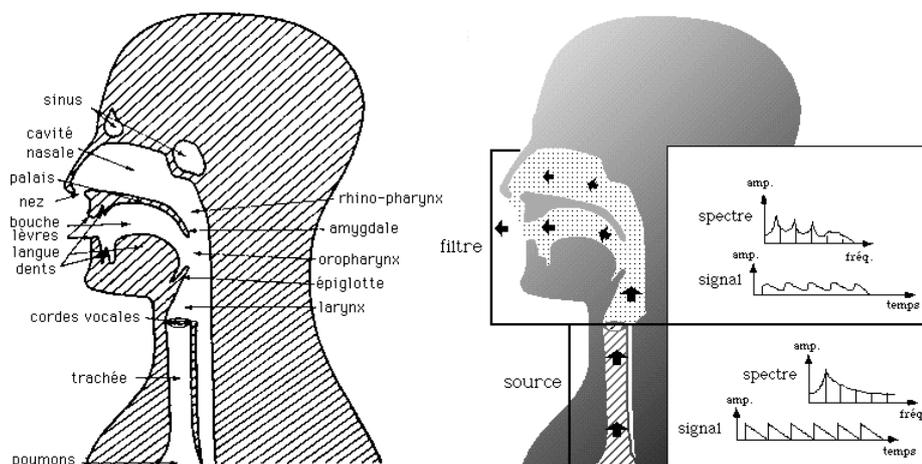


figure I-67 : filtrage du son émis par la glotte

I.4.1. La source

Le flux d'air envoyé par les poumons passe par la glotte en faisant vibrer les cordes vocales : c'est la source sonore. Son spectre est harmonique et comprend de nombreux partiels dont les amplitudes relatives peuvent varier selon la hauteur et la force de la voix.

Le registre de la fondamentale dépend de la longueur et de la masse des cordes vocales, plus les cordes sont longues et fines plus la fondamentale est élevée. En général, la fondamentale pour la voix parlée d'un homme adulte se situe aux alentours de 110 Hz et un peu moins d'une octave au-dessus pour une femme adulte. Un chanteur basse peut descendre jusqu'à 65 Hz (Do1) et monter jusqu'à 330 Hz (Mi3). L'étendue d'une voix de ténor est comprise entre 123 Hz et 520 Hz, celle d'une alto entre 175 Hz et 700 Hz et celle d'une soprano entre 260 Hz et 1300 Hz.

La hauteur des notes n'est pas uniquement dépendante des cordes vocales. La pression de l'air influe également. En s'intensifiant, elle peut augmenter le registre de la fondamentale, ce qui explique pourquoi certains chanteurs ne peuvent chanter des notes aiguës doucement.

La force de la voix est essentiellement contrôlée par la pression de la colonne d'air.

L'amplitude relative de la fondamentale dans le spectre peut également être modifiée en fonction de la position des cordes vocales par rapport à la glotte.

Pour une voix normale chantée, le spectre harmonique est constitué de partiels qui suivent une décroissance de 12 dB/octave. En règle générale, une augmentation du volume sonore s'accompagne d'une augmentation relativement importante des partiels élevés [Sunberg-1991, p. 144].

I.4.2. Les formants

La source se propage à travers le conduit vocal qui agit comme un résonateur et filtre le son avant de l'émettre. Il en résulte l'apparition de formants qui ont une grande influence sur les sons de la voix : les partiels du « son-source » qui sont placés dans la région d'un formant sont amplifiés par rapport aux autres. Les formants apparaissent comme de larges pics dans l'enveloppe spectrale du son produit.

Chaque voyelle présente un spectre caractéristique dépendant du type de voix (cf. figure I-61). En modifiant la forme du conduit vocal, un homme adulte peut faire varier la fréquence du premier formant entre 150 Hz et 900 Hz, celle du second entre 500 Hz et 3000 Hz et celle du troisième entre 1500 Hz et 4500 Hz. Dans la plupart des cas, ce sont les valeurs des deux premiers formants qui déterminent la nature de la voyelle.

Les consonnes contiennent également des formants mais ceux-ci jouent un rôle différent de ceux des voyelles. Pour les consonnes, les caractéristiques du son proviennent des changements particuliers des formants. Pour la plupart des consonnes, les trois formants les plus graves présentent des évolutions permettant leur reconnaissance et leur identification.

Un trait caractéristique de la voix des chanteurs basse, ténor et alto est la présence d'un pic particulièrement marqué dans la partie aiguë du spectre, situé entre 2 kHz et 3 kHz. Il proviendrait de la fusion des 3^e, 4^e et 5^e formants. Ce formant est appelé le formant des chanteurs. Il permet de rendre la voix plus distincte car il est placé dans un registre où l'oreille est particulièrement sensible et où l'orchestre est moins présent.

Pour les voix de soprano, la technique de chant consiste à faire coïncider la fréquence de la fondamentale avec celle du premier formant de façon à permettre son amplification.

En général, lorsque les chanteurs atteignent les aigus et qu'ils doivent choisir entre la qualité de la voyelle et la qualité de la voix, ils sacrifient les voyelles et placent le premier formant sur la fondamentale puis le deuxième formant sur le deuxième partiel.

I.5. Conclusion sur la nature du son

Les premières utilisations des théories de l'acoustique appliquées à la synthèse sonore ont parfois été décevantes. Par exemple, on s'est vite aperçu que la liste des harmoniques d'un son ne suffisait pas à donner une description satisfaisante de ce son et qu'il faut utiliser des lois dynamiques, plus complexes, pour obtenir des résultats satisfaisants. Le timbre est la résultante de relations fonctionnelles entre de multiples facteurs.

« L'acoustique issue de l'informatique parvient à des conclusions inverses de celles auxquelles tendaient la plupart des acousticiens du XIX^e siècle : elle abolit la conception de l'onde à spectre fixe.[...] On sait maintenant que c'est l'évolution d'ensemble de l'onde sonore qui est significative, et non seulement un transitoire initial précédant une phase stationnaire » [Dufourt-1991a, p. 331]

Les informations apportées par l'acoustique pour aider à la synthèse sont généralement employées de deux façons. La première voie consiste à étudier les modes de vibration des objets physiques donnant lieu à la production de sons pour tenter de les reproduire à l'aide de programmes informatiques. Elle donne lieu à la synthèse dite par « **modèle physique** ». Cette technique permet la synthèse de sons réalistes. Par contre, elle restreint le champ des possibilités de la synthèse à des sons obéissant aux lois de l'acoustique physique, alors qu'un des avantages de la synthèse numérique est justement de pouvoir s'affranchir de ces contraintes.

La seconde méthode consiste à analyser les sons produits par différentes sources sonores et à en tirer des règles utiles à la synthèse. Elle correspond à la synthèse par « **modèle de signal** ». La quantité de données nécessaire à la reconstitution d'un son réaliste par analyse/synthèse est considérable, ce qui rend les manipulations difficiles. Il est donc impératif de regrouper les données en fonction de critères ayant une signification pour la perception auditive. Les recherches effectuées dans le domaine de la psycho-acoustique tendent à définir de tels critères et à les classer en fonction de leur pertinence.

Les analyses des propriétés microtemporelles du son, effectuées par les scientifiques, portent souvent sur des événements sonores simples comme une note de musique. Sur des sons instrumentaux, ces propriétés sont liées d'une part à la nature de la structure résonante qui entre en vibration et d'autre part au moyen d'excitation utilisé. Mais à quelles caractéristiques du son peut-on attribuer les qualificatifs que nous employons couramment pour parler d'un son et le décrire ?

Les recherches de Jean-Baptiste Barrière sur la simulation des instruments et de la voix en synthèse des sons l'ont amené à définir des « bibliothèques de structures conceptuelles » formées de couples de concepts dont les éléments sont placés en opposition [Barrière-1990b, p. 63]. Les quelque 250 termes qu'il a choisis sont autant de points de départ servant au compositeur à définir des chemins à parcourir.

Par ailleurs, le tableau I-9 regroupe les qualificatifs employés par Eric Mauer pour décrire les sons produits par les instruments à vent et montre à quel point les termes peuvent être variés.

Aéré	Ample	Aspiré	Assourdi	Avec écho
Bouché	Bouillonnant	Bourdonnant	Brillant	Bruyant
Calmant	Cannelé, strié	Cassant	Caverneux	Chantant
Chaud	Choqué	Chuintant	Clair	Claquant
Coup de langue	Coupant	Courbe	Dense	Discordant
Doux	Dur	Dynamique	Elastique	Elevé, sublime
En gros	Épais	Éthéré	Étouffé	Étroit
Faible (niveau)	Fêlé	Fluet	Flûté	Fondu
Fort (niveau)	Fracassant	Frappé	Frotté	Goutte à goutte
Gras	Gratouillé	Grattant	Grinçant	Grondant
Guttural	Haché	Haut	Humide	Hurlant
Impulsif	Indolent	Infléchi	Large	Lié
Lisse	Lointain	Lourd	Louré	Lumineux
Martelé	Mélodieux	Métallique	Mielleux	Mobile
Moelleux, doux	Mordant	Mou	Nasal	Nasillard
Ondulé	Perçant	Perçant aigu	Percussif	Pincé
Piqué, staccato	Plat	Plein	Pleurant	Pointu
Propre	Pur	Raquant	Râpeux	Rauque
Résonnant	Retentissant, net	Réverbéré	Rond	Rugueux
Sanglotant	Sec	Secoué	Sifflé	Solide
Sombre	Sonnant	Sonore	Soufflant	Soufflé
Strident	Surnaturel	Suspendu	Sustain, maintenu	Tendu
Tintant	Tintinnabulant	Tremblotant	Vapoureux	Velouté
Vibrant	Vigoureux	Violent, détonnant	Voyé, Vocal	

tableau I-9 : qualificatifs employés par Eric Mauer pour une classification des sons des instruments à vent

Les travaux de John Grey [Grey-1977] ont mis en évidence trois paramètres primordiaux permettant de comparer des sons instrumentaux : **la brillance** du son qui est proportionnelle au barycentre de l'enveloppe spectrale, **le flux spectral** indiquant le degré de fluctuation de l'enveloppe spectrale pendant la durée du son et **le taux de synchronisme des attaques des différents harmoniques**.

Pour la reconnaissance d'un son instrumental, la phase la plus importante est généralement l'attaque. C'est une phase transitoire qui caractérise à la fois l'excitation et la mise en vibration de l'instrument.

Dans la partie stable d'un son, le vibrato joue un rôle prépondérant pour l'identification du son (cf. § III.3). Max Mathews a montré que sur les violons, le vibrato s'accompagne d'une modulation spectrale qui leur confère une qualité sonore très particulière [Roads-1989a]. Dans le cas de la synthèse de la voix chantée, c'est lorsque l'on utilise un vibrato sur la fondamentale que le son acquiert un caractère vocal par dissociation entre l'excitateur et le résonateur et la mise en évidence des formants. L'oreille étant très sensible aux aspects fréquentiels des sons, c'est la cohérence des partiels composant un son qui permet la séparation des sons joués à l'unisson par des instruments différents. Les travaux de Stephen MacAdams ont montré qu'en plaçant un vibrato sur les partiels impairs d'un son harmonique, on entend deux sons indépendants [McAdams-1985, p. 14].

Pour la plupart des sons instrumentaux entretenus, l'extinction du son est une phase de courte durée et peu significative. Dans le cas des sons percussifs, par contre, elle représente pratiquement la totalité de la durée du son. La technique d'analyse/synthèse par « Modèle de résonance » permet de décrire et de synthétiser de nombreux types de sons de façon satisfaisante après avoir calculé leurs modes de résonance.

Un des points importants qui donne une identité au son musical est la façon dont le timbre varie en permanence en fonction de son intensité et de sa hauteur.

Les modes d'articulation, c'est-à-dire les transitions entre les différentes notes successives produites par l'instrument, tiennent une place importante dans la reconnaissance des instruments.

Les sons acoustiques que nous entendons contiennent des informations relatives à la position spatiale de la source qui les a émis. La détection de l'éloignement d'une source sonore est liée à l'intensité perçue ainsi qu'aux variations du spectre et du taux de réverbération. La nature du lieu modifie les qualités de la réverbération qui colorent le son. Enfin, la détection des mouvements des sources sonores est liée à la reconnaissance d'indices témoignant des variations de position, de distance et d'effet Doppler. De nombreux dispositifs de spatialisation du son permettent de simuler les mouvements de sons de synthèse [Jullien, Jot-1995] et [Pottier, Stalla-2000].

En ce qui concerne les recherches actuelles sur l'étude du son, on remarque deux démarches. D'un côté, des scientifiques analysent les sons pour mieux les comprendre et en extraire des informations pertinentes pour des activités de synthèse sonore. De l'autre, ils analysent les sons afin d'en extraire les informations non pertinentes pour réaliser des algorithmes de compression. A l'heure actuelle, une énorme poussée commerciale, liée au développement des échanges par réseau, tend à favoriser la deuxième voie. Certains chercheurs œuvrent pour les deux démarches en proposant des algorithmes de compression du son basés sur des techniques de synthèse. C'est le cas des recherches menées au MIT [Wright, Scheirer—1999] pour le développement de la norme de codage multimédia MPEG4 basée sur la synthèse en utilisant le synthétiseur SAOL (Structured Audio Orchestra Language).

Les algorithmes de détection de fondamentales ont eu pour principal objectif l'analyse de phrases musicales pour réaliser l'écriture automatique de partition. Maintenant, la tendance est de rechercher des analyses capables d'écrire une partition à partir d'enregistrements polyphoniques. Ces analyses passent par des techniques de suivi de partiels et des algorithmes d'apprentissages de styles musicaux.

Le problème de la caractérisation du timbre est un sujet complexe sur lequel les chercheurs n'ont pas fini de s'interroger. Il semble aujourd'hui indispensable d'avancer dans ce domaine pour permettre de mieux utiliser les outils qui servent à la synthèse sonore et pour mettre au point des codes servant à décrire les sons de synthèse de façon à pouvoir transmettre le savoir. Notre connaissance des sons instrumentaux harmoniques progresse mais pour les autres, elle est encore trop parcellaire. Les analyses en ondelettes qui semblaient être une solution au problème de l'antagonisme temps/fréquence, n'ont encore rien résolu. Les sons de feu, d'eau, de verre qui se brise et la plupart des sons non harmoniques très mouvementés restent encore totalement irréductibles à quelques paramètres et à une équation.

I.	LA NATURE DU SON MUSICAL-----	11
I.1.	Introduction-----	11
I.2.	L'acoustique physique -----	11
	I.2.1. Historique.....	11
	I.2.2. Les mécanismes mis au cause pour la production du son	17
I.3.	Les techniques d'analyse des sons -----	31
	I.3.1. Les grandeurs physiques du son	32
	I.3.2. Les représentations de l'amplitude du son par rapport au temps	35
	I.3.3. Les représentations spectrales du son.....	38
	I.3.4. Les techniques d'analyses spécialisées	49
I.4.	Principes de production de la voix chantée -----	56
	I.4.1. La source	57
	I.4.2. Les formants.....	57
I.5.	Conclusion sur la nature du son -----	58

