**Discrétisation de l'espace sonore : construction d'une gamme.**

Les échelles sonores obéissent à un principe de symétrie géométrique et non arithmétique.

L'étendue du spectre des fréquences audibles est énorme, il s'étend typiquement entre 16 et 20000 Hz, soit un rapport entre les extrême (ambitus) valant 1250 environ. En musique il est rare qu'on flirte avec les fréquences les plus élevées : l'orgue fait à peine exception en couvrant la plage allant de 16.35 (do0) à 15804.27 Hz (si9) mais le plus grand piano de concert, de Stuart & Sons, "ne" comporte "que" 102 touches, allant de 16.35 Hz (do0) à 5587.65 Hz (fa8). Si l'on déroule ce spectre (spectrepiano), on réalise qu'il est inutile de dépasser ces limites.

Il n'est pas question de traiter continûment ce spectre, une discrétisation s'impose à plusieurs titres :

Deux fréquences ne différant que de moins de 1% sont indiscernables pour l'immense majorité des auditeurs, il est donc inutile de s'embêter à prévoir qu'un instrument puisse jouer n'importe quelle fréquence, reste à savoir lesquelles.

Le principe de toute musique évoluée est de faire jouer ensemble des instruments différents qui doivent s'accorder sur les fréquences des sons qu'ils émettent sous peine de cacophonie. L'un des principes de l'harmonie est que, pour être consonantes, ces fréquences doivent être dans des rapports rationnels simples 2/1 (octave), 3/2 (quinte), 4/3 (quarte), 5/4 (tierce majeure), 5/3 (sixte), etc. Cela n'est possible que si les instruments sont préalablement accordés selon un schéma universel. La construction d'une gamme acceptable par tous a occupé les mathématiciens (Euler, d'Alembert, …), les acousticiens (Sauveur, Mersenne, …) et rassurons-nous les musiciens du Siècle des Lumières (Rameau, Bach, …). Deux exigences hélas incompatibles doivent être satisfaites : la consonance et la transposabilité.

La construction d'une échelle sonore maximalement consonante dans la plage 16-5600 Hz (pour fixer les idées) est un problème d'arithmétique. Voyons comment le résoudre lorsqu'on se limite aux 5 rapports, octave, quinte, quarte, tierce majeure et sixte : on part de la fréquence la plus basse, f0, que l'on assimile à la première note de la gamme en construction, par exemple f0 = 16.35 Hz mais la valeur précise importe peu. On impose alors que de nouvelles notes font partie de la gamme, celles dont les fréquences (inférieures à 5600 Hz) valent f0 multiplié par 2/1, 3/2, 4/3 et 5/4. Si on se limite à un ambitus de 350 (celui qui correspond à la plage 16-5600 Hz) on se retrouve avec une gamme comprenant 64 notes. L'opération n'est pas terminée pour autant car il faut recommencer itérativement le même travail sur base de chacune des nouvelles notes obtenues à l'étape précédente. On pourrait penser qu'il s'agit d'un travail sans fin et que le nombre des notes va exploser à l'infini mais il n'en est rien : une première itération fait passer de 64 à 643 notes puis une deuxième fait passer de 643 à 1554 notes. Toute nouvelle itération n'introduit aucune note nouvelle. De plus l'intervalle de sixte est automatiquement respecté puisque 5/3 = 5/4 x 4/3.

La solution trouvée est satisfaisante au plan harmonique : à chaque note de la gamme correspondent une quinte, une quarte, etc, effectivement présentes dans la gamme. Pourtant elle ne convient pas pour plusieurs raisons :

Elle ne respecte pas les intervalles plus exotiques (tierce mineure = 6/5) pour lesquelles une extension de la gamme s'imposerait

Le nombre des notes est exagéré et surtout elles sont très mal réparties s'accumulant inutilement aux hautes fréquences là où elles sont souvent indiscernables à l'oreille. Par exemple les notes 1403 à 1410 ont des fréquences *f0* trop voisines pour être distinguées.

Mais l'objection majeure et pour ainsi dire rédhibitoire est qu'elle complique le problème de la transposition au point de la rendre impossible du moins en toute généralité.

Transposer une phrase musicale, c'est la translater en fréquence vers le grave ou vers l'aigu sans en altérer la ligne mélodique. Qu'une telle transposition soit possible apparaît clairement dans le cas particulier où l'on double les fréquences de chaque note. Un son de fréquence 2f correspond à une vibration des molécules d'air qui reprennent leur position initiales 2f fois par seconde donc forcément aussi f fois par seconde car qui peut le plus peut le moins. Autrement dit un son de fréquence 2f possède aussi la fréquence f d'où une mélodie dont toutes les fréquences sont doublées sonnera comme la mélodie simple sauf qu'elle sera plus aigüe. Cet exemple simple indique que la transposition est de nature géométrique et non arithmétique : ajouter un terme constant aux fréquences des notes d'une mélodie la dénature complètement; ce qu'il faut faire c'est multiplier les fréquences des notes par un facteur constant.

Certes la transposition d'une octave (ou d'une quinte, d'une quarte ou d'une tierce majeure) est possible puisque la gamme a été construite pour qu'il en soit ainsi mais cela cesse d'être vrai pour toute autre transposition.

Si dans la gamme construite précédemment on multiplie n'importe quelle note par le facteur qu'il faut pour la transposer vers une autre note de la gamme, les autres notes ne tomberont plus exactement sur une note existante brisant la parfaite harmonie qui y régnait.

) et, à titre d'exemple, telle qu'une discrétisation s'imposeoreille humaine intacte est sensible aux fréquences comprises entre 16 et 20000 Hz. Cet intervalle est énorme, ses extrêmes étant dans un rapport 1250. En comparaison, l'œil n'est sensible qu'à un intervalle beaucoup plus restreint de fréquences lumineuses, typiquement comprises entre 4.0 1014 et 7.5 1014 Hz, soit un rapport inférieur à 2 !

Il est inconcevable et d'ailleurs inutile de vouloir utiliser l'infinité des notes présentes dans ce continuum. Une discrétisation est non seulement possible mais elle est souhaitable, c'est au fond l'analogue de la pixellisation pour les images. Le pas retenu pour cette discrétisation dépend naturellement du pouvoir de résolution de l'oreille et à cet égard, le demi-ton chromatique est le standard retenu depuis le début du 18ème siècle.

Ces observations ont des conséquences importantes pour la théorie de la musique. La gamme acoustique s'étend continûment sur 10.4 octaves, de 16 à 20000 Hz. Ces limites sont théoriques. En pratique, selon les individus, on entend encore un bourdonnement jusqu'à 8 Hz, par contre il est rare qu'on atteigne la limite supérieure et de toutes façons ces extrêmes ne présentent aucun intérêt musical. L'un des plus grands pianos de concert, le "Grand Concert" de Stuart & Sons comporte 102 touches allant de 16.35 Hz (do0) à 5587.65 Hz (fa8). Si l'on déroule ce spectre (spectrepiano), on réalise qu'il est inutile de dépasser ces limites.

Cette plage est tellement étendue qu'une discrétisation s'impose : il est en effet impossible de gérer un ensemble trop étendu de notes, typiquement une bonne centaine. Une discrétisation s'impose qui revient à pixelliser l'espace sonore. Cela peut être fait d'une infinité de manières : aussi longtemps qu'un interprète est seul il est libre d'accorder son instrument éventuellement sa voix comme il l'entend. De fait les traditions monodiques qui sont nées de par le monde ont construit des modèles de gammes différentes. Les choses se compliquent sérieusement dès que l'on envisage de superposer plusieurs voix car l'accord des instruments est nécessaire pour éviter toute cacophonie.

Deux solutions sont envisageables soit répartir les notes en progression arithmétique soit le faire en progression géométrique. La première solution est à rejeter car le pas de la progression arithmétique serait tellement grand que l'on manquerait un grand nombre de notes utiles.

donc et a deux conséquences physiologiques :

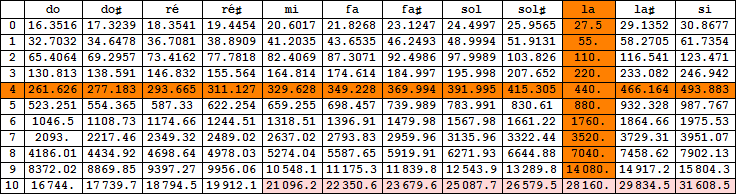
L'oreille obéit doublement à la loi empirique de Weber-Fechner : elle perçoit comme n fois plus intenses des sons d'intensités, In, en progression géométriques, In=I0 10n, d'une part, et n fois plus aigus des sons de fréquences, fn, en progression géométriques, fn=f0 10n, d'autre part. Dans ce deuxième cas, le seul qui retiendra notre attention, l'oreille ne fait aucun lien entre des fréquences disposées en simple progression arithmétique mais bien en progression géométrique. Une façon plus savante d'exprimer les choses revient à dire que l'oreille apporte une réponse logarithmique à la hauteur du stimulus, .

l'oreille est sensible au phénomène d'octave. Un son émis avec la fréquence de 440 Hz se nomme un "la". Sa période, mesurée en secondes, vaut l'inverse de sa fréquence soit, 1/440 s. C'est, par définition d'une période, le temps au bout duquel la vibration sonore reprend sa valeur de départ. Un son émis avec la fréquence double, 880 Hz, reprend sa valeur de départ au bout de 1/880 s, donc il la reprend également après 1/440 Hz. Pour le tympan ces deux sons doivent être apparentés et de fait ils le sont : si l'on chante une mélodie au départ d'une note quelconque et qu'on la rechante ensuite en doublant la fréquence de chaque note, on entend une mélodie semblable à la précédente, seulement décalée vers l'aigu. On dit qu'on l'a transposé à l'octave supérieure.

Même en s'en tenant à une seule octave, par exemple celle couvrant les fréquences allant de 440 à 880 Hz, il est évidemment inimaginable de donner des noms différents à l'infinité des sons possibles. C'est d'ailleurs inutile car l'oreille est généralement incapable de distinguer deux sons qui sont dans un intervalle de fréquences trop proche de 1 (Préciser un seuil est d'ailleurs impossible car il varie selon les individus). Ecoutons ce que donnerait un rapport-seuil fixé conventionnellement à la valeur 1.01 : voici deux sons à 400 Hz et 404 Hz en séquence pendant 4 secondes (soit 2 s par son) (ensequence) puis les deux mêmes sons joués simultanément pendant 2 s (battement). Imaginons que l'on fixe ce seuil à un rapport 1.01, cela aurait pour conséquence que pour couvrir l'octave, il faudrait pas moins de 70 notes distinctes. Multipliant ce nombre par le nombre d'octaves (10.4) cela donnerait un clavier de piano à plus de 750 touches, ce qui est impensable !

Il faut trouver autre chose et la solution retenue revient à pixelliser le spectre des fréquences sonores, c'est-à-dire le partitionner en intervalles nettement plus grands que 1.01. Plusieurs partitions ont été envisagées au cours de l'histoire de la musique et chacune avait ses avantages et ses inconvénients. Une solution s'est cependant imposée au début de 18ème siècle, sous l'impulsion de J-S Bach, consistant à partager l'octave en 12 parties égales (ce qui revient à adopter un rapport de 21/12=1.05946 entre deux notes consécutives).

Si l'on pose conventionnellement un la4 à 440 Hz, on obtient que les fréquences des notes de la gamme chromatique audible valent, f = 16.3516 2k/12, où l'entier k parcourt les valeurs 0 à 131 (au-delà votre oreille sera sans doute sourde) :



*Remarque : dans ce tableau les octaves sont numérotées selon la convention internationale. Cette convention est suivie aux USA mais pas en Europe qui pose que le "la" situé à 440 Hz appartient à la 3ème octave. Les musiciens ignorent généralement ce point de détail qui ne les concerne pas sauf s'ils composent leur musique sur base de programmes informatiques. Dans ce cas, il est commode de respecter la convention internationale, chaque note pouvant être codée par un entier composé de deux chiffres, xy, écrits en base 12 (x, y = 0, 1, 2, …, 8, 9, A, B), où x est le numéro de l'octave et y le numéro de la note au sein de cette octave (do = 0, do# =1, …, la = 9, la# = A, si = B). Ce système ne fonctionne que si le chiffre des douzaines, x, est positif ou nul, ce qui n'est vrai qu'avec la convention internationale.*

Dans ce tableau, les notes consécutives ne sont pas équidistantes en fréquences : l'échelle des fréquences n'est pas linéaire mais logarithmique. Autrement dit, deux notes consécutives sont dans un rapport de fréquences constamment égal à Cela est parfaitement inscrit dans la formule qui donne ces fréquences : Log(f) = 1.213  + 0.025 𝑘. La millième partie de ce logarithme se nomme "Savart" en sorte que l'intervalle entre deux notes consécutives vaut 25 Savarts et que l'intervalle d'octave vaut 12 fois plus soit 300 Savarts. L'oreille qui entend un son est incapable d'en déterminer la fréquence exacte; en particulier elle est incapable de décider si cette fréquence fait partie de la gamme chromatique. L'oreille absolue est, par définition, celle qui est en mesure de reconnaître toute note entendue mais elle est l'exception plutôt que la règle, même chez les musiciens professionnels. L'oreille exercée est cependant relative, capable d'identifier une note si on lui donne préalablement une note de référence pas trop éloignée, par exemple un la3.

Les raisons pour lesquelles l'octave a été partagé en 12 parties égales (et pas 14 ou 16) repose sur quelques coïncidences arithmétiques étonnantes :

* Depuis Pythagore, on sait que les sons se marient bien mélodiquement (en séquence) ou harmoniquement (en parallèle) si leurs fréquences sont dans un rapport simple. Partitionner l'octave en parties égales est peut-être commode mais encore faut-il s'assurer que si l'on part d'une note quelconque celles qui sont consonantes font également partie de cette gamme or cela n'est nullement évident puisque les notes forment un ensemble discret. Les rapports les plus simples se nomment : Quinte (3/2), Quarte (4/3), Tierce (5/4) et Sixte (5/3) dont les valeurs en Savarts sont respectivement 176, 125, 97 et 222 soit proches d'un multiple de 25. Autrement dit si l'on part d'une note quelconque de la gamme chromatique, celles situées respectivement 7 (do-sol), 5 (do-fa), 4 (do-mi) et 9 (do-la) positions plus loin dans la table sont très proches de la quinte (quarte, tierce, sixte) juste. Très proche ne signifie pas exacte mais les musiciens s'en accommodent car le bénéfice est immense d'une gamme uniformément répartie de 25 en 25 Savarts : la transposition d'un air devient un jeu d'enfant.

**Vue et audition.**

Que les ordres de grandeurs des fréquences sensibles soient extrêmement différents n'a guère d'importance en soi car les ondes en question sont de natures complètement différentes et on ne compare pas des vaches à des chevaux. Par exemple, les ondes sonores exigent un support matériel pour se propager (usuellement l'air ambiant) au contraire des ondes électromagnétiques qui se propagent dans le vide. Ce qui importe par contre c'est le rapport existant entre les extrêmes des plages sensibles, 20000/16 = 125 dans le cas des ondes sonores et à peine 7.5/4=1.875 dans le cas lumineux. La note la plus grave perçue par l'oreille intacte est le do-1 qui fait vibrer le tympan 16.35 fois par seconde. Autrement dit le tympan ainsi sollicité reprend sa position initiale tous les 6 centièmes de secondes (exactement 0.06116 = 1/16.35). Le point important est alors le suivant : si une onde de fréquence double sollicite l tympan, celui-ci reprendra sa position initiale deux fois plus souvent mais en tous cas il la reprendra également tous les 6 centièmes de secondes et cela aura pour conséquence que l'oreille percevra cette nouvelle note comme un do appartenant à l'octave suivante (do0). Le phénomène sera particulièrement sensible si l'on enchaîne mélodiquement plusieurs notes de deux manières apparemment distinctes où les fréquences des notes sont simplement doublées : les deux mélodies sonneront à l'identique la deuxième simplement décalée (le terme musical correct est transposée) vers l'aigu.

Notre vue et notre ouïe ont en commun d'être adaptés à l'analyse et au traitement des signaux périodiques. Ces signaux sont cependant de natures différentes : l'œil n'est sensible qu'ondes électromagnétiques de très hautes fréquences, typiquement de 4300 à 7500 THz tandis que l'oreille (jeune !) est sensible aux vibrations (généralement de l'air) entre 16 et 20000Hz.

L'ouïe connaît le phénomène d'octave : on signifie par là qu'un son résultant d'une vibration à la fréquence 440 Hz (qu'on appelle un "la" dans un diapason ordinaire), un son vibrant à la fréquence moitié.

L'oreille connaît de ce fait le phénomène d'octave alors que l'œil ne le connaît pas.

*La vue se comporte très différemment précisément parce que le phénomène d'octave n'existe pas pour l'œil humain. Une lumière qui vibre à 4.0 1014 Hz est perçue comme rouge et on pourrait se demander comment serait perçue une lumière vibrant à la fréquence double 8.0 1014 Hz (une sorte de rouge aigu !?) mais précisément cette question est vide de sens puisque notre œil n'y est pas sensible. Le peintre n'a en conséquence aucun moyen de profiter d'une gamme fréquentielle aussi étendue que le musicien et il faudra qu'il diversifie autrement sa palette chromatique.*

**Peinture et musique.**

Tout sépare l'art du peintre de celui du musicien et pourtant l'ingrédient de base est le même une onde physique périodique. La lumière est de nature électromagnétique et à ce titre elle se propage dans le vide comme dans la matière transparente. Le son lui a impérativement besoin d'un support matériel pour se propager, usuellement l'air qui achemine la vibration vers nos tympans. Le paramètre principal de l'onde purement sinusoïdale est sa fréquence qui détermine la couleur en optique et la hauteur en acoustique. A partir de là tout sépare l'optique et l'acoustique physiologiques.

L'œil n'est sensible qu'à une plage tenue des fréquences lumineuses, typiquement entre 4300 et 7500 THz, soit un ambitus de 1.875, inférieur à 2, ce point aura son importance.

En comparaison l'oreille est sensible à une plage fréquentielle beaucoup plus grande, typiquement entre 16 et 20000 Hz, d'ambitus 1250, largement supérieur à 2.

Le peintre gère une palette continue de couleur : il a pour ce faire à sa disposition tous les dégradés possibles un coup de brosse lui suffit. L'œuvre qu'il produit est immédiatement accessible au spectateur.

Les choses diffèrent essentiellement en musique : l'œuvre doit être notée puis interprétée d'où il est impossible de traiter continûment le spectre sonore, celui-ci doit être discrétisé. La question de savoir comment opérer au mieux cette pixellisation a occupé les mathématiciens du 18ème siècle (Euler, d'Alembert) mais aussi quelques musiciens férus de théorie (Rameau, Bach)

Un essai au carrefour des arts (Peinture et musique) et des sciences (Physique, mathématiques et Théorie de l'information).

L'ensemble des fréquences étant infini (et même non dénombrable) il convient de se restreindre ce qui est rendu possible par le fait que ni l'œil ni l'oreille ne possèdent un pouvoir de résolution infini. Il en résulte qu'une pixellisation de l'image et du son est possible qui discrétise l'objet visuel ou sonore. La pixellisation de l'image est connue de tous et elle se comprend d'autant plus facilement qu'elle se satisfait d'un cadre souvent rectangulaire également partitionné dans les deux directions spatiales x et y.

La pixellisation de la gammes sonore est éventuellement plus délicate. Une solution confortable mais imparfaite a tôt été mise en place à l'époque de Bach qui consiste à partitionner la plage des fréquences audibles en n'en retenant qu'un échantillonnage disposé en progression géométrique de raison 21/12. La raison pour laquelle la progression est géométrique et non arithmétique ou n'importe quoi d'autre est liée à la physiologie de l'oreille qui perçoit les sons de façon logarithmique, c'est le phénomène d'octave.

L'art musical présente cette particularité d'exiger le concours de trois partenaires égaux : le compositeur, l'interprète et l'auditeur.

*Le peintre agit très différemment lui qui ne dispose que d'une octave même pas complète allant du rouge au violet en passant par toutes les couleurs de l'arc-en-ciel. Il explore en conséquence l'infinité des nuances intermédiaires dans le cadre d'un continuum. Rien en théorie n'empêcherait le musicien gourmand d'en faire autant en explorant les quarts voire les seizièmes de tons. Ce genre d'expérience (courante dans le monde arabe) a effectivement ponctuellement été tenté en musique savante occidentale (Aloïs Haba) mais elle ne s'est nulle part imposée comme indispensable.*

Il existe une autre différence entre ces deux arts, qui tient au rôle joué par le facteur temps. Le temps ne joue aucun rôle en peinture : la toile achevée l'est une fois pour toutes aux réfections ultérieures près. Il en va tout autrement en musique.

**Compréhension et redondance.**

Pour être intelligible, un discours doit être redondant - donc compressible sans perte - et la musique n'échappe pas à cette règle. En musique la redondance peut prendre toutes sortes de formes, allant du trivial da capo où une section entière est reprise à l'identique aux répétitions micro-variées qui abondent dans les œuvres de Beethoven ou dans la musique minimaliste. L'adoption d'un mode musical répond également à cette exigence d'intelligibilité. Avec le temps l'usage a prévalu de se contenter de deux modes particuliers, majeur et mineur