

LES MATHÉMATIQUES DE L'ENSEIGNEMENT SCIENTIFIQUE LA NUMÉRISATION ET LE CODAGE D'UN SON

Mots-clés

Échantillonnage ; quantification ; codage ; compression.

Références au programme

Le son, vibration de l'air, peut être enregistré sur un support informatique. Les techniques numériques ont mis en évidence un nouveau type de relations entre les sciences et les sons, le processus de numérisation dérivant lui-même de théories mathématiques et informatiques.

Savoir

Pour numériser un son, on procède à la discrétisation du signal analogique sonore (échantillonnage et quantification).

Plus la fréquence d'échantillonnage est élevée et la quantification est fine, plus la numérisation est fidèle, mais plus la taille du fichier audio est grande.

La reproduction fidèle du signal analogique nécessite une fréquence d'échantillonnage au moins double de celle du son.

La compression consiste à diminuer la taille d'un fichier afin de faciliter son stockage et sa transmission.

Les techniques de compression spécifiques au son, dites « avec perte d'information », éliminent les informations sonores auxquelles l'oreille est peu sensible. Il existe aussi des techniques de compression sans perte d'information.

Savoir-faire

- Justifier le choix des paramètres de numérisation d'un son.
- Estimer la taille d'un fichier audio.
- Calculer un taux de compression.
- Comparer des caractéristiques et des qualités de fichiers audio compressés.

Notions mathématiques mobilisées

- Discrétisation d'une fonction.
- Nombres en écriture binaire.
- Unités de temps, unités de quantité de données, débit.
- Taux (de compression).
- Lecture de courbes.

Histoire, enjeux, débats

Claude Shannon (1916-2001), ingénieur en génie électrique et mathématicien, est considéré comme l'un des pères de la théorie de l'information dont l'une des problématiques consiste à représenter, compresser, stocker, communiquer des informations.

Vues du côté de l'utilisateur, ces informations peuvent prendre la forme de textes, d'images, de sons, de vidéos. On peut donc s'interroger sur la manière dont un ordinateur, dont la mémoire est constituée d'un très grand nombre de circuits électroniques, peut traiter des informations aussi sophistiquées. La numérisation des sons permet d'appréhender cette problématique.

Dans un sens plus général, une théorie de l'information est une théorie visant à quantifier et qualifier la notion de « contenu en information » présent dans un ensemble de données. À ce titre, il existe une autre théorie de l'information : la théorie algorithmique de l'information, créée par Kolmogorov, Solomonov et Chaitin au début des années 1960.

Au plan sociétal, la compression de données et la transmission en flux (streaming) ont fait évoluer les pratiques culturelles d'écoute de la musique.

Les mathématiques et la numérisation et le codage d'un son

La mémoire d'un ordinateur est constituée d'un très grand nombre de circuits électroniques dont chacun peut se retrouver dans l'un ou l'autre de deux états, qu'on désigne conventionnellement 0 et 1. Un bit d'information (Binary digiT) est donc l'un ou l'autre de ces deux chiffres. L'état d'un circuit composé de plusieurs circuits mémoires à un bit se décrit par une suite finie de 0 et de 1.

Par exemple, un octet est un mot formé de huit bits. On voit alors pourquoi il est nécessaire de coder les nombres, non pas en base dix comme on le fait habituellement (à l'origine parce qu'on a dix doigts), mais en base deux. Selon le principe de la numération de position en base 10, l'écriture 235 traduit que :

$$235 = 2 \times 10^2 + 3 \times 10^1 + 5 \times 10^0$$

Selon le même principe :

$$235 = 128 + 64 + 32 + 8 + 2 + 1$$

$$235 = 1 \times 2^7 + 1 \times 2^6 + 1 \times 2^5 + 0 \times 2^4 + 1 \times 2^3 + 0 \times 2^2 + 1 \times 2^1$$

L'écriture en base deux d'un nombre est appelée écriture binaire. L'écriture binaire de 235 est :

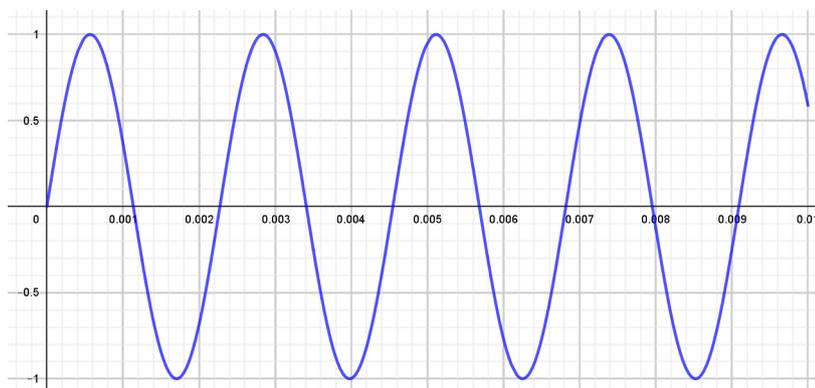
$$11\ 101\ 011.$$

Codage au son

Un son est une vibration mécanique se propageant dans l'air ou dans un autre milieu (fluide, solide...). À l'aide d'un micro, un son peut être capté et converti en un signal analogique modélisé mathématiquement par une fonction représentant par exemple une tension en fonction du temps.

Retrouvez éduscol sur :



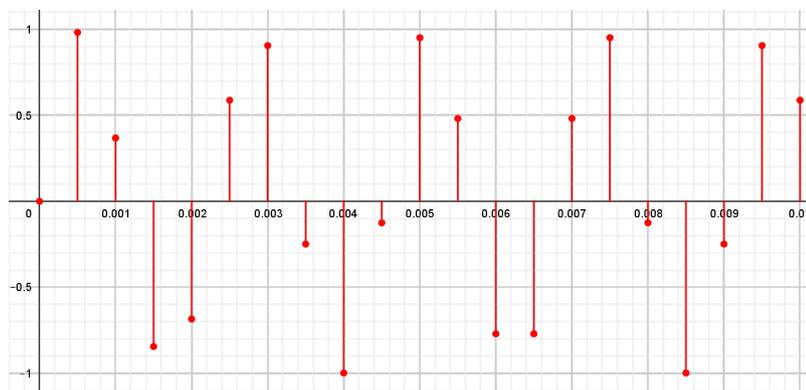


Le traitement numérique de ce signal analogique, réalisé à l'aide d'un convertisseur analogique/numérique, consiste à discrétiser cette fonction en abscisse et en ordonnée pour en extraire un nombre fini de données. L'échantillonnage consiste à relever différentes valeurs de la tension à intervalles de temps réguliers. La quantification, quant à elle, revient à associer à chaque valeur de l'échantillon un nombre dont la longueur de l'écriture binaire est décidée par avance.

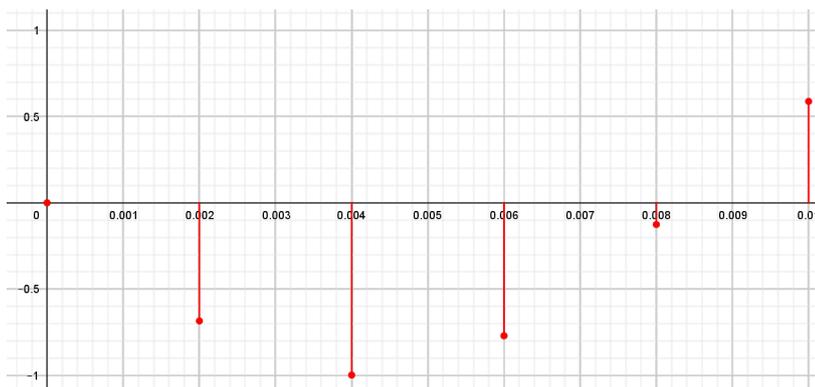
Échantillonnage

Considérons la représentation graphique en fonction du temps de la tension électrique correspondant au signal analogique du La du diapason (La3), son pur de fréquence 440 Hz.

On effectue un échantillonnage en prélevant des mesures de cette tension toutes les 5×10^{-4} s (période d'échantillonnage), comme l'illustre le graphique ci-dessous. La fréquence d'échantillonnage vaut $\frac{1}{5 \times 10^{-4}} = 2000$ Hz.



Un échantillonnage à la fréquence 500 Hz fournit quant à lui l'échantillon suivant :



Retrouvez éducol sur :



On conjecture que le premier échantillon permettra une reconstitution du son analogique initial plus fidèle que le second.

En introduction de l'énoncé du théorème de Nyquist-Shannon, on peut, sur des exemples, faire varier la fréquence du son analogique et celle de de l'échantillonnage. Ce théorème, dont la démonstration est inaccessible à ce niveau mais dont le résultat est à connaître, stipule que la reproduction fidèle d'un signal analogique à partir d'un échantillon nécessite d'avoir échantillonné avec une fréquence au moins double de la fréquence initiale.

Ce théorème, démontré par Shannon en 1949, utilise des outils mathématiques en lien avec la théorie de Fourier (1768 – 1830). Il est à la base de la numérisation de l'information.

Écriture binaire d'un nombre entier

0 et 1 sont les seuls nombres que l'on peut coder sur un seul bit. Sur deux bits, il y a quatre nombres codés : 00, 01, 10 et 11.

Plus généralement, si p est un entier naturel non nul, on peut coder 2^p nombres entiers sur p bits.

Nombre p de bits	1	2	3	4	8	10	16	24
Nombre de nombres codables sur p bits	2	4	8	16	256	1 024	65 536	16 777 216

On voit ainsi qu'un octet, constitué de 8 bits, permet de représenter 256 nombres distincts.

L'octet est l'unité qui, en informatique, permet de mesurer la quantité de données.

Des multiples de l'octet, utilisant les préfixes décimaux du Système International, sont régulièrement employés.

Nom	Kilooctet	Mégaoctet	Gigaoctet	Téraoctet	Pétaoctet
Valeur en octets	10^3	10^6	10^9	10^{12}	10^{15}
Notation	ko	Mo	Go	To	Po

Les ordres de grandeurs s'expriment suivant les puissances de dix en raison de l'écriture décimale ; en informatique, le codage en base deux induit l'utilisation de puissances de deux. On emploie alors des préfixes proches de ceux en bases dix en remarquant que $2^{10} = 1024 \approx 10^3$.

Nom	Kibiocet	Mébiocet	Gibiocet	Tébiocet	Pébiocet
Valeur en octets	2^{10}	2^{20}	2^{30}	2^{40}	2^{50}
Notation	Kio	Mio	Gio	Tio	Pio

Retrouvez éduscol sur :



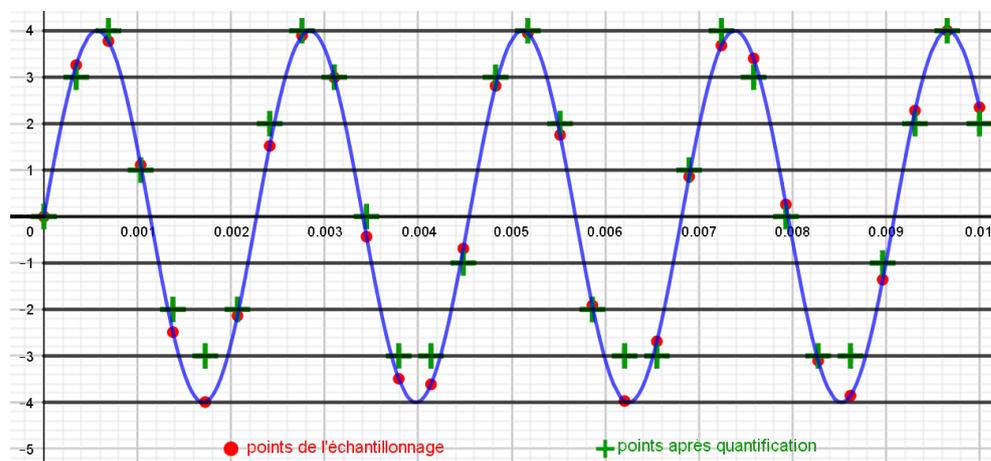
Quantification

Les valeurs des tensions, en nombre fini, obtenues après échantillonnage sont des nombres réels. Il faut les coder en mots uniquement formés de 0 et de 1 sur le nombre de bits utilisés pour le codage.

La quantification consiste alors à approcher au mieux les valeurs des tensions échantillonnées par des nombres, dont la taille de l'écriture binaire est fixée par le nombre de bits retenu.

Si on choisit une quantification sur p bits, alors on peut coder 2^p nombres différents.

Ainsi, si on choisit de coder sur trois bits, il y a seulement huit valeurs possibles. L'amplitude des tensions doit alors être découpée en huit intervalles de même longueur. Toutes les valeurs issues de l'échantillonnage situées dans un même intervalle sont quantifiées par le même nombre écrit en binaire sur trois bits.



Sur l'exemple ci-dessus, il faut représenter sur trois bits les nombres -3, -2, -1, 0, 1, 2, 3 et 4. Le tableau ci-dessous donne une façon de le faire à l'aide d'une représentation dite « complément à 2 » :

Nombres entiers	-3	-2	-1	0	1	2	3	4
Codages binaires	101	110	111	000	001	010	011	100

Dans cette représentation, un nombre x négatif est représenté par le codage du nombre positif $2^3 - |x|$. Cette précision sur la représentation des nombres pourra être réservée aux élèves suivant la spécialité NSI.

Valeurs usuelles

- Les fréquences audibles par l'oreille humaine sont comprises entre 20 Hz et 20 000 Hz. Le théorème d'échantillonnage de Nyquist-Shannon amène à choisir une fréquence d'échantillonnage supérieure au double de la fréquence maximum audible, soit 40 000 Hz.
- La voix humaine oscille entre des fréquences allant de 100 Hz à 3 400 Hz. La fréquence d'échantillonnage pour numériser la voix doit donc être supérieure à 6 800 Hz. Pour transmettre une voix sur une ligne téléphonique, celle-ci est échantillonnée avec une fréquence de 8 000 Hz et est codée sur 8 bits. Diffuser un son sur une ligne téléphonique nécessite donc de transmettre $8 \times 8000 = 64\,000$ bits par seconde.

Retrouvez éducol sur :



- Sur les disques audionumériques (CD) la musique est enregistrée en stéréo. Elle est codée sur 16 bits avec une fréquence d'échantillonnage de 44,1 kHz. Ainsi, l'enregistrement d'une seconde de musique nécessite un espace de stockage de $2 \times 16 \times 44\,100 = 1\,411\,200$ bits (2 car c'est un enregistrement stéréo, 16 est le nombre de valeurs possibles pour la quantification, enfin l'échantillonnage à la fréquence de 44 100 Hz fournit 44 100 échantillons par seconde).

Notons que $1\,411\,200 \text{ bits} = \frac{1\,411\,200}{8} \text{ octets} = 176,4 \times 10^3 \text{ octets} = 176,4 \text{ ko}$. Une seconde de musique nécessite un espace de stockage de 176,4 ko.

- Le débit binaire mesure la quantité de données numériques transmises par unité de temps. L'unité utilisée est le bit par seconde (bit/s) et ses multiples kbit/s, Mbit/s ou Gbit/s. Quelques débits binaires peuvent faire référence :
 - un code Morse : 40 bit/s ;
 - un câble avec prise USB.2 : 480 Mbit/s ;
 - une connexion Bluetooth : 3 Mbit/s ;
 - une connexion ADSL à Internet : 13 Mbit/s.

Compression

Pour réduire les difficultés liées au stockage et à la transmission de fichiers audio, on effectue des compressions des données. Il existe des techniques de compression sans perte et d'autres avec perte.

Une compression est dite sans perte d'information si elle permet de récupérer, après décompression, l'intégralité des sons produits. Elle est réalisée par des algorithmes exploitant les redondances et la prévision de ces redondances dans les fichiers audio. Ainsi, le format FLAC permet de réduire de 30 % à 70 % la taille d'un fichier audio sans perte d'information.

Dans le cas contraire, la compression est dite avec perte. La compression avec perte supprime les sons peu audibles. La compression est effectuée par des algorithmes. Un format très connu de compression de ce type est le MP3.

Les services de musique en ligne proposent en streaming ou en téléchargement des fichiers MP3 à 128 kbit/s. Cela signifie que, pour un tel fichier, une seconde de musique nécessite 128 kbit de données.

Comme pour un CD audio, une seconde de musique nécessite 1411 kbit de données. On en déduit que le taux de compression d'un CD audio vers un fichier MP3 à 128 kbit/s est égal à $\frac{128}{1411} \approx 0,091$.

Le taux de compression est d'environ 9 % ou encore dans le ratio de 1:11 puisque $0,09 \approx \frac{1}{11}$.

Intentions pédagogiques

Le sous-thème de la numérisation des sons fournit l'occasion de mettre en action des concepts et des pratiques issues de différentes sciences (physique, mathématiques, informatique). L'animation GeoGebra®, qui peut être présentée en classe par le professeur ou fournie aux élèves pour faire des manipulations et émettre des conjectures, permet de visualiser les différentes étapes de l'échantillonnage et de la quantification.

Le travail sur les nombres en écriture binaire, en lien avec le programme de SNT, permet à la fois d'appréhender l'un des fondements de l'informatique, de manipuler des puissances de 2 et de réinterroger les algorithmes opératoires en base dix.

Retrouvez éducol sur :

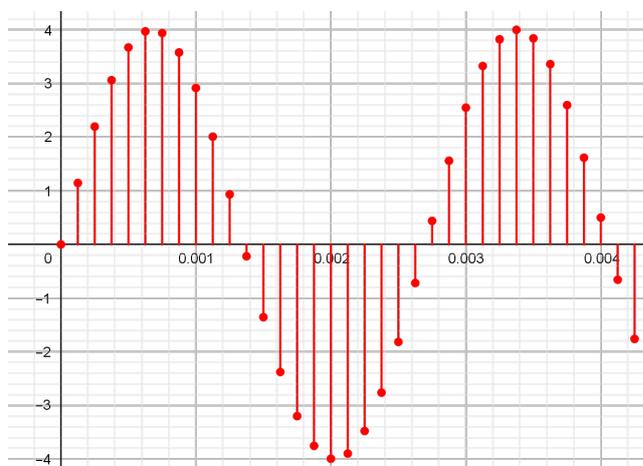


La manipulation de différentes grandeurs et de leurs unités permet de consolider le sens des opérations et les ordres de grandeurs utilisés en informatique, dans un contexte lié aux usages des fichiers numériques.

Propositions d'Activités

Activité 1 : lecture de courbes illustrant l'échantillonnage et la quantification

On considère un ensemble de valeurs obtenues par échantillonnage d'un son pur. L'unité portée sur l'axe des abscisses est la seconde.



1.

1.1. Déterminer la fréquence du son échantillonné.

1.2. Déterminer la fréquence f d'échantillonnage.

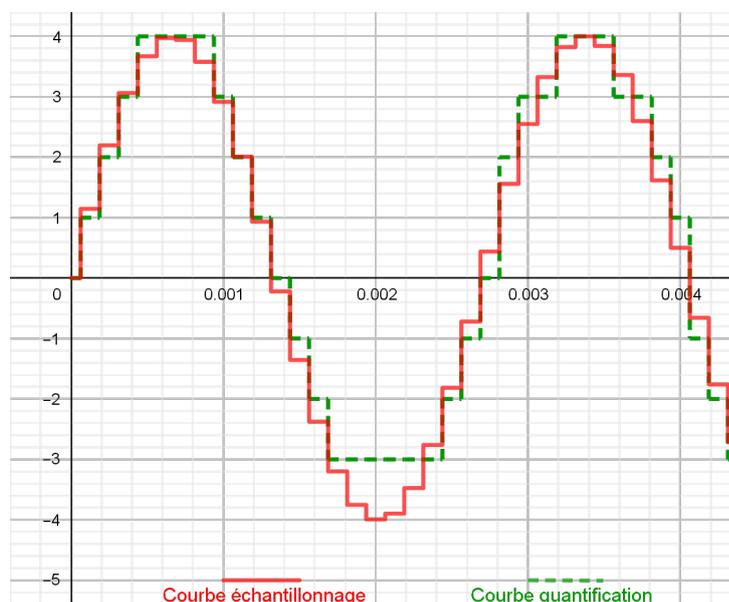
À partir de l'échantillon réalisé ci-dessus, on trace la courbe d'échantillonnage de la manière suivante : on associe à chaque point de l'échantillon un segment parallèle à l'axe des abscisses, centré sur le point, de longueur $\frac{1}{f}$ (période de l'échantillonnage). Cela revient à considérer que la tension du signal est constante sur une durée, exprimée en seconde, égale à $\frac{1}{f}$ autour de chaque prise de tension.

On construit également la courbe obtenue après quantification, sur le même principe, mais avec les valeurs quantifiées.

Retrouvez éduscol sur :



On obtient les graphiques suivants.



2.

2.1. À partir de la courbe d'échantillonnage, déterminer la fréquence d'échantillonnage et indiquer sur combien de bits sont codées les valeurs quantifiées.

2.2. Les nombres positifs sont simplement écrits dans leur écriture binaire : ainsi le nombre 3 a-t-il pour écriture binaire 011.

Pour coder un nombre négatif :

- On considère l'écriture binaire de sa valeur absolue. Par exemple pour -3 on commence par écrire 011.
- On détermine le complément à 1 de ce dernier nombre, ce qui revient à prendre la négation de chaque bit : 0 est remplacé par 1 et 1 est remplacé par 0. Par exemple, 011 devient 100.
- On ajoute 1 au nombre obtenu (ne pas oublier les retenues éventuelles), ce qui constitue le code cherché. Par exemple, en ajoutant 1 à 100, on obtient 101. Le codage de -3 est 101.

Appliquer l'algorithme ci-dessus pour déterminer le codage de -1 et -2 .



Télécharger l'animation GeoGebra® intitulée « [Échantillonnage et quantification d'un son](#) ». Cette animation permet de construire des courbes d'échantillonnage et de quantification d'un son pur dont on peut faire varier la fréquence (encore appelée hauteur).

Retrouvez éducol sur :



Activité 2 : taille de fichiers audio et débit binaire

Les services de musique en ligne proposent en téléchargement de la musique en diverses résolutions. L'un d'eux propose des fichiers « haute résolution » correspondant à un échantillonnage à 96 kHz, un codage sur 24 bits et un enregistrement stéréo.

On fait l'hypothèse totalement irréaliste que les fichiers proposés ne sont pas compressés.

1. Quel est l'espace de stockage nécessaire pour enregistrer une seconde de musique de cette qualité ?
2. Quelle est la taille d'un fichier pouvant contenir un enregistrement de cette qualité des six suites pour violoncelle de Bach, d'une durée totale de 2 h 15 min ?
3. Pour une connexion Internet dont le débit est 8 Mbit/s, peut-on écouter en streaming un enregistrement de cette qualité ?
4. Avec cette même connexion internet, combien de temps faut-il pour télécharger les six suites de Bach ?

Activité 3 : fichiers compressés

Un fichier audio, stéréo, échantillonné à 44,1 kHz et codé sur 16 bits, contenant les six suites pour violoncelle de Bach d'une durée totale de 2 h 15 min est compressé en un fichier MP3 à 320 kbit/s.

1. Combien de données sont utilisées pour coder 1 seconde de musique sur le fichier avant compression ?
2. Calculer le taux de compression pour passer d'un fichier à l'autre.
3. Quelle est la taille du fichier MP3 obtenu ?
4. Combien de temps faut-il pour télécharger ce fichier avec une connexion Internet de 8 Mbit/s ?
5. Un mélomane possède dans son audiOTHÈQUE de nombreux CD, tous enregistrés en stéréo, échantillonnés à 44,1 kHz et codés sur 16 bits. Il estime qu'en moyenne, chaque CD dure une heure et décide de stocker sa musique sous forme de fichiers MP3 à 320 kbit/s.
Combien peut-il stocker de fichiers MP3 sur un disque dur ayant une capacité de 1 To ?

Pour aller plus loin : calcul en binaire

Addition et soustraction en écriture binaire. Les opérations entre nombres écrits en base deux permettent de prendre conscience de certains processus opératoires (notamment ceux qui sont liés à la retenue) devenus automatiques en base dix.

Bibliographie et sitographie

- Site internet lors du [centenaire Claude E. Shannon](#).
- Article « [Représenter les sons musicaux](#) » sur le site de la revue pour la science.
- Site interstices [Joseph Fourier](#) et les sons.

Retrouvez éducol sur :

