



Traitement numérique du signal

Livrable final

Groupe G10C

MANGIN Rémi  
LE MONNIER Romain  
TEOFILOVIC Kristina  
TON Sophie  
RISY Théo  
YANG Chen

15 décembre 2023

# Table des matières

<b>1</b>	<b>Chapitre I : Problème I</b>	<b>3</b>
1.1	Introduction . . . . .	3
1.2	Méthode . . . . .	4
1.3	Résultats Expérimentaux . . . . .	6
1.4	Conclusion . . . . .	8
<b>2</b>	<b>Chapitre II : Problème II</b>	<b>9</b>
2.1	Introduction . . . . .	9
2.2	Méthode . . . . .	10
2.2.1	Acquisition et Pré-traitement du Signal . . . . .	10
2.2.2	Détection des Notes . . . . .	11
2.2.3	Calcul de la puissance moyenne . . . . .	11
2.2.4	Détermination de la Fréquence Fondamentale . . . . .	12
2.2.5	Analyse des Harmoniques . . . . .	12
2.2.6	Extraction et classification des caractéristiques des sons d'instruments de musique . .	13
2.3	Résultats Expérimentaux . . . . .	14
2.4	Conclusion . . . . .	14
<b>3</b>	<b>Chapitre III : Problème III</b>	<b>15</b>
3.1	Introduction . . . . .	15
3.2	Méthode . . . . .	15
3.2.1	Calcul du gain $G$ à appliquer au signal analogique . . . . .	15
3.2.2	Schéma fonctionnel de numérisation . . . . .	16
3.2.3	Calcul du nombre de bits . . . . .	16
3.2.4	Calcul de la capacité de stockage pour enregistrer une heure d'audio en stéréo . . . .	18
3.2.5	Compatibilité avec une qualité et un stockage sur CD . . . . .	18
3.2.6	Respect de la fidélité du timbre des instruments d'un orchestre symphonique . . . .	18
3.2.7	Réflexion sur l'utilité d'un pas de quantification qui suit une loi logarithmique . . . .	19
3.3	Conclusion . . . . .	19
<b>4</b>	<b>Chapitre IV : Problème IV</b>	<b>20</b>
4.1	Introduction . . . . .	20
4.2	Méthode . . . . .	20
4.3	Résultats Expérimentaux . . . . .	20
4.4	Conclusion . . . . .	20

# 1 Chapitre I : Problème I

## 1.1 Introduction

Nous présentons une approche de détection de signaux audio, dont les caractéristiques sont les suivantes :

- Un micro avec une sensibilité (S) de - 48 dBV et un Gain (G) de 40 dB.
- Les sons pénibles sont définis comme des sons de plus d'une seconde, dépassant une puissance en dBm de référence.
- Les sons acceptables sont définis en dessous de ce seuil.
- Chaque signal sera caractérisé par sa durée, sa puissance moyenne, tension RMS, et classifié en conséquence.

Les expérimentations seront réalisées sur les signaux de la question F, signaux qui nous ont été fournis (bruit de marteau piqueurs, bruit de ville et de jardin). Dans un premier temps, nous implémenterons le code de détection sur le micro-contrôleur, en choisissant une fréquence d'échantillonnage adaptée aux caractéristiques matérielles. La caractérisation des bruits détectés fera l'objet d'une étape ultérieure.

### Problème I - Énoncé

Proposer un algorithme qui permet de détecter la présence/absence d'un signal audio enregistrant le bruit ambiant. Le micro utilisé a une sensibilité S (dBV) et amplifie le signal électrique avec un gain G. On définit un son "son pénible", comme un signal sonore de durée supérieure à  $D_t$  (s) et de niveau supérieur ou égal à  $P_{SPL}$  (dB SPL). A contrario, un "son acceptable", correspond à un signal sonore de niveau inférieur à  $P_{SPL}$ .

Caractériser chaque son détecté par :

- Sa durée
- sa puissance
- sa tension RMS en V et le classifiez en "son pénible" ou "son acceptable"

Paramètres S = -48 dBV, G = 40 dB,  $P_{SPL} = 80$  dB SPL,  $D_t = 1$ s.

Les expérimentations seront menées sur les signaux de la question F.

## 1.2 Méthode

Tout d'abord, il faut définir  $P_{dBm}$ . Le schéma du microphone qui nous est donné permet de réaliser les calculs.

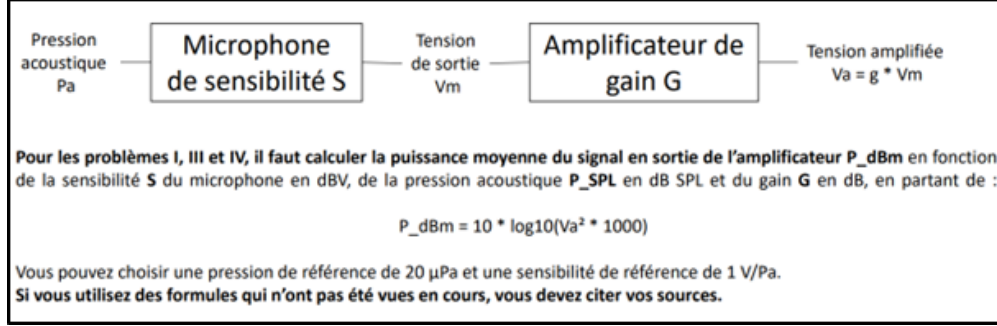


FIGURE 1 – Schéma bloc microphone

En effet, nous avons la relation suivante :

$$P_{dBm} = 10 \times \log(V_a^2 \times 1000) \text{ avec } V_a = S \times P_a \quad (1.1)$$

Cependant,  $S$  la sensibilité<sup>1</sup> que l'on nous donne est en dBV et pas en  $V/P_a$ , il faut donc la convertir. Nous savons que :

$$S = 20 \times \log\left(\frac{S_{v/P_a}}{S_{ref}}\right) \text{ avec } S_{ref} = 1V/P_a \quad (1.2)$$

donc :

$$S_{V/P_a} = S_{ref} * 10^{\frac{S}{20}} \quad (1.3)$$

Nous avons donc :

$$V_m = S_{V/P_a} \times P_a \quad (1.4)$$

Pour  $P_a$ , nous utilisons la formule de  $P_{SPL}$ <sup>2</sup> :

$$P_{SPL} = 20 \times \log\left(\frac{P_a}{P_{ref}}\right) \quad (1.5)$$

d'où :

$$P_a = P_{ref} * 10^{\frac{P_{SPL}}{20}} \quad (1.6)$$

Il ne reste que le gain à traiter, celui-ci s'exprime en dB dans les paramètres donnés, or, nous voulons l'exprimer sans unité, nous obtenons alors

$$G_{SU} = 10^{\frac{G}{20}} \text{ où } G_{SU} \text{ est sans unité}^3 \quad (1.7)$$

Cela nous permet de trouver la formule finale :

$$P_{dBm} = 10 \times \log((G_{SU} \times S_{VPA} \times P_a)^2 \times 1000)$$

En substituant avec les résultats précédents, on obtient :

- 
1. Pour la formule de la sensibilité  $S$
  2. Pour la formule de  $P_{SPL}$
  3. Pour la formule du gain sans unité  $G_{SU}$

$$P_{dBm} = 20 \times \log(10^{\frac{G}{20}} \times S_{REF} \times 10^{\frac{S}{20}} \times P_{REF} \times 10^{\frac{P_{SPL}}{20}}) + 30$$

Enfin, avec :

- $S_{REF} = 1$
- $P_{REF} = 20\mu Pa$
- $G = 40dB$
- $S = -48dBV$
- $P_{SPL} = 80dB SPL$

On obtient :

$$P_{dBm} = 8dBm.$$

À présent, nous avons la valeur de  $P_{dBm}$  qui va nous servir de référence. Pour chaque signal, nous allons le décomposer en plusieurs « sous-signaux » à chaque fois que la valeur de référence est franchie. Par exemple, sur la figure 2, on distingue trois « sous-signaux » ou plages de signal différentes, car la valeur de référence est franchie deux fois. Ensuite, chaque plage est traitée comme un nouveau signal et, si cette plage se trouve au-dessus de la valeur référence et dure plus d'une seconde, alors le son de cette plage est catalogué dans la catégorie bruit pénible. Sur cet exemple, il y a donc deux sons acceptables (plages 1,3) et un son pénible (plage 2). Enfin, pour chaque signal, on règle la taille des fenêtres d'analyse pour obtenir un résultat correct, contenant  $2k+1$  échantillons. Pour déterminer la puissance instantanée, nous utilisons la formule suivante :

$$P(n) = \frac{1}{2K+1} \sum_{n-K}^{n+K} x(k)^2 \quad (1.8)$$

L'identification du paramètre  $k$  ainsi que le choix de sa valeur retenue dépendent de la nature du signal et des exigences spécifiques de l'analyse. Admettons que nous avons un signal de 5s et d'une fréquence de 20k Hz, si  $k = 80$ , on aura 161 échantillons par fenêtre. Pour un signal sonore, il faut que la taille de la fenêtre soit petite, car c'est un fichier son. Pour avoir un bon vecteur puissance à analyser, il faut donc une fenêtre inférieure à la seconde, comme on a  $T_e$ , il faut déterminer  $k$  tel que :

$$k < \frac{(\frac{1}{T_e} - 1)}{2} \quad (1.9)$$

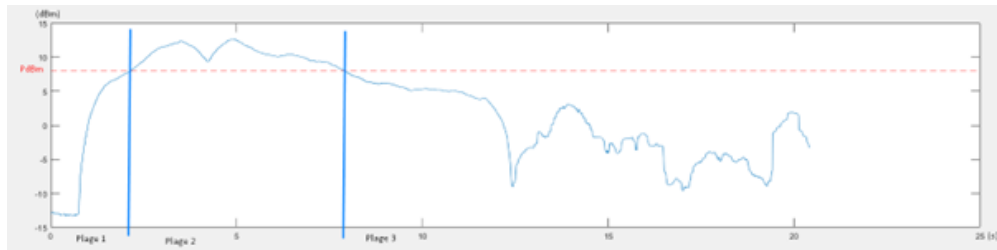


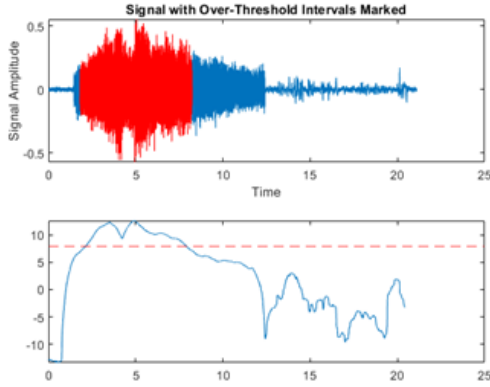
FIGURE 2 – Exemple signal avec plages acceptables et pénibles

### 1.3 Résultats Expérimentaux

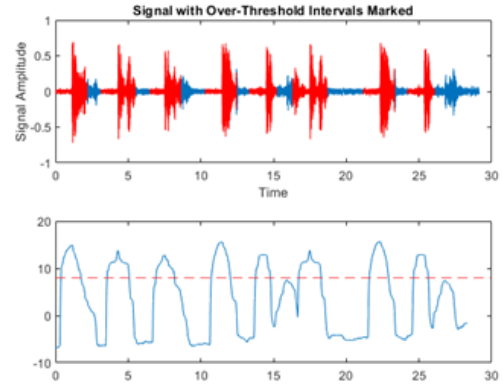
Après avoir appliqué la fonction du problème aux signaux Jardin01.mp3, Jardin02.mp3, Ville01.mp3 et MarteauPiqueur01.mp3, voici ce que nous obtenons :

**Légende :**

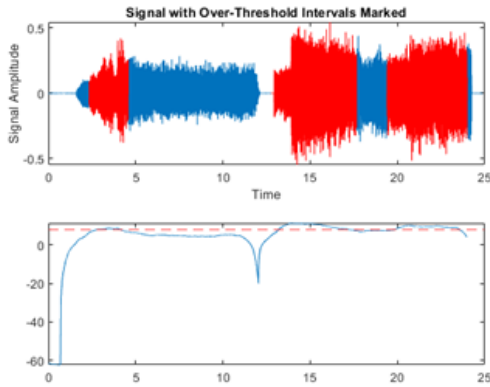
**Rouge :** dépasse la durée et la puissance seuil, **Bleu :** dépasse la puissance seuil mais plus petit que dt



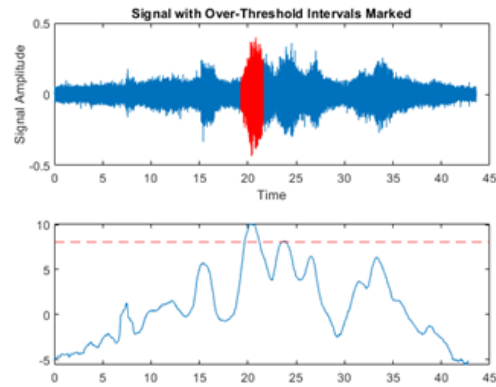
(a) Application sur le signal Jardin01



(b) Application sur le signal Jardin02



(c) Application sur le signal MarteauPiqueur



(d) Application sur le signal Ville01

FIGURE 3 – Application sur les signaux donnés

Pour chacun des signaux, l'amplitude en fonction du temps est présentée en haut, tandis que la puissance en dBm en fonction du temps est affichée en bas. Les pointillés rouges désignent la puissance de référence à 8 dBm et on observe bien que lorsque le signal dépasse 8 dBm et dure plus d'une seconde, on a une plage de son pénible (en rouge).

Avant amplification

	Ville01	MarteauPiqueur01	Jardin01	Jardin02
Puissance moyenne (dBm)				
Puissance moyenne (mW)				
Durée				
Tension RMS (V)				

Dans ce tableau des caractéristiques des différents signaux, nous avons déterminé la puissance moyenne. Elle mesure la puissance d'un signal sur une période de temps donnée et elle est définie comme la quantité totale d'énergie transférée ou consommée divisée par la durée de cette période. Entre deux instants, elle est l'énergie dissipée par la durée d'observation et sa formule est :

$$P(x, t_1, t_2) = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} x_c(t)^2 dt \quad (1.10)$$

## 1.4 Conclusion

Pour conclure, le problème 1 a été abordé à travers la conception d'un algorithme dédié à la détection de la présence ou de l'absence de signaux audio, enregistrant ainsi le bruit ambiant. Nos expérimentations ont débuté en utilisant des sons préalablement définis, fournis dans la question F, tels que des bruits de marteau-piqueur, de ville, et de jardin 1 et 2. L'analyse approfondie de ces sons a permis de classer les signaux en deux catégories distinctes : les "sons pénibles", identifiés en rouge, définis par une durée excédant à 1 seconde et une puissance en dBm égale ou supérieure à 8, et les "sons acceptables" distingués en bleu, caractérisés par une puissance en dBm inférieure à ce seuil.

Nos résultats présentent des points forts notables. La méthode de classification des signaux a démontré une clarté visuelle, facilitée par l'utilisation de couleurs, renforçant ainsi la compréhension des différentes catégories. De plus, la caractérisation détaillée de chaque signal fournit une base solide pour évaluer l'impact des sons sur l'environnement.

Cependant, des points faibles nécessitent une attention particulière. La détermination du seuil de puissance pour les "sons pénibles" pourrait être améliorée pour refléter davantage la variabilité des environnements sonores. De plus, la limitation actuelle de nos expérimentations à des enregistrements préétablis soulève des questions sur la généralisation de notre algorithme à des situations en temps réel et dynamique.

Pour perfectionner notre recherche, nous envisageons d'optimiser la calibration du seuil de puissance en tenant compte de la variabilité environnementale. L'extension de nos expérimentations à des conditions réelles d'environnement sonore permettra une validation plus approfondie de notre algorithme. Ces perspectives d'amélioration visent à renforcer la fiabilité et l'applicabilité de notre méthode dans des situations plus diversifiées et changeantes.



## 2 Chapitre II : Problème II

### 2.1 Introduction

Nous proposons une approche visant à élaborer un algorithme de détection et de caractérisation des notes de musique jouées par divers instruments tels que le violon, la flûte et le piano, tout en déterminant précisément leur hauteur. Notre algorithme est conçu pour générer des sorties exhaustives, incluant les instants de début et de fin de chaque note ( $t_d, t_f$ ), la puissance moyenne du signal en dBm ( $P_{dBm}$ ), la fréquence fondamentale ( $f_0$ ) accompagnée du nom et de l'octave de la note jouée, la fréquence haute  $f_h$  définie par  $[0, f_h]$  englobant 99.99% de la puissance, le nombre d'harmoniques dans cette bande de fréquences ( $n_h$ ), ainsi que toutes autres caractéristiques jugées pertinentes pour classer les notes selon l'instrument de musique. Nous adopterons la méthode de détection d'un son utile du PBI (Produit de l'analyse du Bruit Instantané) pour élaborer notre algorithme, en nous appuyant sur une approche en trois étapes. Cette méthodologie implique la détection du signal sonore, la synchronisation du récepteur par la détection de l'en-tête, et enfin, le décodage des informations pertinentes pour caractériser chaque note.

#### Problème II - Énoncé

Proposer un algorithme qui permet de détecter, de caractériser une note de musique jouée par un instrument (violon, flûte, piano) et de déterminer sa hauteur.

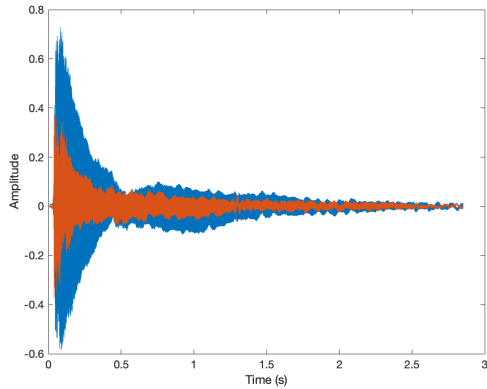
Cet algorithme devra produire les sorties suivantes :

- $t_d, t_f$  : instants de début et de fin de note
- $P_{dBm}$  : la puissance moyenne du signal en dBm
- $f_0$  : la fréquence fondamentale et le nom et l'octave de la note jouée
- $f_h$  : la fréquence haute telle que  $[0, f_h]$  contient 99.99% de la puissance
- $n_h$  : le nombre d'harmoniques dans cette bande de fréquences
- Toutes autres caractéristiques qui vous paraît pertinente pour classer les notes par instrument de musique.

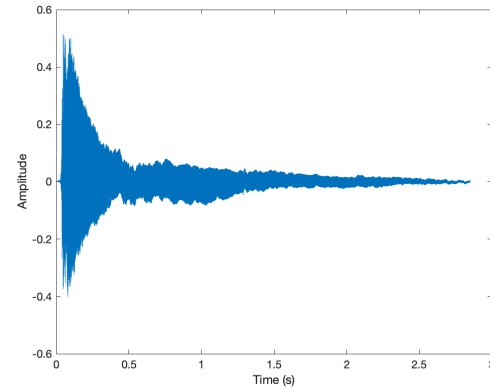
On reprendra la méthode de détection d'un son utile du PBI et on justifiera les paramètres retenus.

## 2.2 Méthode

Les signaux originaux sont en stéréo, il faut les moyenner pour retrouver le cadre du cours et ainsi avoir des signaux en mono. fig.4(a) représente le signal original et fig.4(b) représente le signal moyenné.



(a) Exemple de signal audio original



(b) Exemple de signal moyenné

FIGURE 4 – Exemple de signal

### 2.2.1 Acquisition et Pré-traitement du Signal

Dans cette première étape, nous chargeons une série de fichiers audio correspondant à différentes notes jouées par des instruments tels que le violon, la flûte et le piano. Ces fichiers audio sont ensuite convertis en signaux mono pour simplifier l'analyse. Pour chaque signal, nous calculons la puissance du signal en utilisant une fenêtre glissante de taille fixe. La puissance ainsi obtenue est ensuite convertie en décibels (dB). Ce processus est essentiel pour préparer le signal en vue de l'analyse ultérieure.

### 2.2.2 Détection des Notes

La détection se fait en trouvant les points où la puissance du signal dépasse un seuil défini (1% de la puissance maximale). Les indices correspondants sont ensuite convertis en temps, fournissant ainsi les instants de début et de fin de chaque note. Les notes d'une durée inférieure à une seconde sont éliminées pour garantir la qualité des données. Fig.5 représente le signal avec les notes détectées (partie rouge).

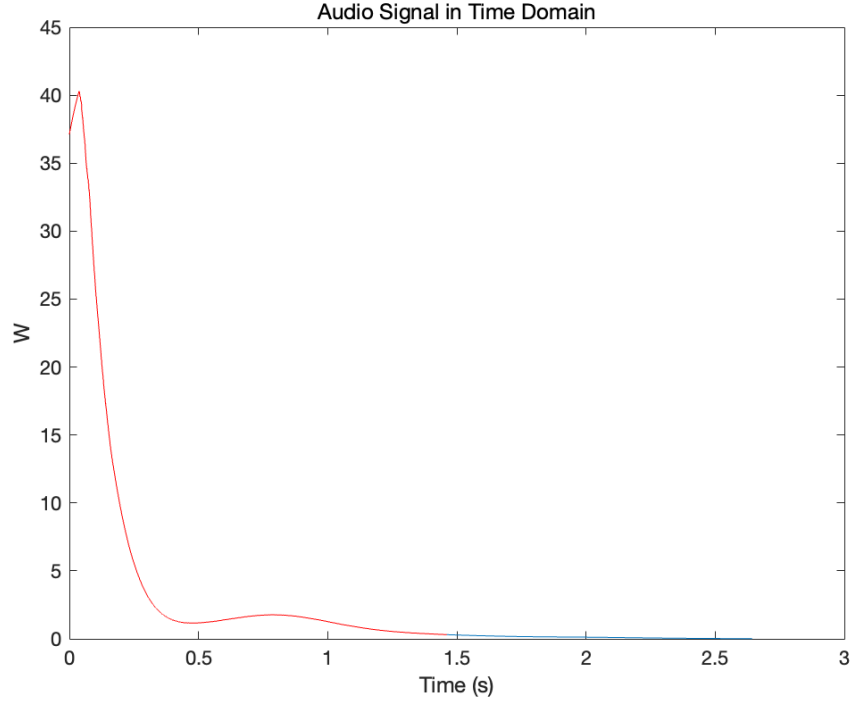


FIGURE 5 – Exemple signal avec les notes détectées

### 2.2.3 Calcul de la puissance moyenne

En cette partie, nous allons calculer la puissance moyenne de la partie des notes précédemment isolées.

$$\bar{P} = \frac{1}{t_{\text{end}}f_s - t_{\text{start}}f_s} \sum_{n=t_{\text{start}}f_s}^{t_{\text{end}}f_s} P(n)^2 \quad (2.1)$$

Où :

- $f_s$  est la fréquence d'échantillonnage
- $t_{\text{start}}$  est l'heure de début
- $t_{\text{end}}$  est l'heure de fin
- $n$  est l'indice de l'échantillon

La puissance moyenne en dBm peut alors être calculée en utilisant :

$$\text{avgPower}_{\text{dBm}} = 10 \cdot \log_{10} \left( \frac{\bar{P}}{0.001} \right) \quad (2.2)$$

### 2.2.4 Détermination de la Fréquence Fondamentale

Le décodage des caractéristiques de chaque note s'effectue en calculant la puissance moyenne, la fréquence fondamentale ( $f_0$ ), la fréquence de la plus haute harmonique ( $f_h$ ), et le nombre d'harmoniques dans la bande de fréquences. La détection précise de la fréquence fondamentale ( $f_0$ ) est essentielle dans l'analyse des signaux sonores, notamment pour l'étude des notes musicales. Ces informations sont extraites à partir du signal audio, utilisant diverses fonctions telles que le calcul de la moyenne de puissance, la fréquence fondamentale par auto-corrélation, et le comptage des harmoniques. Les résultats sont affichés pour chaque note, fournissant une caractérisation détaillée des éléments musicaux contenus chaque fichier audio ainsi que la note et l'octave associée à la fréquence.

L'ensemble de ces étapes contribue à élaborer un algorithme complet de détection et de caractérisation des notes de musique, avec une option facultative pour estimer la direction de la source sonore sur des signaux acquis par deux micros différents.

Dans notre étude, nous avons appliqué deux méthodes distinctes pour calculer cette fréquence fondamentale, puis nous les avons comparées afin de vérifier leur exactitude.

La première méthode s'appuie sur la détection de la fréquence qui présente la densité de puissance la plus élevée après avoir effectué une transformation de Fourier rapide (FFT). Cette approche est basée sur l'hypothèse que, dans la majorité des cas, la fréquence fondamentale correspond à la fréquence la plus basse ayant la plus grande densité de puissance.

$$f_0 = \arg \max_f \left( \frac{2}{N} |FFT(signal)| \right) \quad (2.3)$$

Quant à la seconde méthode, nous utilisons le principe de l'auto-corrélation dans le domaine temporel. Cette technique permet de détecter les répétitions périodiques au sein du signal, ce qui facilite l'identification de la fréquence fondamentale, même en présence de harmoniques complexes ou de bruit.

$$f_0 = \frac{f_s}{Lag_{peak}} \quad (2.4)$$

Où :

- $f_0$  est la fréquence fondamentale
- $f_s$  est la fréquence d'échantillonnage
- $Lag_{peak}$  est le temps de retard correspondant au premier pic significatif de la fonction d'auto-corrélation.

Ces méthodes contribuent à élaborer un algorithme complet de détection et de caractérisation des notes de musique. En combinant et en comparant les résultats obtenus par ces deux méthodes, nous pouvons augmenter la fiabilité de notre analyse et assurer une détection plus précise de la fréquence fondamentale, ce qui est crucial pour une compréhension approfondie des propriétés acoustiques des notes analysées.

### 2.2.5 Analyse des Harmoniques

Dans l'analyse des signaux audio, la détection précise des harmoniques est essentielle pour comprendre les caractéristiques spectrales du signal.

Notre recherche utilise une méthode basée sur la transformée de Fourier rapide (FFT) pour identifier les composantes harmoniques d'un signal. Ce processus consiste à calculer la FFT du signal et à en extraire les points de fréquence qui répondent à des conditions spécifiques et qui représentent les composantes harmoniques.

Tout d'abord, nous effectuons une FFT sur le signal original pour obtenir son spectre, et nous ne considérons que la partie positive de la fréquence. Dans le spectre, nous identifions d'abord le point de fréquence ayant la puissance la plus élevée ( $P_{max}$ ), et à partir de là, un seuil relatif est fixé (typiquement 40dB en dessous du point de puissance maximale). Tous les points de fréquence supérieurs à ce seuil sont initialement reconnus comme des harmoniques possibles.

Ensuite, pour chaque multiple entier de la fréquence fondamentale ( $n \times f_0$ , où  $n$  est un nombre naturel), nous recherchons les points du spectre qui sont les plus proches de ces fréquences. Si les fréquences de ces points ne dépassent pas un seuil de fréquence maximale fixé ( $f_{max}$ ) avec une puissance d'au moins 10% de la

puissance maximale, ces points de fréquence sont identifiés comme des harmoniques. De cette manière, nous pouvons extraire efficacement les composante harmoniques du signal.

$$Harmonics = \{f_n | f_n = n \times f_0, n \in \mathbb{N}, f_n \leq f_{max}, P(f_n) \geq 0.1 \times P_{max}\} \quad (2.5)$$

où :

- $f_n$  est la fréquence de la nième harmonique
- $f_0$  est la fréquence fondamentale
- $P(f_n)$  est la densité spectrale de puissance à la fréquence  $f_n$
- $P_{max}$  est la maxime dans le spectrale de fréquence
- $f_{max}$  est le seuil de fréquence le plus élevé pour la détection des harmoniques

### 2.2.6 Extraction et classification des caractéristiques des sons d'instruments de musique

Dans le domaine de l'analyse audio numérique, l'extraction des caractéristiques sonores des instruments de musique est essentielle pour la recherche d'informations musicales et la classification des instruments. Nous avons étudié trois caractéristiques clés : le centroïde spectral, le taux de croisements par zéro et l'enveloppe ADSR, ainsi que leur application dans l'identification des types d'instruments.

**Le centroïde spectral** décrit le centre de gravité du spectre sonore, reflétant la « brillance » du son. Sa représentation mathématique est la suivante :

$$C = \frac{\sum_{n=0}^{N-1} f(n) \cdot S(n)}{\sum_{n=0}^{N-1} S(n)} \quad (2.6)$$

où :

- $f(n)$  représente la fréquence
- $S(n)$  l'amplitude correspondante à cette fréquence

**Le taux de croisements par zéro** mesure la fréquence à laquelle la forme d'onde du signal traverse l'axe temporel, et est liée à la hauteur et au timbre du son. Sa formule de calcul est la suivante :

$$ZCR = \frac{1}{T-1} \sum_{t=1}^{T-1} \frac{1}{2} |\text{sgn}(s(t)) - \text{sgn}(s(t-1))| \quad (2.7)$$

**L'enveloppe ADSR** analyse les variations dynamiques temporelles du son, où chaque phase est définie comme suit :

- **Attack Time(A) :**

$$A = \min\{t | s(t) \geq 0.9 \cdot P_{peak}, t \in [0, \frac{T}{f_s}]\} \quad (2.8)$$

- **Decay Time(D) :**

$$D = \min\{t | s(t) \leq S_{level}, t \in [A, \frac{T}{f_s}] - A\} \quad (2.9)$$

- **Sustain Level (S) :**

$$S = \alpha \cdot P_{peak} \quad (2.10)$$

Où :

$\alpha$  est un pourcentage de l'amplitude maximale

$P_{peak}$  est l'amplitude maximale du signal

- **Release Time (R) :**

$$R = \min\{t | s(t) \leq \beta \cdot P_{initial}, t \in [t_{end}, \frac{T}{f_s}]\} - t_{end} \quad (2.11)$$

Où :

$\beta$  est le coefficient qui détermine le seuil d'extinction sonore

$P_{initial}$  est l'amplitude au début de la phase de libération

$t_{end}$  est le point où la note se termine

$T$  est la taille totale de l'échantillon

$f_s$  est le taux d'échantillonnage

À travers ces caractéristiques, nous avons tenté de classer les instruments de musique. Cependant, nous avons constaté que les résultats de classification n'étaient pas idéaux. Les principales raisons en sont le choix arbitraire des seuils de caractéristiques, la possible superposition des caractéristiques entre différents instruments, ainsi que la complexité due à la diversité des échantillons.

Compte tenu de ces défis, nous estimons qu'il serait plus approprié d'utiliser des méthodes d'apprentissage automatique pour la classification des instruments. L'apprentissage machine, en particulier les techniques d'apprentissage profond, peut automatiquement apprendre des caractéristiques complexes pour distinguer les instruments à partir de grandes quantités de données. Ces méthodes réduisent non seulement le besoin de fixer manuellement des seuils, mais elles permettent également, en apprenant les modèles intrinsèques dans les données, d'améliorer la précision et la capacité de généralisation de la classification.

## 2.3 Résultats Expérimentaux

Nom de Fichier	Note	$f_0$ (Hz)	$f_h$ (Hz)	$n_h$ <sup>4</sup>	$t_d$ (s)	$t_f$ (s)	$P_{dBm}$
Pi_A_96K.wav	A3	440.367	876.648	1	1.0417e-05	1.3567	9.0004
Pi_C_96K.wav	C3	261.580	-	0	1.0417e-05	2.9677	5.9068
Vi_A3_96K.wav	A3	440.367	3963.157	7	0.0781	1.8450	0.7896
Vi_C3_96K.wav	C3	262.295	2360.814	4	0.0286	1.5718	4.3884
Vi_G4_96K.wav	G4	780.488	3902.377	2	0.0593	1.5783	6.9208
Fl_A4_96K.wav	A4	880.734	-	0	0.0231	1.6218	19.8006
Fl_B3_96K.wav	B3	494.845	1979.1543	2	0.0079	1.8108	10.2884

TABLE 1 – Résumé des caractéristiques acoustiques extraites pour chaque note.

## 2.4 Conclusion

En conclusion, nous avons réussi par le code via Matlab à répondre aux attentes du cahier des charges, c'est à dire à caractériser des notes de musiques. Cette caractérisation passe par la détection de la note jouées, de sa fréquence fondamentale et de son harmonique et de sa puissance moyenne en dBm. Le programme détecte aussi les instants de début et de fin ainsi que le nombre d'harmoniques.

Les résultats à l'exécution du programme sont cohérent avec ce que l'on attend (avec le nom des fichiers qui donnent le nom des notes) et aussi de ce l'on entend.

---

4. Seules les quantités harmoniques sont comptées, la fréquence fondamentale n'est pas incluse

## 3 Chapitre III : Problème III

### 3.1 Introduction

Nous envisageons la réalisation d'un enregistrement numérique en exploitant un microphone doté d'une sensibilité de  $S = -47$  dBV et capable de capturer des niveaux sonores atteignant 130 dB SPL dans une plage de fréquence de 200 Hz à 20 kHz. En supposant l'absence totale de distorsion et de bruit introduit par le microphone, le signal analogique subit une amplification dont le gain, exprimé en décibels, doit être déterminé. La plage dynamique cible après amplification est fixée à  $[-1V, +1V]$ . Après cette étape, le signal est numérisé avec des paramètres spécifiques visant à garantir un rapport signal à bruit d'au moins 40 dB pour un signal acoustique de PdB SPL = 60 dB SPL. Nous entreprenons de déterminer le gain  $G$  nécessaire à l'amplification du signal analogique, de fournir un schéma fonctionnel détaillé pour la numérisation, de spécifier tous les paramètres requis, et de calculer le débit du signal numérique résultant. Enfin, nous examinerons la capacité de stockage nécessaire pour enregistrer une heure d'audio en stéréo, évaluerons la compatibilité des paramètres avec la qualité et le stockage sur CD, et analyserons leur aptitude à préserver le timbre des instruments d'un orchestre symphonique.

#### Problème III - Énoncé

On souhaite effectuer un enregistrement numérique en utilisant un micro de sensibilité  $S = -47$  dBV, de niveau sonore maximal de 130 dB SPL de bande passante 200 Hz - 20 kHz. On suppose que ce micro n'introduit aucune distorsion sur la bande considérée et aucun bruit. Le signal analogique est ensuite amplifié avec un gain  $G$  (en dB). La dynamique après amplification est  $[-1V, +1V]$ . Ce signal est ensuite numérisé. On choisit des paramètres de numérisation afin de garantir un rapport signal à bruit d'au moins 40 dB pour un signal acoustique de PdB SPL = 60 dB SPL.

Déterminer le gain  $G$  à appliquer au signal analogique et donner le schéma fonctionnel de numérisation de ce signal, déterminer tous les paramètres et calculer le débit du signal numérique.

Quelle capacité de stockage faut-il pour enregistrer une heure d'audio en stéréo ? Les paramètres choisis sont-ils compatibles avec une qualité et un stockage sur CD ? Permettent-ils de respecter le timbre des instruments d'un orchestre symphonique ?

Réfléchir sur l'utilité d'adopter un pas de quantification non pas uniforme, mais qui suit une loi logarithmique.

### 3.2 Méthode

#### 3.2.1 Calcul du gain $G$ à appliquer au signal analogique

Reprenons la formule du problème I :

$$P_{dBm} = 20 \log(S_{REF} * 10^{\frac{S}{20}} * P_{REF} * 10^{\frac{P_{SPL}}{20}} * 10^{\frac{G}{20}}) \quad (3.1)$$

pour la simplifier :

$$P_{dBm} = 20 \log(S_{REF}) + S + 20 \log(P_{REF}) + P_{SPL} + G + 30 \quad (3.2)$$

avec

- $S_{REF} = 1 \frac{V}{P_a}$
- $P_{REF} = 20 \mu P_a$
- $S = -47$  dBV

donc nous avons :

$$P_{dbm} = 10 \log(V_a^2 * 1000) = 30 \quad (3.3)$$

(comme la dynamique après amplification est  $[-1V, +1V]$ , donc  $V_a = 1V$ ) ; nous obtenons :

$$G = P_{dBm} - 20 \log(1) - 20 \log(20 * 10^{-6}) - S - P_{SPL} - 30 \quad (3.4)$$

soit,

$$G = P_{dBm} - S - P_{SPL} - (30 + 20 \log(20 * 10^{-6}))$$

Avec l'application numérique, nous trouvons notre gain.

$$G = 30 + 47 - 130 - (30 + 20 \log(20 * 10^{-6})) \approx 11$$

### 3.2.2 Schéma fonctionnel de numérisation



FIGURE 6 – Schéma fonctionnel

### 3.2.3 Calcul du nombre de bits

Pour ce faire, nous utiliserons la formule du rapport signal bruit SNR qui est définie comme :

$$SNR = 10 \log \left( \frac{P_s}{P_b} \right). \quad (3.5)$$



Avec un SNR d'au moins 40 dB, nous déterminons d'abord la puissance du signal utile  $P_s$  puis la puissance du bruit  $P_b$ , ce qui nous permet de calculer le nombre de bits.

Reprenons la formule du gain pour déterminer la puissance en dBm :

$$G = P_{\text{dBm}} - S - P_{\text{SPL}} - (30 + 20 \log(20 \times 10^{-6})) . \quad (3.6)$$

Ce qui nous donne alors :

$$P_{\text{dBm}} = G + S + P_{\text{SPL}} + 30 + 20 \log(P_{\text{ref}}). \quad (3.7)$$

Avec  $P_{\text{SPL}} = 60$  dB SPL et les mêmes valeurs que pour la partie 3.2.1, nous avons :

$$P_{\text{dBm}} = 11 - 47 + 60 + 30 + 20 \log(20 \times 10^{-6}) \quad \text{donc} \quad P_{\text{dBm}} = -40 \text{ dBm}. \quad (3.8)$$

Pour convertir la puissance dBm en Watt, appliquons la formule :

$$P_W = 10 \log \left( \frac{P_{\text{dBm}}}{10^{-3}} \right) \quad \text{qui peut également être exprimée comme } P_{\text{dBm}} = 10 \log(1000 P_w) . \quad (3.9)$$

En isolant  $P_w$ , nous obtenons :

$$P_W = 10^{\left( \frac{P_{\text{dBm}}}{10^{-3}} \right)} \quad \text{ce qui nous donne} \quad P_W = 10^{-7} \text{ Watt}. \quad (3.10)$$

De plus, nous avons :

$$SNR = 10 \log \left( \frac{P_s}{P_b} \right), \quad (3.11)$$

donc

$$P_b = \frac{P_s}{10^{\frac{SNR}{10}}}. \quad (3.12)$$

La puissance moyenne du bruit de quantification  $P_b$  est donc :

$$P_b = \frac{q^2}{12}, \quad (3.13)$$

soit,

$$q^2 = \frac{12 \times P_s}{10^{\frac{SNR}{10}}}. \quad (3.14)$$

De plus,  $q$  est donné par :

$$q = \frac{2A}{2^b}, \quad (3.15)$$

donc

$$\left( \frac{2A}{2^b} \right)^2 = \frac{12 \times P_s}{10^{\frac{SNR}{10}}}, \quad (3.16)$$

et

$$A^2 \times 2^{2(1-b)} = \frac{12 \times P_s}{10^{\frac{SNR}{10}}}. \quad (3.17)$$

Ainsi, nous avons :

$$2^{2(1-b)} = \frac{12 \times P_s}{A^2 \times 10^{\frac{SNR}{10}}}, \quad (3.18)$$

et en terminant les calculs :

$$2(1-b) = \log_2 \left( \frac{12 \times P_s}{A^2 \times 10^{\frac{SNR}{10}}} \right), \quad (3.19)$$

$$b = 1 - \frac{\log_2\left(\frac{12 \times P_s}{A^2 \times 10^{-10} \text{ SNR}}\right)}{2}$$

puis

Avec  $A = 1V$ ,  $\text{SNR} = 40 \text{ dB}$ , et  $P_s = 10^{-7}W$ , nous obtenons  $b \approx 17.47$ .

18 bits.

Nous savons que les bits prennent des valeurs entières, ainsi, nous avons donc

### 3.2.4 Calcul de la capacité de stockage pour enregistrer une heure d'audio en stéréo

Nous avons  $D = F_e * b$

Où  $D$  est le débit binaire,  $F_e$  la fréquence d'échantillonnage,  $b$  le nombre de bits.

En considérant nos paramètres, avec une bande passante 200 Hz à 20 kHz, la fréquence d'échantillonnage  $F_e$  doit être d'au moins 40 000 Hz, conformément au théorème d'échantillonnage de Shannon, où nous devons respecter  $F_e \geq 2F_{max}$  avec  $F_{max}$  étant égal à 20 kHz. Ainsi  $F_e$  doit être d'au moins 40 kHz. Pour le nombre de bits  $b$ , nous avons 18 bits, ensuite, il faut multiplier le débit par 2, car nous sommes en stéréo, enfin, il faut effectuer une division par 8 pour obtenir le résultat en octet.

$$C_s = \frac{3600 * 40000 * 18 * 2}{8} = 648 M_o$$

Ainsi, après calculs, nous avons

Cependant, nous ne sommes pas certains d'être en stéréo, alors nous effectuons le même calcul en considérant le cas où nous sommes en mono.

$$C_m = \frac{3600 * 40000 * 18}{8} = 324 M_o$$

Après les calculs, nous obtenons

### 3.2.5 Compatibilité avec une qualité et un stockage sur CD

La fréquence d'échantillonnage  $F_e$  pour un CD standard est déjà établie à 44 100 Hz, tout comme le nombre de bits  $b$  qui est 16 bits et l'enregistrement sur CD se fait en stéréo.(1)

Ainsi,

$$D = 44.1 * 10^3 * 16 * 2 = 1411200 \text{ bit/s}$$

La capacité  $C_{CD}$  est exprimée par  $C_{CD} = D * d$ , où  $D$  représente le nombre binaire et  $d$  la durée que nous avons sélectionnée, correspondant à 1 heure, soit 3 600 secondes.

En effectuant une division par 8 pour obtenir le résultat en octet, la capacité  $C_{CD}$  se calcule comme suit :

$$C_{CD} = \frac{1411200 * 3600}{8} \approx 635 M_o$$

Ainsi, la capacité de stockage nécessaire pour enregistrer une heure d'audio en stéréo sur un CD standard est d'environ 635  $M_o$ .

Par conséquent, avec nos paramètres choisis, l'enregistrement n'est pas conforme aux spécifications d'un CD audio standard, si on est dans le cas où nous sommes en stéréo. Il pourrait être envisageable d'explorer d'autres supports ou formats offrant une flexibilité plus grande pour répondre à des exigences spécifiques en termes de bande passante et de qualité audio. Sinon, dans le cas où **nous sommes en mono**, cela est conforme aux spécifications d'un CD audio standard.

source : 1. Cours de cycle préparatoire associé avec Stanislas

### 3.2.6 Respect de la fidélité du timbre des instruments d'un orchestre symphonique

Pour récupérer de la musique avec fidélité, il convient de conserver le timbre musical, autrement dit, garder le plus d'harmoniques possibles, et pas uniquement le fondamental. Autrement dit, plus la fréquence

d'échantillonnage est élevée, et plus nous avons de chance de respecter le timbre d'un instrument (ici source). Prenons un orchestre symphonique possédant une harpe, la plus haute note jouée avec cet instrument est un sol dièse 7 (3322.4 Hz). Notre bande passante est de 20 000 Hz, nous pouvons donc récupérer 6 harmoniques d'un sol dièse 7 en comptant la fondamentale (nous atteignons alors 19 934.4 Hz). De plus, comme précisé dans la partie 3.2.4, notre fréquence d'échantillonnage doit être au moins deux fois plus grande pour éviter le repliement spectral, ainsi, nous pouvons estimer que l'on respectera le timbre des instruments de musique dans un orchestre symphonique. L'oreille humaine ne pouvant pas réellement distinguer les sons au-delà de 14 000 Hz, cela est suffisant. De surcroît, nous avons calculé un nombre de bits de quantification égale à 18. Ce nombre de bits est suffisamment élevé pour réduire au plus possible le bruit de quantification. En vérité, si le CD reste peu aimé pour l'écoute de musique classique (ici source), c'est plus par appréciation de l'unité des performances d'un orchestre, ou par le fait de ne pas aimer le support lui-même, mais la qualité d'enregistrement ne semble pas un problème. Le CD standard a 16 bits de quantification, on peut conclure que le timbre des instruments d'un orchestre sera respecté.

### 3.2.7 Réflexion sur l'utilité d'un pas de quantification qui suit une loi logarithmique

En règle générale, on adopte un pas de quantification uniforme pour échantillonner de l'information qui suit elle-même une distribution uniforme. Cependant, les signaux sonores ne suivent pas de distribution uniforme. Cependant, les signaux sonores ne suivent pas de distribution uniforme, il peut être utile d'adopter un pas de quantification suivant une loi logarithmique pour minimiser les distorsions (bruit de quantification). Il existe d'ailleurs un algorithme permettant de déterminer le pas de quantification idéal, l'algorithme de Lloyd-Max.

source : [Ce lien](#)

## 3.3 Conclusion

En conclusion, notre démarche vise à réaliser un enregistrement numérique à partir d'un microphone. Après avoir déterminé un gain optimal de 11 dB et un nombre de bits de 18, nous avons calculé une capacité de stockage estimée de 648  $M_o$  pour 1 heure d'audio en stéréo.

Cependant, l'estimation de la capacité de stockage, à 648  $M_o$  pour une heure d'audio en stéréo, présente une légère divergence par rapport à la norme d'un CD standard, fixée à 635  $M_o$ . Cette différence souligne la complexité de l'équilibre entre la qualité audio et la contrainte de stockage, ce qui pourrait rendre une telle configuration incompatible avec un CD standard.

Il est important de noter toutefois que si l'enregistrement est effectué en mono, notre approche pourrait être compatible avec les spécifications d'un CD standard. Cela met en lumière l'importance de considérer la configuration audio choisie lors de l'évaluation de la compatibilité avec les normes de stockage existantes. De plus, la réflexion sur l'adoption d'un pas de quantification logarithmique a ouvert des perspectives innovantes pour optimiser la représentation des signaux, en particulier à faible amplitude, tout en minimisant les distorsions.

Notre approche méthodologique, du choix de gain à la réflexion sur la quantification, établit une base solide pour des ajustements futurs visant à concilier au mieux qualité sonore et contraintes de stockages dans des contextes spécifiques.

## 4 Chapitre IV : Problème IV

### 4.1 Introduction

#### **Problème IV - Énoncé**

On souhaite détecter les signaux environnementaux qui comprennent des composantes spectrales très aiguës et pénibles pour l'oreille humaine. Ces signaux sont définis ainsi : 20% de leur puissance est comprise dans les fréquences supérieures à 2 kHz, avec une puissance sonore totale supérieure à 110 dB SLP.

Concevoir un système de détection de ces signaux pénibles fondé sur filtrage numérique. On prendra un micro de sensibilité égale à -67 dBV et de gain égal à 16 dB. Discuter les critères énoncés.

### 4.2 Méthode

### 4.3 Résultats Expérimentaux

### 4.4 Conclusion