



Laboratoire Traitement du Signal et de l'Image

# Codage audio

Adrien Llave

14 novembre 2023

① Traitement du signal

② Compression - intro

③ Compression sans pertes

④ Compression avec pertes

⑤ Compression neuronale

⑥ Résumé

# Objectifs



- Démystifier le traitement du signal (et la théorie de l'information)
- Donner quelques intuitions sur la compression de données
- Donner un peu de culture technique

## D'où je parle



- 2010 - 2012 : IUT GEII
- 2012 - 2015 : Master métiers du son à l'ENS Louis-Lumière
- 2015 - 2016 : Master traitement du signal à Grenoble-INP Phelma
- 2017 - 2022 : Doctorat traitement du signal à CentraleSupélec
- 2022 - 2023 : Chercheur à Orange
- 2023 : Enseignement à l'IUT, recherche au LTSI
- 2024 - : Chercheur à Orange

Traitement du signal  
●ooooooooo

Compression - intro  
ooo

Compression sans pertes  
oooooooooooo

Compression avec pertes  
oooooooooooo

Compression neuronale  
oooooooooooo

Résumé  
oo

# Traitement du signal

## « Traitement du signal » ???

Pensons à une science...



## « Traitement du signal » ???

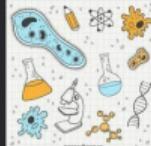
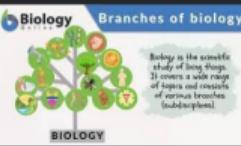
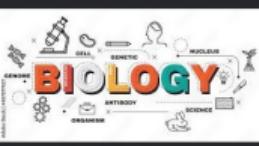
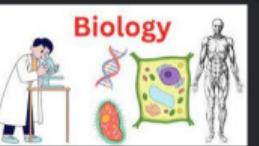
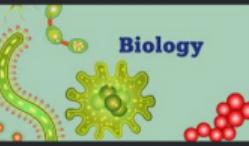
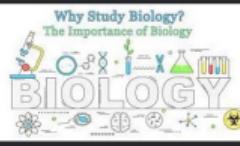
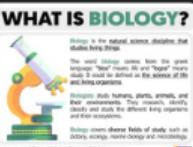
Pensons à une science... la **biologie** :



# « Traitement du signal » ???



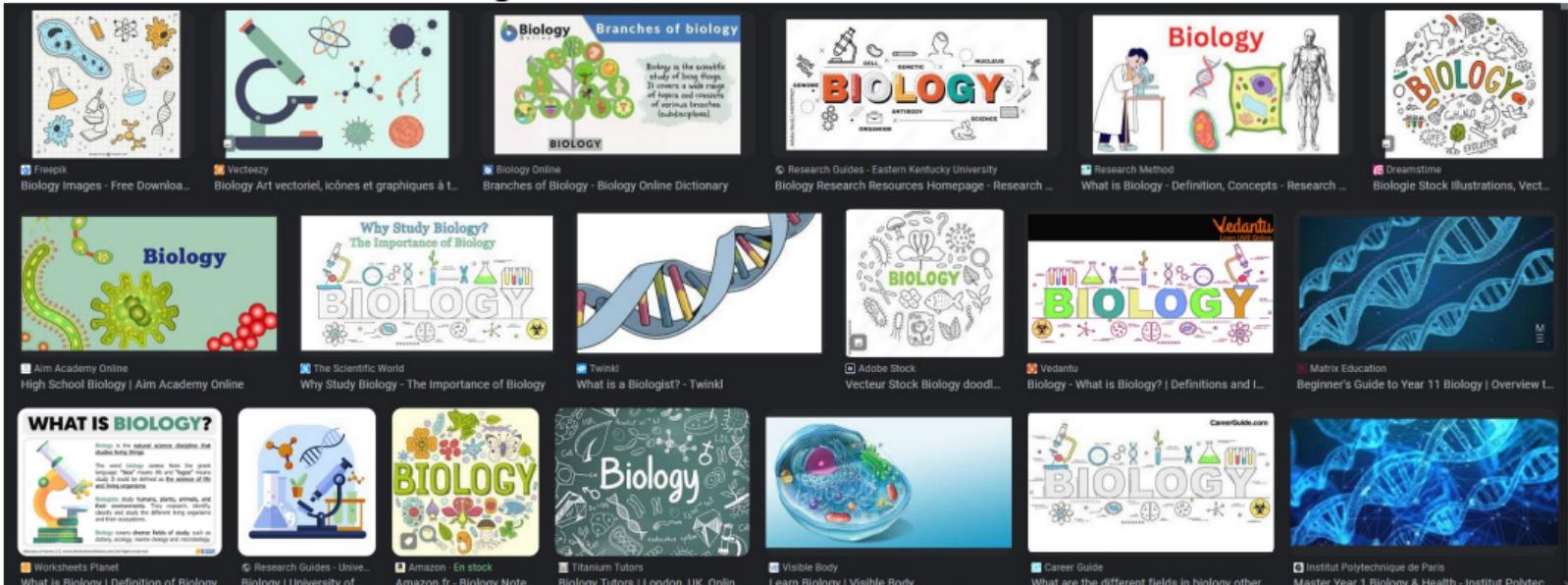
Pensons à une science... la **biologie** :

					
Freepik Biology Images - Free Download...	Vecteezy Biology Art vectoriel, icônes et graphiques à t...	Biology Online Branches of Biology - Biology Online Dictionary	Research Guides - Eastern Kentucky University Biology Research Resources Homepage - Research ...	Research Method What is Biology - Definition, Concepts - Research ...	Dreamstime Biologie Stock illustrations, Vect...
					
Aim Academy Online High School Biology   Aim Academy Online	The Scientific World Why Study Biology - The Importance of Biology	Twinkl What is a Biologist? - Twinkl	Adobe Stock Vecteur Stock Biology doodl...	Vedantu Biology - What is Biology?   Definitions and I...	Matrix Education Beginner's Guide to Year 11 Biology   Overview ...
					
Worksheets Planet What is Biology? Definition of Biology	Research Guides - Univ... Biology   University of...	Amazon - En stock Amazon.fr - Biology Note...	Titanium Tutors Biology Tutors.UU.london, UK. Onlin...	Visible Body I Learn Biology   Visible Body	CareerGuide.com What are the different fields in biology other...

# « Traitement du signal » ???



Pensons à une science... la **biologie** :



- la biologie est faite par les... biologistes :)
- mots-clés : **ADN, cellule, microscope, organisme, vivant...**

Traitement du signal  
○○●○○○○○

Compression - intro  
○○○

Compression sans pertes  
○○○○○○○○○○

Compression avec pertes  
○○○○○○○○○○

Compression neuronale  
○○○○○○○○○○

Résumé  
○○

## « Traitement du signal » ???

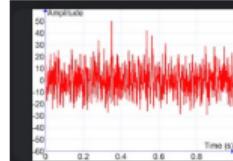


Et maintenant le **traitement du signal** :

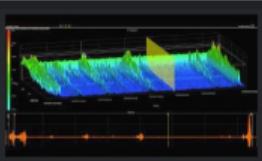
# « Traitement du signal » ???



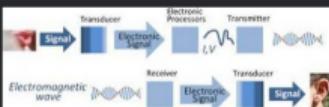
Et maintenant le **traitement du signal** :



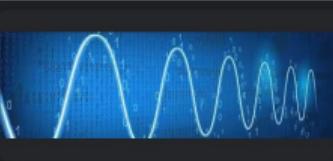
Wikipedia  
Signal processing - Wikipedia



Dewesoft  
What is Signal Processing? | Dewesoft



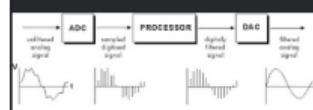
Wikipedia  
Signal processing - Wikipedia



SINTEF  
Signal processing and machine learning



Software Testing Help  
Digital Signal Processing - Complete Guide With Ex...



legacy.cs.indiana.edu  
Digital Signal Processing Overview



Electronics For You  
Signal Processing Algorithms, Applications, ...



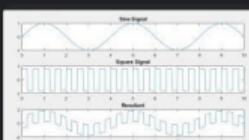
Stanford Online - Stanford University  
Digital Signal Processing Course I Stanford ...



YouTube  
Introduction to Signal Processing - ...



www.uu.nl  
Hundred times faster signal processing ...



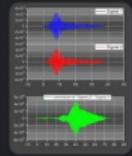
SSLA  
What is basic signal processing? And Defin...



U.S. ASIC  
What is Digital Signal Processing?



Medium  
Filters as an image processing tool ...



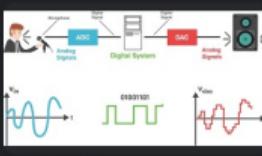
OriginLab  
Signal Processing



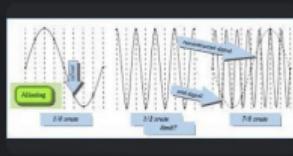
MIT OpenCourseWare ...  
Digital Signal Process...



Wolfram  
Signals, Systems and Signal Proces...



All About Circuits  
An Introduction to Digital Signal Processing - T...

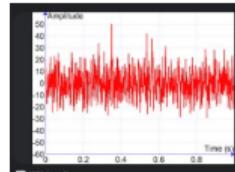


legacy.cs.indiana.edu  
Digital Signal Processing Overview

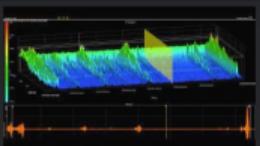
# « Traitement du signal » ???



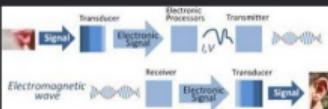
Et maintenant le **traitement du signal** :



Wikipedia  
Signal processing - Wikipedia



Dewesoft  
What is Signal Processing? | Dewesoft



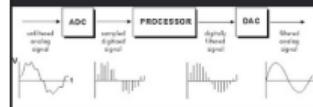
Wikipedia  
Signal processing - Wikipedia



SINTEF  
Signal processing and machine learning



Software Testing Help  
Digital Signal Processing - Complete Guide With Ex...



legacy.cs.indiana.edu  
Digital Signal Processing Overview



Electronics For You  
Signal Processing Algorithms, Applications, ...



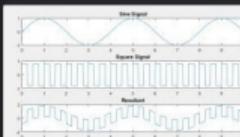
Stanford Online - Stanford University  
Digital Signal Processing Course I Stanford ...



YouTube  
Introduction to Signal Processing - ...



www.uu.nl  
Hundred times faster signal processing ...



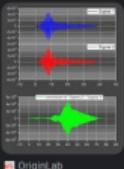
SSLA  
What is basic signal processing? And Defin...



U.S. ASIC  
What is Digital Signal Processing?



Medium  
Filters as an image processing tool ...



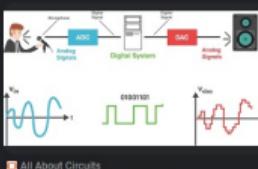
OriginLab  
Signal Processing



MIT OpenCourseWare  
Digital Signal Process...



Wolfram  
Signals, Systems and Signal Proces...



All About Circuits  
An Introduction to Digital Signal Processing - T...



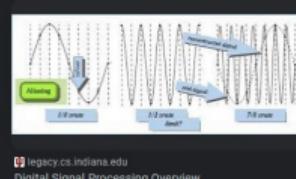
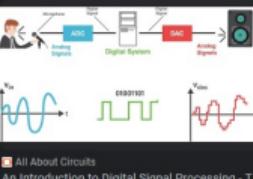
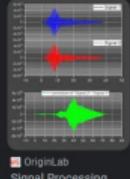
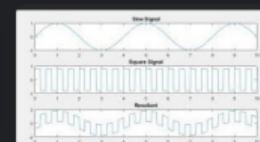
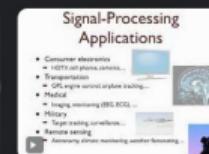
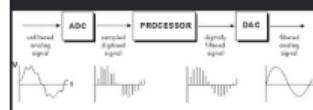
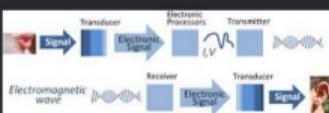
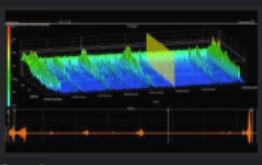
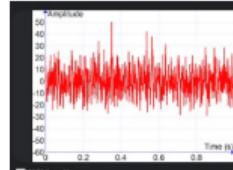
legacy.cs.indiana.edu  
Digital Signal Processing Overview

- le traitement du signal est fait par les... ?

# « Traitement du signal » ???



Et maintenant le **traitement du signal** :

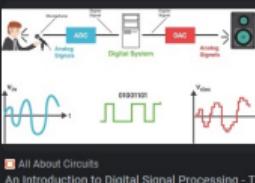
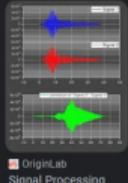
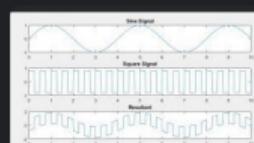
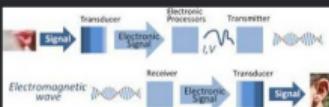
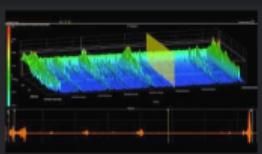
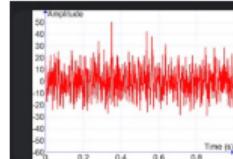


- le traitement du signal est fait par les... ? traiteur-euses ?

# « Traitement du signal » ???



Et maintenant le **traitement du signal** :

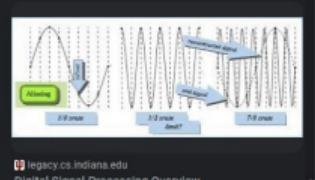
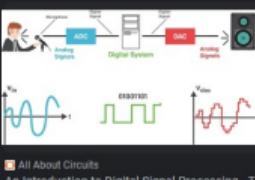
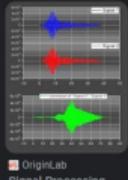
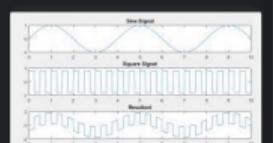
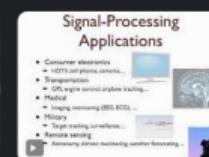
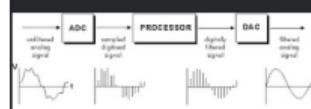
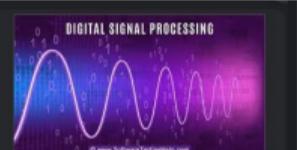
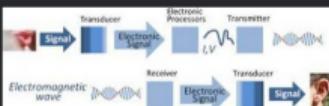
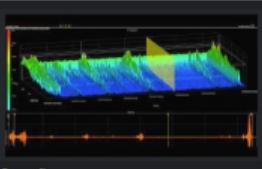
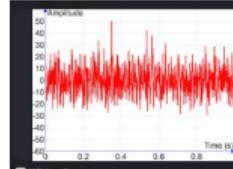


- le traitement du signal est fait par les... ? traiteur-euses ? sémaphoricien-nes ?

# « Traitement du signal » ???



Et maintenant le **traitement du signal** :



- le traitement du signal est fait par les... ? traiteur-euses ? sémaphoricien-nes ?
- mots-clés : **sinusoïde, quantification, binaire**. C'est tout ? :(

## Quelques définitions : « traitement du signal »



### traitement

- analyse
- communication (stockage)
- synthèse

### signal

mesure physique contenant de l'**information**

**information** : ce qui aide à la **décision**

## Quelques définitions : « Décision »



- **binnaire** : gauche/droite, thé/café, fromage/dessert, vrai/faux, ...
- **discrète** (discontinue) : "quel étage?", "quel âge?", ...  
→ décisions binaires en cascade
- **continue** : "quelle température?"  
→ seuil de perception → discontinue

## Fromage ou dessert ?

Le plus simple : un arbre de décision à deux branches  
→ notation binaire ! 0 ou 1



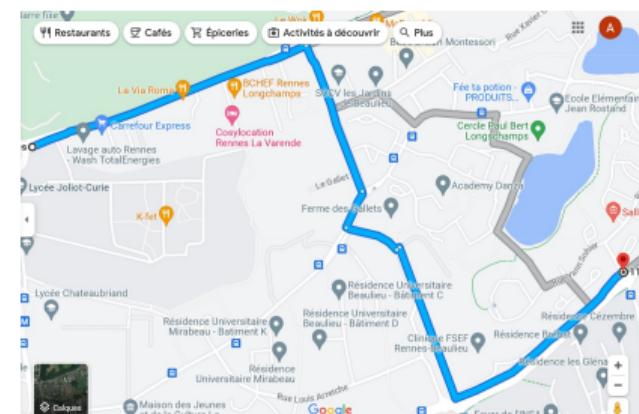
## Itinéraire



Guidage itinéraire : droite/gauche/tout-droit/STOP

Décision	1er bit	2eme bit
Droite	0	0
Gauche	1	0
Tout-droit	0	1
STOP	1	1

droite-droite-gauche-gauche-STOP = 00 00 10 10 11  
→ 10 bits



## Qui est-ce ?



Règle du jeu :

- poser une question fermée (oui/non) à chaque tour
- trouver la personne cachée en un **minimum** de tour

Exemples de questions :

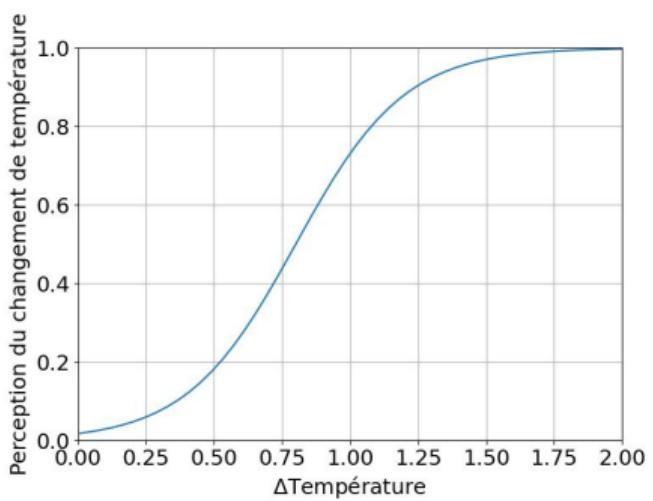
- blond ?
- homme/femme ?
- lunette ?
- ...



# Température



Effet de seuil de perception



Traitement du signal  
oooooooooo

Compression - intro  
●○○

Compression sans pertes  
oooooooooooo

Compression avec pertes  
oooooooooooo

Compression neuronale  
oooooooooooo

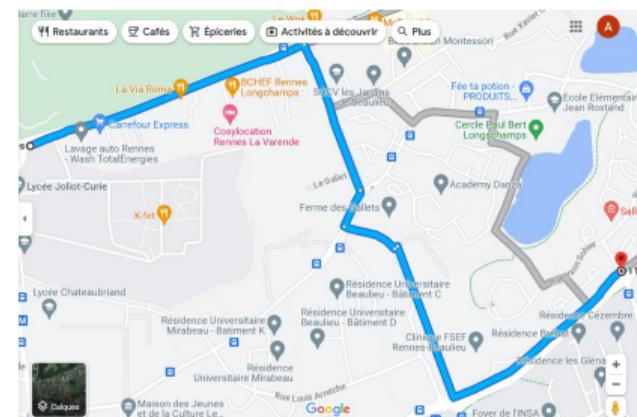
Résumé  
○○

# Compression - intro

## Quelques ordres de grandeur...



- Exemple : itinéraire IUT → Parc des Gayeulles
- Capture d'écran :
  - ▶ **bitmap 24 bits** :  $3270 \text{ ko} = 26'160'000 \text{ bits}$  !
  - ▶ **png** (sans perte) :  $572 \text{ ko} = 4'576'000 \text{ bits}$
  - ▶ **jpg** (avec perte) :  $297 \text{ ko} = 2'376'000 \text{ bits}$
- **old school** : "droite/gauche/tout-droit/STOP" :
  - ▶ droite-droite-gauche-gauche-STOP = 10 bits



## « Codage » ?



### Exemples de <https://fr.wikipedia.org/wiki/Codage>

- Le codage de l'information désigne les moyens de formaliser l'information.
- Un codage des caractères définit une manière de représenter les caractères (lettres, chiffres, symboles) dans un système informatique.
- En psychologie cognitive de la mémoire, le codage est le processus par lequel une information est mise en mémoire.
- Le codage de source, qui permet de faire de la compression de données.
- le codage de canal, qui permet une représentation des données de façon à être résistant aux erreurs de transmission.

Traitement du signal  
oooooooooo

Compression - intro  
ooo

**Compression sans pertes**  
●oooooooooo

Compression avec pertes  
oooooooooo

Compression neuronale  
oooooooooo

Résumé  
oo

# Compression sans pertes

## Itinéraire



Devine un nombre (réponse + ou -) : navigation dans un arbre

- meilleure stratégie avec a priori équiprobable
- meilleure stratégie si 0 la moitié du temps



#2609233

## Quantité d'information



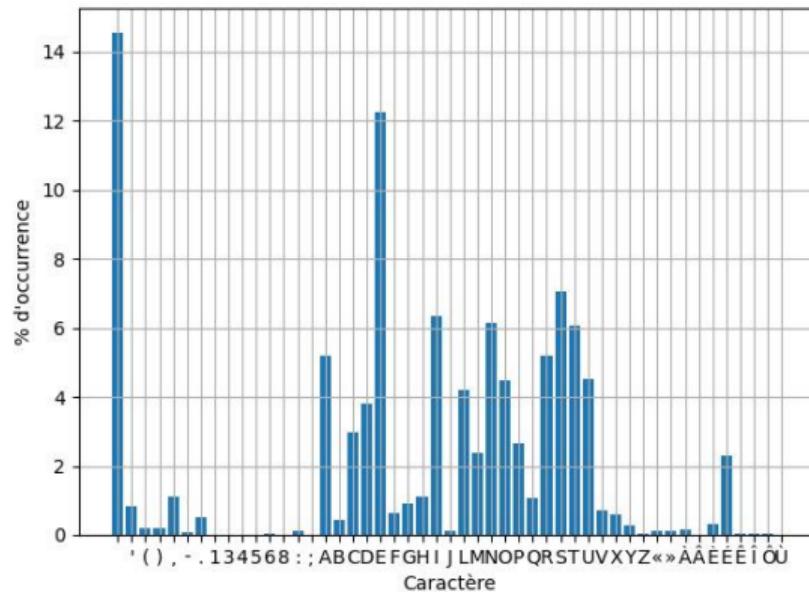
On note  $H(X)$  la quantité d'information contenue dans un message binaire  $X$  :

$$H(X) = p \log_2 \left( \frac{1}{p} \right) + (1 - p) \log_2 \left( \frac{1}{1 - p} \right)$$



## Texte : redondance des caractères

Codage de texte = « Qui est-ce ? »



WIKIPÉDIA  
L'encyclopédie libre

### Rennes

Article Discussion

Lire Modifier Modifier le code Voir l'historique Outils

Ne doit pas être confondu avec Rennes Métropole.

Pour les articles homonymes, voir Rennes (homonymie).

**Rennes** (*ʁɛn*)<sup>1</sup> est une commune du nord-ouest de la France, chef-lieu du département d'Ille-et-Vilaine et de la région Bretagne, la deuxième ville du Grand Ouest et la quinzième commune la plus peuplée de France en nombre d'habitants.<sup>2</sup> L'unité urbaine est peuplée de 367 622 habitants en 2020<sup>3</sup> et son aire d'attraction, qui comprend 763 749 habitants en 2020<sup>4</sup>, est la dixième au niveau national. Rennes est le siège d'une métropole d'462 590 habitants en 2020<sup>5</sup>, faisant ainsi partie des onze grandes métropoles françaises de droit commun (depuis janvier 2015).

À l'époque proto-romaine, la cité fondée par les *Riedones* porte le nom gaulois de *Condate*. La ville voit son pouvoir politique s'accroître au Moyen Âge en devenant successivement forteresse des Marches de Bretagne puis capitale du duché de Bretagne. Sous l'Ancien Régime, fusion de la Bretagne à la France, Rennes progresse rapidement au rang de grande ville provinciale. L'implantation du Parlement de Bretagne à Rennes au XVII<sup>e</sup> siècle puis du palais du Parlement de Bretagne au XVIII<sup>e</sup> siècle a cependant permis à la Bretagne de conserver jusqu'à la Révolution française une certaine autonomie à l'égard du pouvoir royal de l'époque. Rennes a notamment joué un rôle important dans la Révolte du Peuplé limbre en 1675. Victime d'un terrible incendie en 1720, le centre médiéval en bois de la ville est partiellement reconstruit en pierre (granit et tuffeau). Restée majoritairement rurale jusqu'à la Seconde Guerre mondiale, Rennes se développe véritablement au XX<sup>e</sup> siècle.

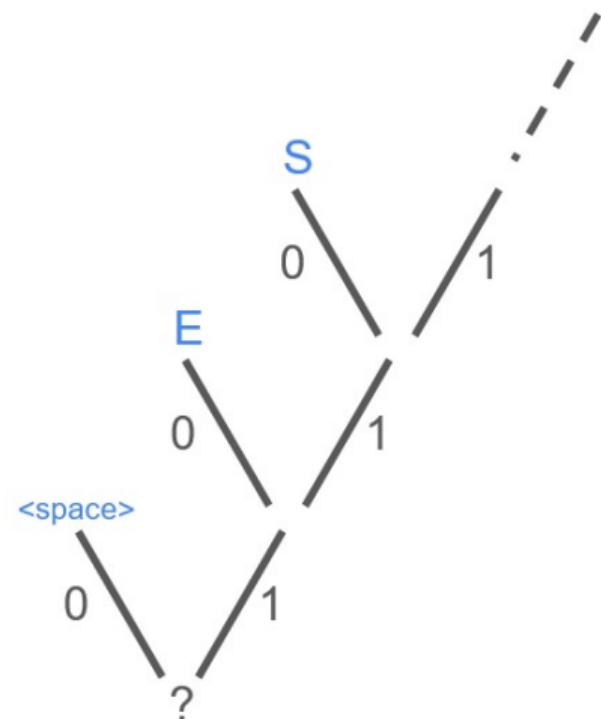
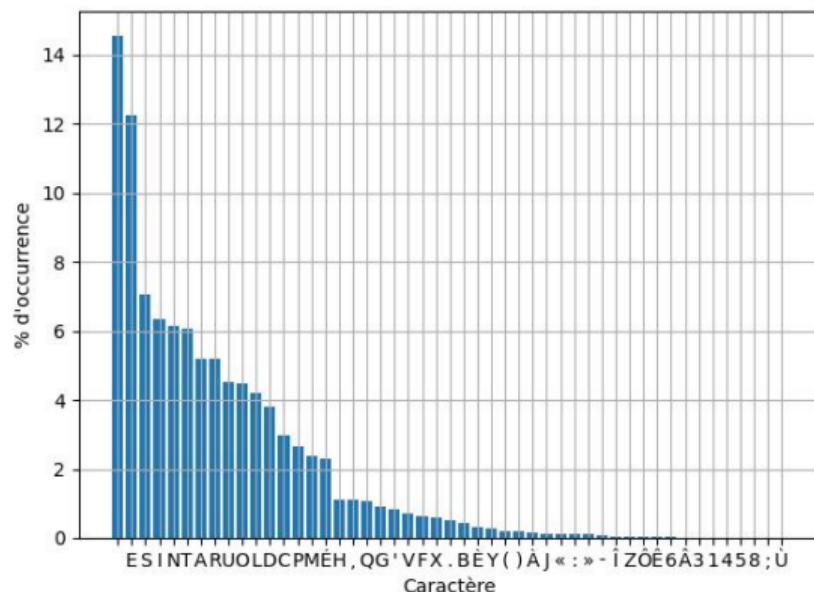
À partir des années 1950, la commune surnommée « ville des administrations » connaît un essor



## Texte : redondance des caractères



Codage de texte = « Qui est-ce ? »





## Applications

Codage entropique :

- Image : PNG (Portable Network Graphics)
- Audio : FLAC (Free Lossless Audio Codec)
- tout type de fichier : ZIP, RAR, ...

## Redondance temporelle



Problème : considérons le message "**ABCDABCDABCDAB**".

Quelle sera la lettre suivante ?

- "A", "B", "C" et "D" sont équiprobables
- Le message n'est pas "redondant" au sens de l'entropie de Shannon
- Pourtant, on devine que la lettre suivante sera "C"



## Quantifier la redondance temporelle

- Pour les redondances de symbole : entropie de Shannon
- Pour les redondances temporelle : entropie de Wiener

$$\xi = \frac{\left(\prod_{i=1}^N P_i\right)^{\frac{1}{N}}}{\frac{1}{N} \sum_{i=1}^N P_i}$$

avec  $P_i$  la puissance du spectre à la fréquence  $i$ .

- $\xi \in [0, 1]$  quantifie l'aplatissement du spectre



## Codage par prédition linéaire

Idée :

- Je peux prédire l'échantillon présent à partir du passé :

$$x_n = \underbrace{f(x_{n-1}, x_{n-2}, \dots, x_{n-P})}_{\hat{x}_n} + \epsilon_n$$

- Quelle approximation pour  $f(\cdot)$ ? → combinaison linéaire

$$\hat{x}_n = \sum_{i=1}^N a_i \cdot x_{n-i}$$

- Décorrélérer le signal  
→ trouver un filtre « blanchisseur »
- le signal « blanchi » est moins puissant que l'original  
→ moins de bits pour une même précision

## Codage par prédition linéaire - exemple



Idée :

- Décorreler le signal  
→ trouver un filtre  
« blanchisseur »
- le signal « blanchi »  
est moins puissant que  
l'original  
→ moins de bits pour  
une même précision



## Applications

Codage entropique :

- Image : PNG (Portable Network Graphics)
- Audio : FLAC (Free Lossless Audio Codec)
- tout type de fichier : ZIP, RAR, ...

Traitement du signal  
oooooooooo

Compression - intro  
ooo

Compression sans pertes  
oooooooooooo

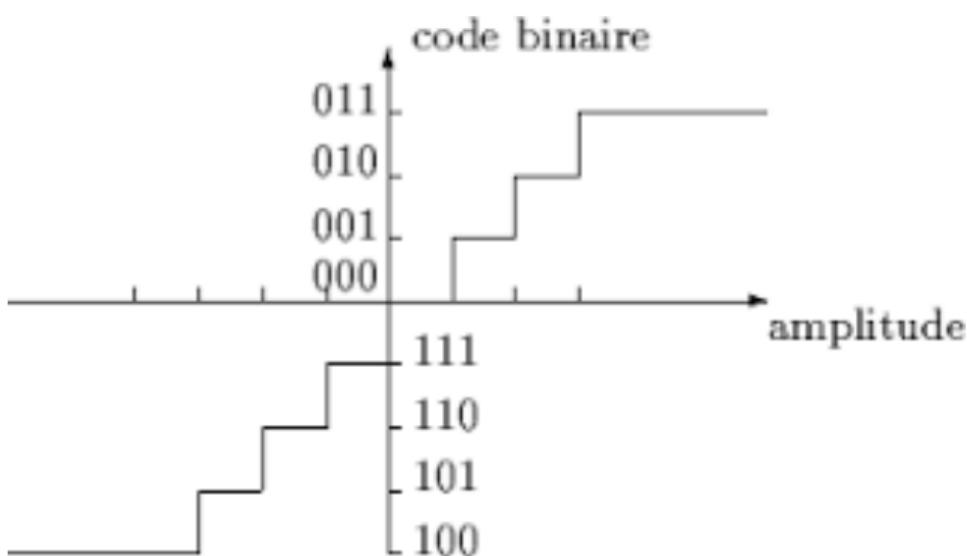
Compression avec pertes  
●ooooooooo

Compression neuronale  
oooooooooo

Résumé  
oo

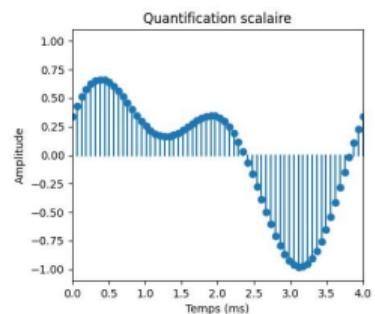
# Compression avec pertes

## Quantification

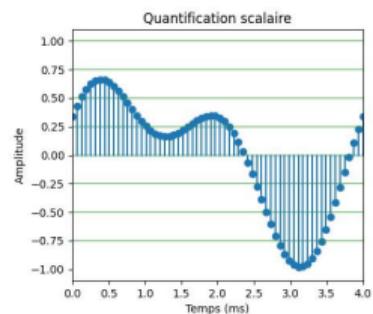


Ouvrage de référence : [Moreau, 2011]

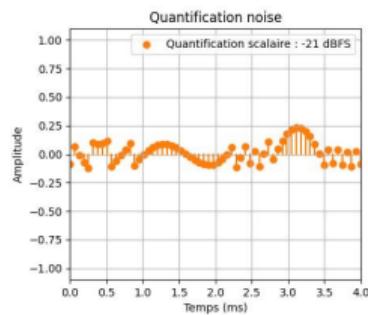
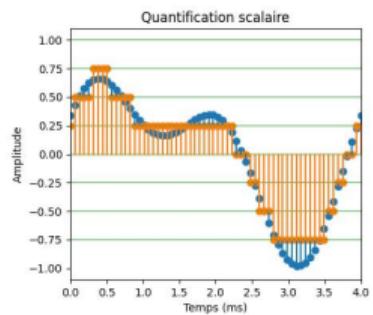
# Codage par transformée : changer de point de vue



## Codage par transformée : changer de point de vue



# Codage par transformée : changer de point de vue



Traitement du signal  
oooooooooo

Compression - intro  
ooo

Compression sans pertes  
oooooooooooo

Compression avec pertes  
oo●oooooooo

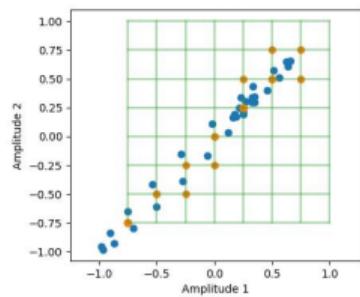
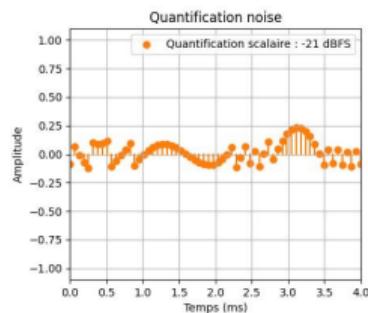
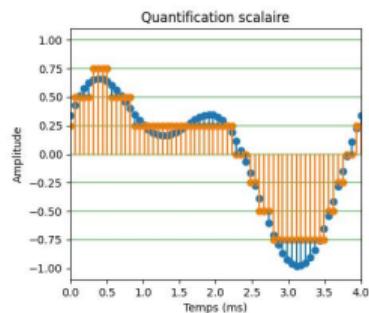
Compression neuronale  
oooooooooo

Résumé  
oo

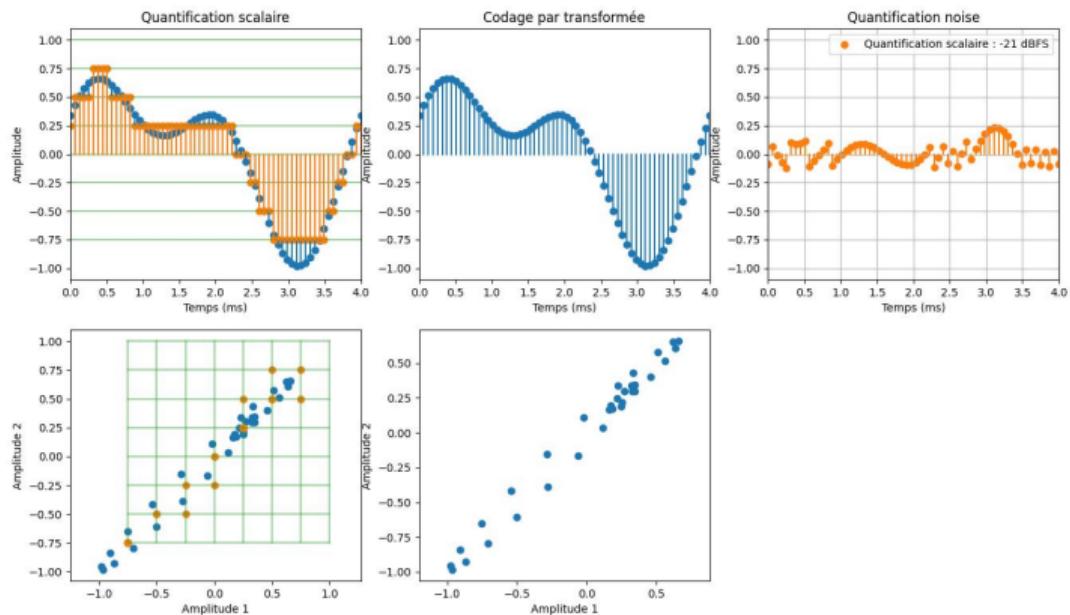
## Codage par transformée : changer de point de vue



## Codage par transformée : changer de point de vue

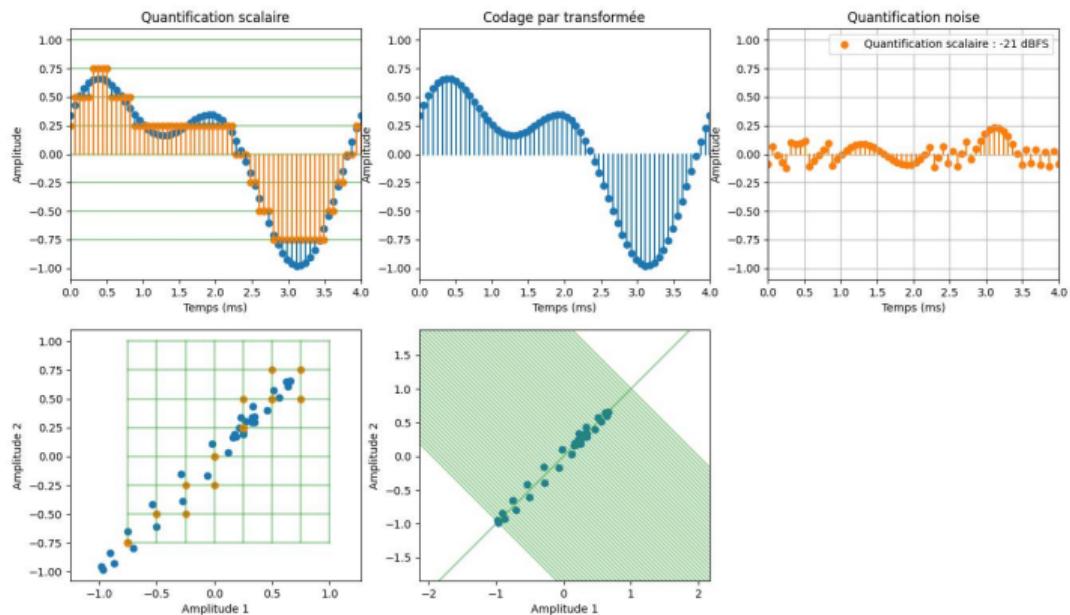


# Codage par transformée : changer de point de vue

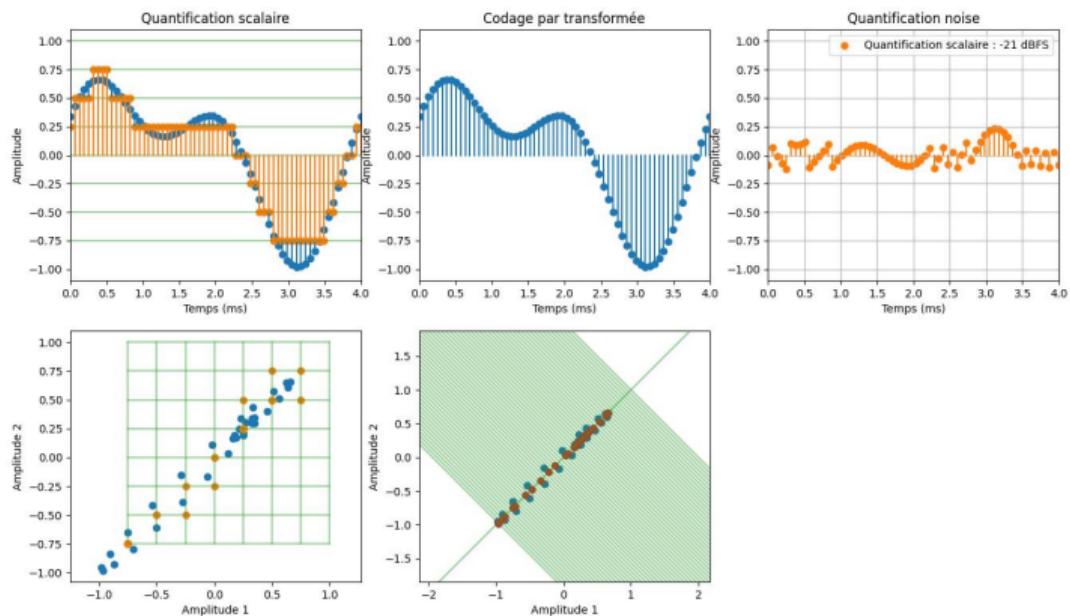




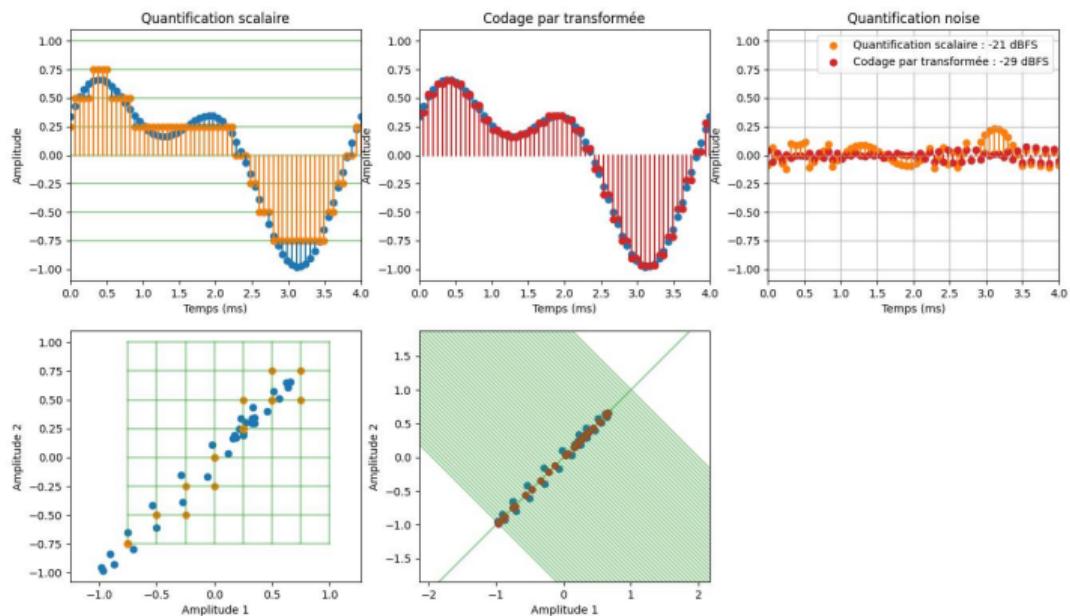
## Codage par transformée : changer de point de vue



# Codage par transformée : changer de point de vue



# Codage par transformée : changer de point de vue



Traitement du signal  
oooooooooo

Compression - intro  
ooo

Compression sans pertes  
oooooooooooo

Compression avec pertes  
oooo●ooooo

Compression neuronale  
oooooooooo

Résumé  
oo

## Codage par transformée : exemple réel sur de la parole



## Codage par transformée : exemples



	WAV 16 bits	Quantif. 8 bits	Transform. 8 bits	Quantif. 4 bits	Transform. 4 bits	MP3 320	MP3 128	MP3 64	MP3 32
Techno									
Voice									
Banjo									

## Quoi perdre ?



- Idée : s'appuyer sur les faiblesses de nos sens
- Exemple (visuel) :
  - ▶ On est plus sensible à la luminosité que la couleur
  - ▶ On entend pas un son faible proche d'un son fort  
[Zölzer, 2008]

## Sous-échantillonnage des couleurs I



RVB → Luminance-chrominanceB-chrominanceR



Traitement du signal  
oooooooooo

Compression - intro  
ooo

Compression sans pertes  
oooooooooooo

Compression avec pertes  
oooooooo●●○

Compression neuronale  
oooooooooo

Résumé  
oo

## Sous-échantillonnage des couleurs II



## Codage par transformée : changer de point de vue



Applications :

- Image : JPEG
- Audio : MP3, AAC

Traitement du signal  
oooooooooo

Compression - intro  
ooo

Compression sans pertes  
oooooooooooo

Compression avec pertes  
oooooooooooo

Compression neuronale  
●oooooooooooo

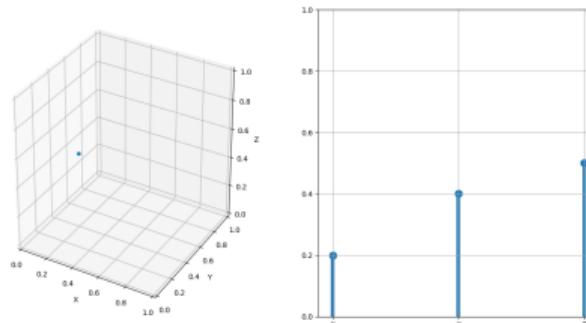
Résumé  
oo

# Compression neuronale



## Notion d'espace d'optimisation

Imaginez, un espace...

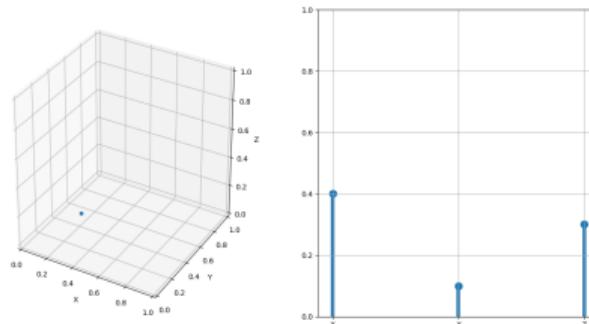


Un vecteur/point de l'espace = un signal



## Notion d'espace d'optimisation

Imaginez, un espace...

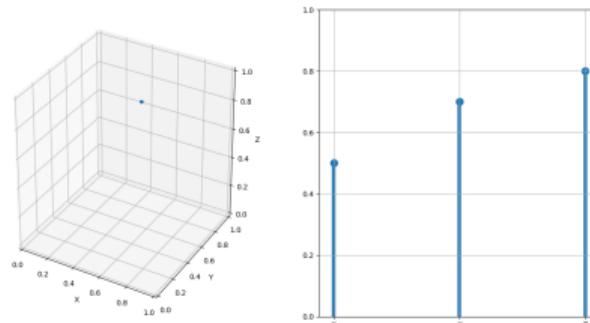


Un vecteur/point de l'espace = un signal



## Notion d'espace d'optimisation

Imaginez, un espace...

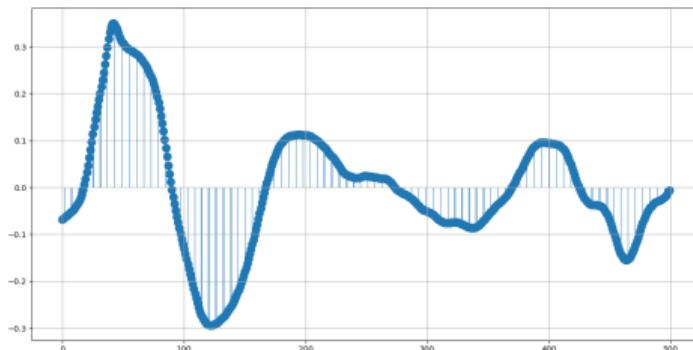


Un vecteur/point de l'espace = un signal



## Notion d'espace d'optimisation

Imaginez, un espace à N dimensions...



- notion importante : le produit scalaire
  - ▶ ressemblance entre deux vecteurs (signaux)
- "Malédiction des grandes dimensions" :
  - ▶  $44.1 \text{ kHz} \times 60 \text{ sec.} \times \text{stereo} = 5,3 \text{ millions de dim. !}$

### Produit scalaire

Soit 2 vecteurs  $x$  et  $y$  en 3D :

$$x = [x_1, x_2, x_3]$$

$$y = [y_1, y_2, y_3]$$

Le prod. scalaire entre  $x$  et  $y$  :

$$\langle x, y \rangle = x_1 \cdot y_1 + x_2 \cdot y_2 + x_3 \cdot y_3$$

$$= \sum_{i=1}^3 x_i \cdot y_i$$

$$= \|x\| \|y\| \cos(\widehat{x, y})$$



# Perceptron (i)

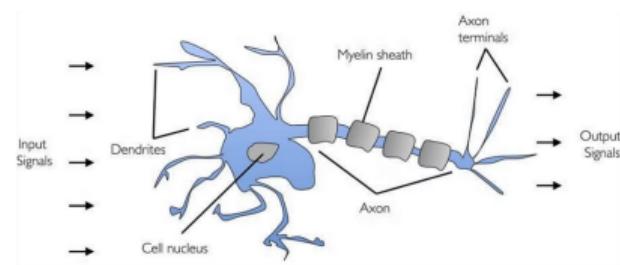
Modèle de neurone de grenouille (Rosenblatt, 1958) :

$$y = f(\theta_0 x_0 + \dots + \theta_{n-1} x_{n-1} + b)$$

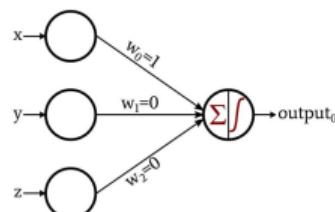
$$y = f \left( \sum_{i=0}^{n-1} \theta_i x_i + b \right)$$

avec  $f(\cdot)$  une fonction non-linéaire.

Analogie : "séparer" l'espace des " $x$ " en deux



*Illustration d'un neurone*



*Modèle simpliste de neurone*

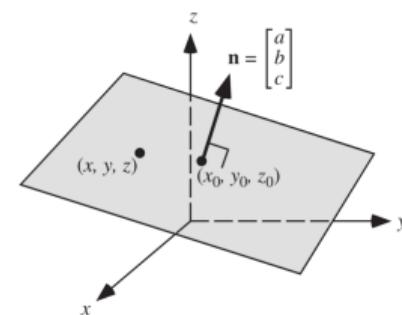
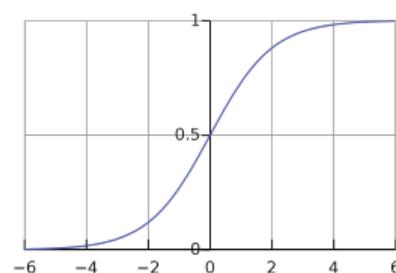


## Perceptron - (ii)

Modèle de neurone de grenouille (Rosenblatt, 1958) :

$$y = f(\theta_0 x_0 + \dots + \theta_{n-1} x_{n-1} + b)$$

$$y = f \left( \sum_{i=0}^{n-1} \theta_i x_i + b \right)$$

avec  $f(\cdot)$  une fonction non-linéaire.Analogie : "séparer" l'espace des " $x$ " en deuxPlan défini par le vecteur  $n$ 

Fonction d'activation sigmoïde



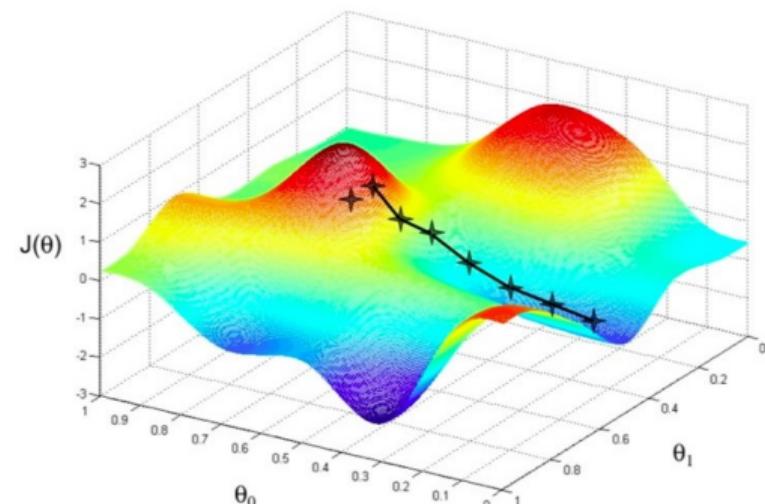
## Procédure d'apprentissage

On cherche la combinaison des paramètres  $\theta_0, \theta_1, \dots$  qui minimise la mesure d'erreur  $J(\theta_0, \theta_1, \dots)$

### Algorithme de la "descente de gradient"

- suivre la pente descendante de la métrique d'erreur
- (besoin d'un réseau de neurone "dérivable")
- (on rétropropage le gradient de l'erreur pour actualiser les  $\theta_i$ )

En pratique, on a  $\sim 10$  millions de  $\theta_i$  !



**Descente de gradient**



## Procédure d'apprentissage

On cherche la combinaison des paramètres  $\theta_0, \theta_1, \dots$  qui minimise la mesure d'erreur  $J(\theta_0, \theta_1, \dots)$

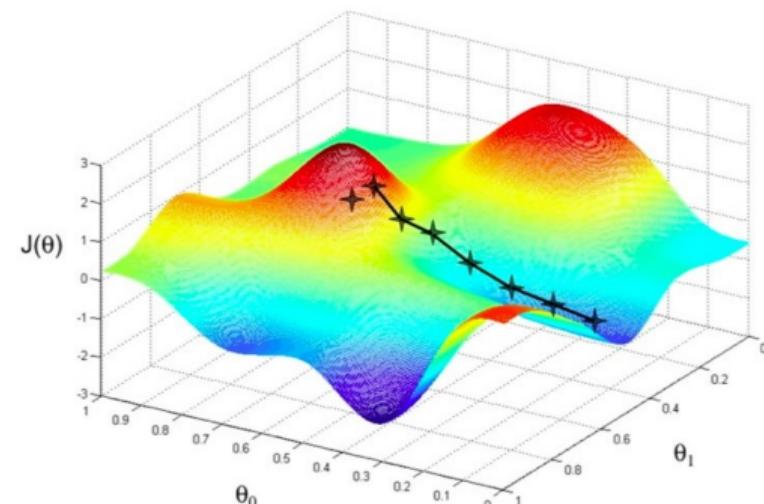
### Algorithme de la "descente de gradient"

- suivre la pente descendante de la métrique d'erreur
- (besoin d'un réseau de neurone "dérivable")
- (on rétropropage le gradient de l'erreur pour actualiser les  $\theta_i$ )

En pratique, on a  $\sim 10$  millions de  $\theta_i$  !

Challenges :

- trouver une mesure d'erreur  $J(\theta_0, \theta_1, \dots)$  pertinente ET "dérivable"
- avoir beaucoup de données

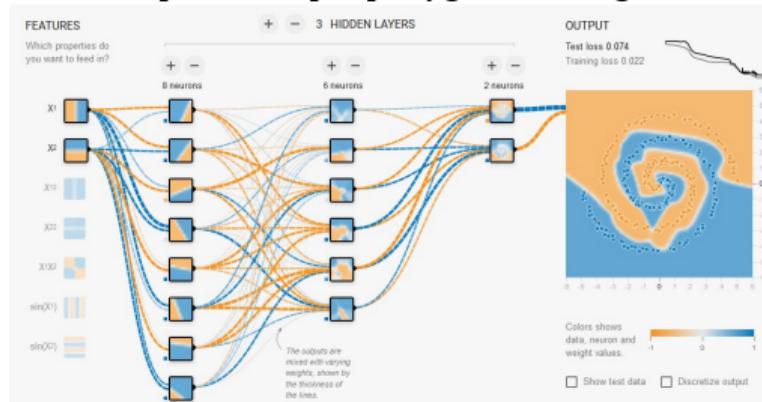


Descente de gradient

# Réseau de neurones : approximateur de fonction



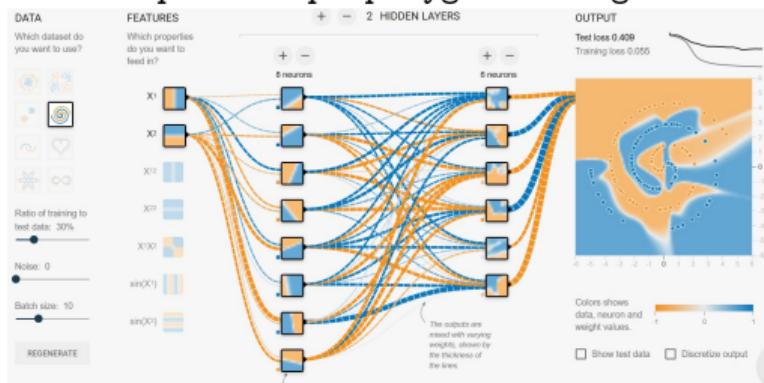
Voir <https://deeperplayground.org/>



# Réseau de neurones : approximateur de fonction



Voir <https://deeperplayground.org/>

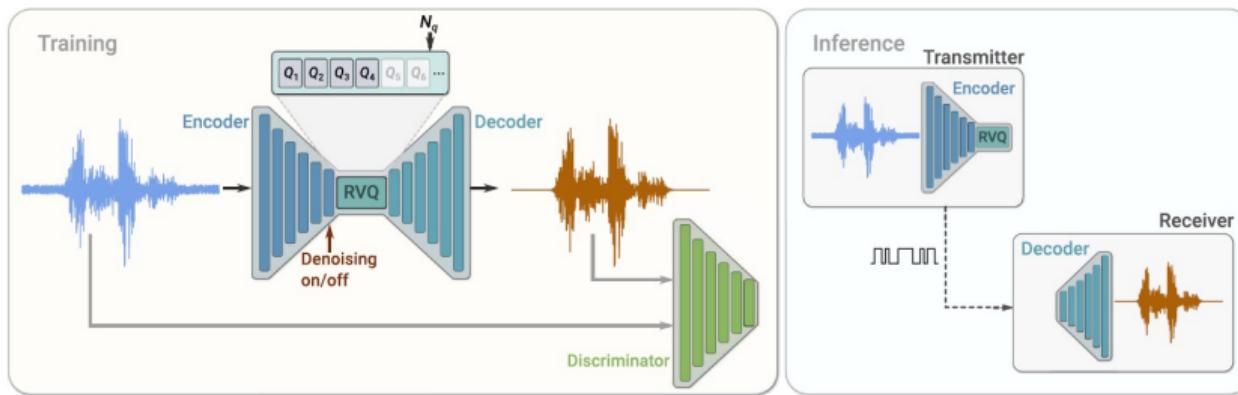


Sur-apprentissage (overfitting) :

- réseau complexe
- peu de données



## Réseau de neurones auto-encodeur



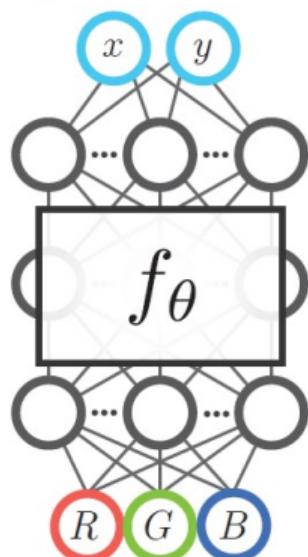
[Zeghidour et al., 2022], [Kumar et al., 2023]

- Dernière innovation de Google et Descript
- Exemples audio : <https://descript.notion.site/Descript-Audio-Codec-11389fce0ce2419891d6591a68f814d5>

# Réseau de neurones : perspectives



Input Coordinate



Value at  
Coordinate

Stratégie similaire au codage par prédiction linéaire

Traitement du signal  
oooooooooo

Compression - intro  
ooo

Compression sans pertes  
oooooooooooo

Compression avec pertes  
oooooooooooo

Compression neuronale  
oooooooooooo

Résumé  
●○

# Résumé

## Taux de compression : ordres de grandeur



Codec	Débit (kbits/s)	Taux de compression
WAV	1411	-
FLAC (sans perte)	420 - 700	2 - 5
MP3 (avec pertes)	128 - 320	4.5 - 11
Audio Descript (DNN avec pertes)	2 - 8	175 - 700

$$\text{WAV} : \underbrace{2}_{\text{stereo}} \times \underbrace{44100}_{f_s} \times \underbrace{16}_{\text{nb. bits}} = 1411 \text{ kbits/s}$$

## Taux de compression : ordres de grandeur



Codec	Débit (kbits/s)	Taux de compression
WAV	1411	-
FLAC (sans perte)	420 - 700	2 - 5
MP3 (avec pertes)	128 - 320	4.5 - 11
Audio Descript (DNN avec pertes)	2 - 8	175 - 700

$$\text{WAV} : \underbrace{2}_{\text{stereo}} \times \underbrace{44100}_{f_s} \times \underbrace{16}_{\text{nb. bits}} = 1411 \text{ kbits/s}$$

- Compression sans pertes : supprimer les redondances
- Compression avec pertes : supprimer les informations (décisions) les moins significatives
- Compression neuronale (avec pertes) : basé sur l'apprentissage (données et critères d'évaluation)

-  **Kumar, R., Seetharaman, P., Luebs, A., Kumar, I., and Kumar, K. (2023).**  
**High-Fidelity Audio Compression with Improved RVQ GAN.**  
arXiv :2306.06546 [cs, eess].
-  **Moreau, N. (2011).**  
*Tools for signal compression.*  
ISTE, London.  
OCLC : 723524173.
-  **Tandik, M., Mildenhall, B., Wang, T., Schmidt, D., Srinivasan, P. P., Barron, J. T., and Ng, R. (2021).**  
**Learned Initializations for Optimizing Coordinate-Based Neural Representations.**  
In *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2845–2854, Nashville, TN, USA. IEEE.
-  **Zeghidour, N., Luebs, A., Omran, A., Skoglund, J., and Tagliasacchi, M. (2022).**  
**SoundStream : An End-to-End Neural Audio Codec.**  
*IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 30 :495–507.
-  **Zölzer, U. (2008).**  
*Digital Audio Signal Processing*.  
Wiley, 1 edition.