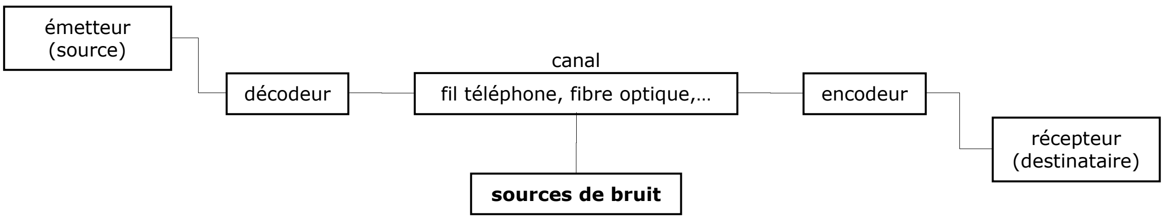
4.3. TRANSMISSION DE L'INFORMATION.

4.3.1. SHANNON ET SON ŒUVRE :

Biographie :
Claude Shannon (1916-2001) était un mathématicien. Il a travaillé dans les labos de recherche de la Bell Labs (compagnie de téléphone américaine). Il est l'auteur de « A Mathematical Theory of Communication ». Il a également donné cours au MIT (Boston). Voir annexe pour plus de renseignements…

Domaine :
- transmission d'un signal (data rate=taux de transmission)
- codage, compression
- échantillonnage

Modèle de Shannon :
Le processus de transmission est essentiellement aléatoire. L'information provenant d'une source est transmise à travers un canal (il y a une certaine probabilité que l'information parvienne à la sortie). Le signal est encodé à l'entrée et décodé à la sortie. Il y a des sources de bruit qui se mêlent au signal transmis.
Les mathématiques et les concepts intervenant dans son modèle sont relativement complexes, en particulier Shannon utilise le concept d'entropie (utilisé en thermodynamique). Chaque variable (séquence) aléatoire possède une entropie, « l'information mutuelle » entre deux variables peut être calculée.



On se pose la question de savoir à quelle vitesse et avec quelle fiabilité l'information sera transmise par le canal, on agira sur le processus d'encodage/décodage afin d'obtenir les meilleures performances.
Shannon donne une limite théorique à la performance atteinte, limite qui dépend de la source et du canal : quel que soit le processus d'encodage utilisé, on ne pourra franchir cette limite.

Nombre de bits par seconde, bande passante :

BPS : nombre de bits/seconde = $\boxed{BW.\log_2\left(1+10^{SNR/10}\right)}$ [bits/s]

BW : bandwith (bande passante) [Hz]

SNR : Signal to Noise Ratio (rapport signal bruit) [dB]

Décibel : rapport logarithmique entre 2 valeurs (ici, le signal et le bruit)

Avec cette formule, on peut calculer le taux de transmission théorique d'un canal (ligne) de bande passante connue et perturbé par un certain bruit.

Rappel :

$$\log_2 x = \frac{\log_{10} x}{\log_{10} 2} \approx \frac{\log_{10} x}{0{,}3}$$

Il existe une autre formule permettant de calculer le BPS : $\boxed{BW.\log_2(1+SNR)}$ avec SNR=rapport signal/bruit donné sous forme de rapport de puissances

$$\left.\frac{U = R.I}{P = R.I^2}\right\} P = \frac{U^2}{R}$$

$$SNR = 20.\log_{10}\frac{U_S}{U_N}$$

U_S=DDP (en volts) du signal et U_N=DDP (en volts) du bruit
P_S=puissance (en watts) du signal et P_N=puissance (en watts) du bruit

$$SNR = 10.\log_{10}\frac{P_S}{P_N}$$

ex :
Une ligne téléphonique a une bande passante de 300 à 3400 Hz et son SNR est de 30 dB, quel est son BPS ?
BW = 3100 Hz

$$BPS = 3100.\log_2\left(1+10^{30/10}\right) = 3100.\frac{\log_{10}(1001)}{\log_{10}2} = \frac{9300}{0{,}3} = 31000 \text{ bits/sec}$$

ex avec le SNR en puissance :

si $\frac{P_S}{P_N} = 1000 \rightarrow SNR = 10.\log_{10}1000 = 10.3 = 30$ dB

4.3.2. Compression de données :

La signification de l'information (théorie de Shannon) :
Revenons à l'époque du télégraphe, on cherchait déjà à obtenir une utilisation optimale de l'infrastructure existante (fils de cuivre aériens). On transmettait des messages du genre : « arriverai jeudi soir Paris gare Nord 20h prière venir chercher ». L'information transmise reste compréhensible par le destinataire (on a omis les articles : ce, à, la, me …).
Si on pousse un peu trop loin, le destinataire risque de se retrouver avec un message incompréhensible, on dira que les symboles ou lettres manquant(e)s sont imprédictibles (ex : « seul infrmatn esnil à comred dt et trami », le message de départ étant « seule l'information essentielle à comprendre doit être transmise »).

Les symboles manquants mais prédictibles sont dits redondants. Le concept d'imprédictibilité ou incertitude est primordial, la source pourra être décrite par son entropie, on peut déterminer le plus petit nombre de bits par symbole requis pour représenter la sortie (on peut montrer que les lettres d'un texte ne peuvent être comprimées en deçà de 1,5 bit par lettre).

En pratique, une partie de l'information transmise est redondante, ce qui permet une réception fiable de l'info effective (hors redondances). C'est le code de correction d'erreur qui inclura la redondance. A nouveau, la théorie nous montre qu'il est impossible d'atteindre une probabilité d'erreur aussi petite que l'on veut (la théorie fixe une limite), cependant la complexité du code augmente en fonction de la compression atteinte.

ex : comptage de voitures circulant sur une autoroute

Couleur	Quantité	Codage binaire	Binaire optimisé
Noires	50%	00	0
Blanches	25%	01	10
Rouges	12,5%	10	110
Bleues	12,5%	11	111

Avec le codage optimisé, on obtient une moyenne de 1,75 bit pour coder l'information (1 bit x 0,5 + 2 bits x 0,25 + 2 x 3 bits x 0,125) au lieu de 2 bits pour un codage classique. Serait-il possible d'encore réduire le nombre de bits? Non, la théorie nous dit que la valeur théorique optimale pour ce problème est 1,75 bit/couleur. Il est donc inutile de chercher un autre codage.

En théorie, lors du transport d'informations, il existe une valeur maximale (BPS) et une valeur minimale (compression).
En pratique, il faut trouver un procédé de modulation pour le BPS et un procédé de codage (compression) pour atteindre la valeur minimale.

4.3.3. Bauds et bits :

Jusqu'à présent, on a mesuré la capacité du canal en termes de débit d'information (BPS). Une autre unité souvent rencontrée est la rapidité de modulation ou vitesse de transmission qui mesure le nombre de variations du signal par seconde, l'unité utilisée est le baud/seconde. Une variation est une transition entre 2 états (niveaux d'amplitudes, niveaux de phases, niveaux de fréquences).

Lien entre le débit d'information et la vitesse de transmission :
BPS=BR.n avec BR le baud rate et n le nombre de bit(s) transmissible(s) par variation du signal (dépend de la technique de modulation utilisée).

ex :
QAM, 1200 bauds, codage de 8 valeurs par baud
BPS=1200.3=3600 bits/s

Standards :

Norme	bauds/s	Technique	bits/baud	BPS
V26	1200	4 φ	2	2400
V27	1600	8 φ	3	4800
V29	2400	8 φ & 2 A	4	9600

ex :
BR=1000bauds/s, QAM avec 4 valeurs de phase et 2 valeurs d'amplitudes, BPS=?

Nombre de valeurs possibles=4.2=8 ($\rightarrow$ 3 bits)
BPS=1000.3=3000 bits/s

4.4. Echantillonage - Digitalisation des signaux.

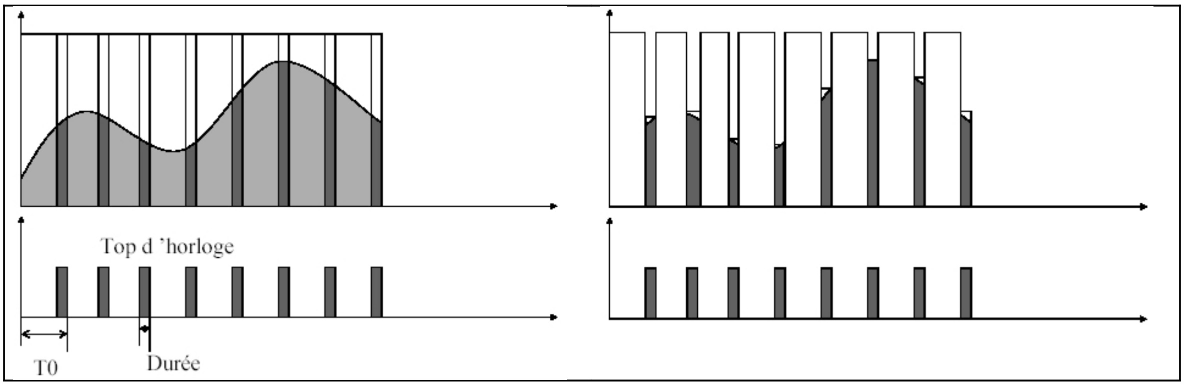
4.4.1. Théorème de Shannon-Nyquist :

On peut reconstruire un signal à partir des échantillons si la fréquence d'échantillonnage ϕ_e vaut au moins le double de la fréquence maximale ϕ_{max} contenue dans la signal.

4.4.2. Echantillonage et quantification :

On calculera le data rate (BPS) comme suit :
BPS=ϕ_e.n avec n le nombre de bits utilisés pour la quantification

NB : $T_e=1/\phi_e$ (période d'échantillonnage)



Tous les k_{Te} ($k \in N_0$), on « mesure » la valeur analogique $x(k_{Te})$. Cette valeur analogique est ensuite représentée par un nombre binaire tel que $x(k_{Te}) \approx 2^n$, n dépend de la valeur maximale de x(t) et de la précision qu'on désire atteindre.
Pour permettre la quantification, il faut bloquer la valeur du signal pendant le temps nécessaire, ce temps doit être << T_e.

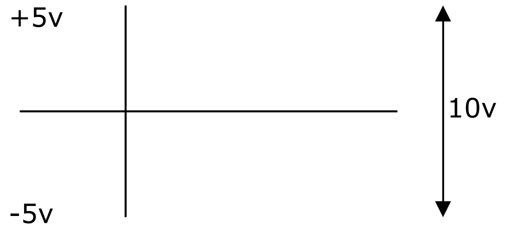
ex :
Si U_{max}=8v et si on choisit n=3, on peut représenter $2^n=2^3=8$ valeurs différentes, on aura une précision de 1v.
Si par exemple, $x(T_e)$=5,2v, on la représente par 101 en binaire ($101_2=5_{10}$) et on commet une erreur absolue de 0,2v.
Quelque soit le n utilisé (si grand soit-il), on aura toujours une erreur appelée bruit de quantification (traduit à l'oreille par un « bourdonnement »). Le rapport signal/bruit obtenu est donné par

SNR=6.n+2[dB].

<u>Quantification linéaire :</u>

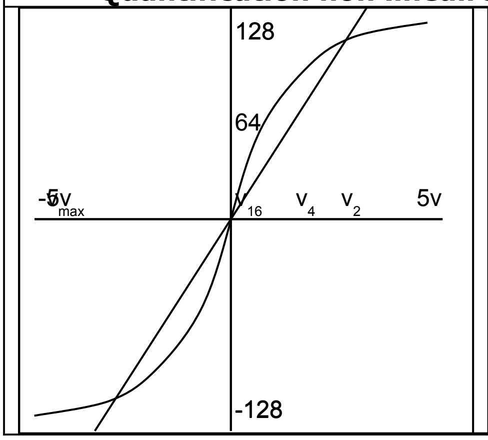
La valeur binaire est proportionnelle à la valeur analogique (111 pour 7V et 011 pour 3V par exemple si on travaille avec n= 3 bits), si le signal est faible, le rapport signal bruit effectif sera médiocre (surtout si n est petit).

ex : signal de « qualité téléphonique » amplitude faible

	n=8 bits → 256 niveaux bruit de quantification (linéaire) 10V/256 ≈ 40 mV

Le SNR à l'amplitude max : SNR=6.8+2=50 dB
Si le signal est faible (Y=1v), le SNR est d'environ 32dB

Quantification non linéaire :
Pour résoudre ce problème, on utilise une quantification non linéaire favorisant les faibles amplitudes.

Quantification non linéaire	
	

L'avantage est que, même pour un niveau faible, la valeur binaire est élevée.

ex :
pour $U_{max}/16$=5v/16≈ 0,3v
64 niveaux → 2^6 → n=6
SNR=6.6+2=38dB

Le standard de quantification non linéaire « loi A » est un standard européen, il existe un standard US similaire appelé « μ law ».

4.4.3. Applications

CD Audio :
f_{max}signal = 20 kHz ϕ_e = 44,1 kHz (> 2 x 20 kHz)
n=16bits, son stéréo (2 signaux)
BPS = **2**.44100.16 = 1,41 Mbits/s
SNR = 6.16+2 = 98 dB

Provenance de cette ϕ_e : standard de codage des couleurs TV appelé PAL
(625 lignes – 37 lignes non utilisées) x 3 échantillonnages x 25 images/seconde
→ 44100 échantillons/seconde

<u>Son en « qualité téléphonique » :</u>
f_e = 8 kHz, n=8 bits, son mono
BPS = 8000.8 = 64 kbits/s
SNR = 6.8+2 = 50 dB

<u>µ law → « voix » :</u>
Différentes possibilités de fréquence : f_e =8,012 kHz ou 11,025 kHz ou 44,1 kHz
n=8 ou 16 bits

<u>GSM :</u>
f_e =8 kHz (voix à l'origine), n=13 bits
BPS = 8000.13 = 104 kbits/s

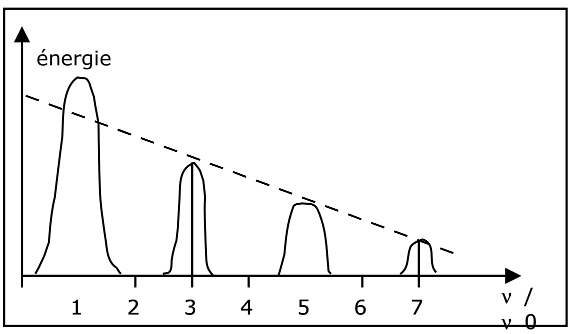
L'introduction des corrections d'erreurs et du cryptage (64 bits → assurant une relative confidentialité) portent le débit à environ 200 kbits/s. Pour réduire ce débit, on utilise un codec (vocodeur) qui ramène le débit initial de 104 kbits/s à 13 kbits/s.

4.4.4. Explication intuitive du théorème de Shannon :

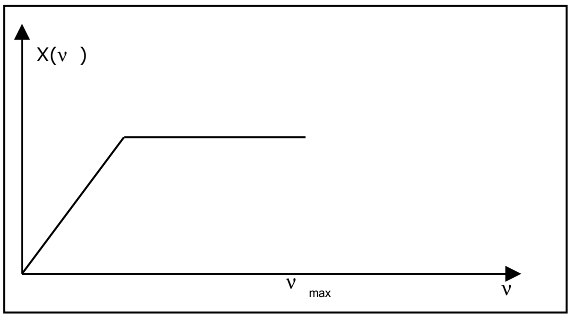
Rappel :
Un signal est une série infinie de termes multiples de f parfois notée ν . On peut donner une représentation temporelle du signal mais aussi une représentation en fréquence (spectre) du signal.

ex :

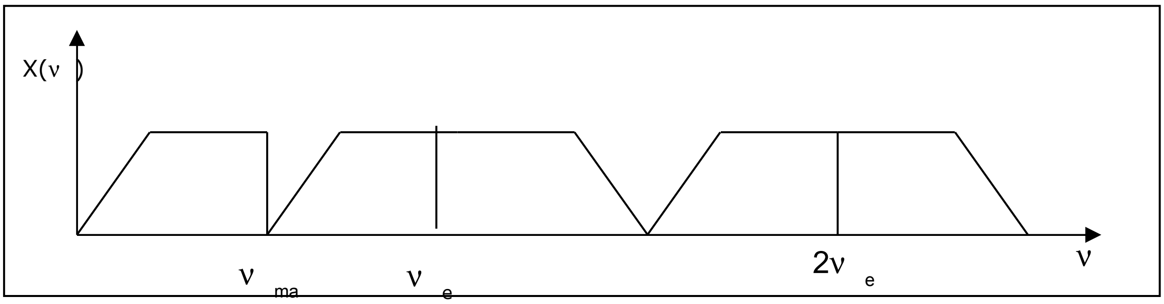
Ce schéma est adéquat pour des sinusoïdes pures. En pratique, le spectre est « dispersé ».



Supposons que le spectre du signal à échantillonner soit le suivant :



On peut montrer que le spectre du signal échantillonné aura la forme suivante



Si $\nu_e \geq 2\nu_{max}$, alors^{x}je peux utiliser un filtre passe-bas pour reconstituer le signal original. Si $\nu_e < 2\nu_{max}$, le spectre du signal échantillonné va présenter des superpositions et on va obtenir de la distorsion (on ne peut plus reconstituer « proprement » le signal original).