

République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université Abderahmane MIRA - BEJAIA -

Département de GENIE ELECTRIQUE



جامعة بجاية
Tasdawit n'Bgayet
Université de Béjaïa

Communications Numériques 1 (UEF 21)

Master 1 Télécom

Dr. Mohamed AZNI

Table des matières

Liste des figures	iii
1 Introduction aux communications numériques	1
1.1 Introduction	1
1.2 Système de communications numériques	1
2 Codage de sources analogiques	5
2.1 Introduction	5
2.2 Échantillonnage	5
2.3 Quantification	6
2.3.1 Quantification scalaire	6
2.3.1.1 Quantification classique	7
2.3.1.2 Quantification optimale	9
2.3.2 Quantification vectorielle	10
3 Notions de théorie de l'information	11
3.1 Introduction	11
3.2 Rappels sur les probabilités	11
3.2.1 Moyenne statistique et moments	12
3.3 Introduction à la théorie de l'information	13
3.3.1 Quantité d'information d'un symbole	13
3.3.2 Quantité d'information moyenne (Entropie)	14
4 Codage de source discrètes	17
4.1 Introduction	17

4.2	Codes a longueur fixe	17
4.3	Codes à longueur variable	18
4.3.1	Codes sans préfixe	19
4.3.2	Construction de codes sans préfixe	20
4.3.3	Inégalité de Kraft pour les codes sans préfixe	21
4.3.4	Algorithme de Huffman pour le codage entropique	22
4.4	Codage de sources avec mémoire	23
4.4.1	Sources de Markov	23
4.4.2	Codage de sources de Markov	25
5	Canaux de transmission	27
5.1	Introduction	27
5.2	Bruit additif	27
5.3	Modélisation du bruit	28
5.3.1	Processus stochastique	28
5.3.1.1	Processus gaussien	29
5.3.1.2	Processus stochastique stationnaire	31
5.3.1.3	Densité de probabilité d'un ensemble de variables conjointement gaussiennes	31
5.3.1.4	Bruit blanc gaussien	32
5.4	Rapport signal sur bruit	32
5.5	Capacité d'un canal de transmission	33
5.5.1	Efficacité spectrale	33
6	Détection	35
6.1	Introduction	35
6.2	Principes de la détection	35
6.3	Détection en binaire	36
6.4	Détection sur canal AWGN	37
	Bibliographie	41

Liste des figures

1.1	Systeme de communication numérique	2
1.2	Modèle de canal de transmission	3
2.1	Quantification scalaire	7
4.1	Arbre de décomposition pour la construction de codes sans préfixe . . .	20
4.2	Réduction de l'arbre de la figure 4.1	20
4.3	Algorithme de Huffman	23
4.4	Exemple de source de Markov	24
5.1	Modèle de canal avec bruit additif	27
5.2	Modèle de canal de transmission à bande passante limitée	28
5.3	Distribution gaussienne ou normale	30
5.4	Allure de la capacité du canal en fonction de la bande passante W . . .	33
6.1	Détection sur canal AWGN en utilisant une modulation BPSK	37
6.2	Distributions normales conditionnées sur les hypothèses $+a$ et $-a$. . .	39

Introduction aux communications numériques

1.1 Introduction

La plupart des applications de communication (radio, télévision, réseaux de communication cellulaires, etc...) ont été numérisées depuis maintenant quelques dizaines d'années. Les communications numériques offrent deux avantages principaux par rapport aux communications analogiques.

- Immunité au bruit.
- Se prêtent plus facilement aux différents traitements de l'information.

Remarque 1 *Discuter les observations ci-dessus*

1.2 Système de communications numériques

L'architecture générique d'une chaîne de communication numérique est donnée par la figure 1.1. Ci-dessous, nous décrivons brièvement la fonction de chacun des blocs qui la constituent.

1. Source d'information : On distingue trois formes de sources d'information :
 - **Sources discrètes** : la sortie d'une source discrète est une séquence de symboles extraits à partir d'un certain alphabet discret et *fini*. Il est souvent convenable, dans ce cas, de supposer que la séquence de symboles est délivrée par la source à une certaine fréquence. Notons que si la source est discrète, il n'est pas nécessaire de la quantifier. Une source discrète est directement codable.
 - **Sources analogiques** : la sortie d'une source analogique est un signal *continu* $x(t)$. C'est le type de source le plus répandu. $x(t)$ représente une grandeur

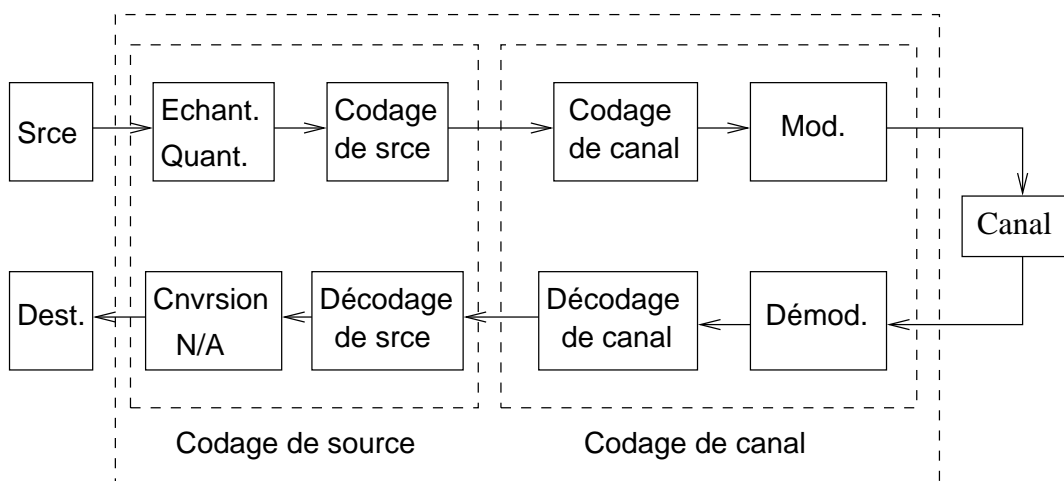


Figure 1.1 : *Système de communication numérique*

physique. Par exemple, il peut représenter un signal audio ou vidéo. D'une manière formelle, $x(t)$ est une application de \mathbb{R} dans \mathbb{R} .

$$x(t) : \mathbb{R} \mapsto \mathbb{R} \quad (1.1)$$

- **Sources semi-discrètes** : ce sont des sources dont la sortie est analogique et dont la valeur n'est spécifiée qu'à des instants discrets. Ce sont, par exemple, les sources échantillonnées mais qui ne sont pas encore quantifiées.
2. **Echantillonnage et quantification** : c'est l'opération de discrétisation d'une source analogique d'information. Après échantillonnage d'une source analogique on obtient une source semi-discrète. La quantification consiste alors à associer à chacun des échantillons un symbole, extrait d'un certain alphabet fini, suivant une certaine loi.
 3. **Codage de source** : c'est la représentation binaire des différents niveaux de quantification. Le choix du code binaire à associer à chacun des symboles se base sur la théorie de l'information. On retrouve ici principalement les codes à longueur fixe et les codes à longueur variable.
 4. **Codage de canal** : le codage de canal est une opération de transcodage. A chacun des codes de source on associe un autre code qui est généralement plus long que celui-ci. Le but de cette opération est la protection de l'information contre les erreurs de transmission, en introduisant dans le code source une certaine redondance contrôlée. La séquence des codes de canal forme un signal en bande de base qui est transmissible directement ou après modulation. Notons que le

codage de canal sert aussi à adapter le message numérique à transmettre aux caractéristiques du canal de transmission.

- Canal de transmission : c'est la liaison câblée ou Hertzienne qui relie l'émetteur au récepteur. Le modèle le plus commun pour modéliser le canal de transmission est le canal linéaire à bruit blanc gaussien additif (AWGN : Additive White Gaussian Noise). Il est schématisé comme montré par la figure 1.2.

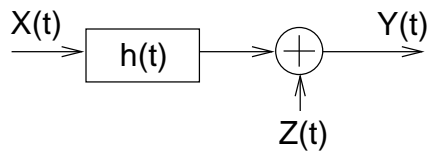


Figure 1.2 : *Modèle de canal de transmission*

Dans ce schéma $h(t)$ est la réponse impulsionnelle du filtre caractérisant le canal et $Z(t)$ est un bruit blanc additif. Nous pouvons écrire donc :

$$\begin{aligned} Y(t) &= X(t) * h(t) + Z(t) \\ &= \int_{-\infty}^{+\infty} X(t - \tau)h(\tau)d\tau + Z(t) \end{aligned} \quad (1.2)$$

où "*" représente la convolution.

- Récepteur : le récepteur effectue les opérations inverses de celle effectuées par l'émetteur. C'est-à-dire, le décodage de canal, le décodage de source et la conversion numérique-analogique, jusqu'à la restitution du message original.

2.1 Introduction

Dans le présent chapitre, nous allons nous intéresser au codage de source analogiques. Celui-ci consiste essentiellement en l'échantillonnage et la quantification. Ces principes ont déjà été étudiés dans le programme de licence. Cependant, l'étude faite en licence était assez simplifiée dans le sens où les propriétés statistiques de la source n'ont pas été prises en compte. De plus, notre étude était limitée à la quantification scalaire. Dans ce chapitre, nous introduisons la quantification optimale, d'abord dans le cas scalaire, puis nous l'étendrons à la quantification vectorielle. Les nouvelles notions introduites seront harmonieusement reliées aux connaissances déjà acquises.

2.2 Échantillonnage

Soit $u(t)$ un signal analogique. L'échantillonnage consiste à prélever des échantillons de $u(t)$ à des intervalles réguliers espacés de T secondes.

T est la période d'échantillonnage et $f = \frac{1}{T}$ est la fréquence d'échantillonnage.

Pour pouvoir formaliser cette notion d'échantillonnage, définissons une fonction $P(t)$ comme suit :

$$P(t) = \begin{cases} 1; & 0 \leq t \leq \tau \text{ avec } 0 < \tau \leq T \\ 0; & \text{ailleurs} \end{cases} \quad (2.1)$$

L'expression du signal échantillonné est le produit de $u(t)$ et du train d'impulsions formé à partir de $P(t)$

$$u(nT) = u(t) \sum_{-\infty}^{\infty} P(t - nT) \quad (2.2)$$

$u(t)$ est supposé être un signal en bande de base, à bande passante limitée, vérifiant :

$$|U(f)| \leq \frac{W}{2} \text{ où } W = \frac{1}{T} \text{ et } U(f) \leftrightarrow u(t) \quad (2.3)$$

C'est à dire que $U(f)$ est la transformée de Fourier de $u(t)$.

Echantillonnage idéal Un échantillonnage idéal est obtenu lorsque $\tau \rightarrow 0$. Dans ce cas,

$$u(nT) = \sum_{-\infty}^{\infty} u(nT)\delta(t - nT) \quad (2.4)$$

Echantillonnage réel En pratique, $\tau \neq 0$ et on effectue ainsi un échantillonnage réel. Nous essayerons de maintenir alors la largeur τ des impulsions $P(t)$ aussi petite que possible.

2.3 Quantification

$u(t)$ étant le signal analogique à coder, il est d'abord échantillonné, ce qui engendre une séquence $u(0), u(T), u(2T), \dots, u(nT), \dots$. Les échantillons prélevés sont aléatoires et la séquence $\{u(nT)\}$ peut être vue comme une séquence de variables aléatoires continues, $\{U_n\}$, chacune caractérisée par une *pdf* (fonction densité de probabilités) $f(u)$. La quantification consiste à représenter les échantillons de la séquence par un nombre fini de points représentatifs.

Nous pouvons distinguer entre la *quantification scalaire* et la *quantification vectorielle*.

2.3.1 Quantification scalaire

La quantification scalaire consiste à traiter chaque variable aléatoire U_n , $n \in \mathbb{N}$ de $\{U_n\}$ indépendamment des autres variables U_m ($m \in \mathbb{N}$ et $m \neq n$) et à lui associer une variable aléatoire discrète V_n qui prend ses valeurs dans un alphabet fini de points représentatifs $\{q_1, q_2, \dots, q_M\}$ ($M \in \mathbb{N}$), appelés niveaux de quantification.

Comme nous pouvons le constater, une telle association engendre une certaine distorsion D que nous pouvons évaluer en utilisant un certain critère, par exemple, le critère M.S.E (Mean Squared Error) qui donne l'erreur quadratique moyenne (E.Q.M). Selon ce critère, l'expression de la distorsion s'écrit :

$$D = E[(U - V)^2] \quad (2.5)$$

Dans le cas d'une quantification scalaire, et pour M niveaux de quantification, l'ensemble \mathbb{R} est partitionné en M intervalles $\{I_1, I_2, \dots, I_M\}$ et chaque intervalle I_j ($1 \leq j \leq M$) est représenté par un niveau représentatif q_j , comme illustré par la figure 2.1.

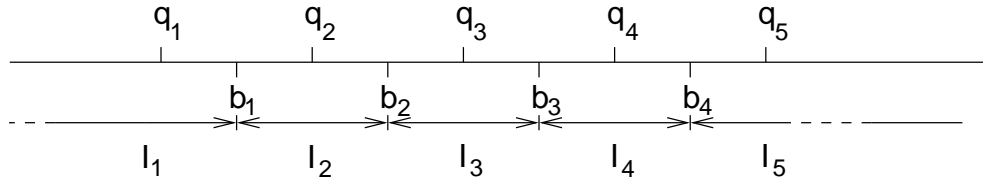


Figure 2.1 : *Quantification scalaire*

Les éléments $\{b_j\}$ délimitent les frontières entre les intervalles. Nous pouvons voir donc que :

$$Pr\{V_n = q_j\} = Pr\{U_n \in I_j\} \quad (2.6)$$

Le problème qui nous intéresse ici est comment choisir les intervalles I_j et les niveaux q_j de façon à *minimiser* la distorsion de quantification D . Deux approches sont possibles :

1. **Approche 1 :** étant donné $\{q_j\}$ comment choisir $\{I_j\}$.
2. **Approche 2 :** étant donné $\{I_j\}$ comment choisir $\{q_j\}$.

La première approche, que nous appelons ici *quantification classique*, va nous conduire aux résultats qui ont été obtenus dans le programme de licence. Nous noterons que cette approche ne tient pas compte des statistiques de la source à coder. La seconde approche, que nous appelons *quantification optimale*, tient compte des statistiques de la source. Ainsi, dans la pratique, les méthodes de quantification modernes sont la combinaison des deux approches qui est effectuée par des algorithmes de compression efficaces.

2.3.1.1 Quantification classique

Dans la première approche, considérons un échantillon quelconque $u \in \mathbb{R}$. La distorsion par rapport à q_j est $D = E[(u - q_j)^2]$. D est minimale en assignant à u le niveau q_j le plus proche. Donc les intervalles $\{I_j\}$ doivent être choisis de telle sorte à ce que les frontières entre I_j et I_{j+1} se situent au milieu du segment $q_j \longleftrightarrow q_{j+1}$. D'où

$$b_j = \frac{q_j + q_{j+1}}{2} \quad (2.7)$$

Notons que dans cette première approche, le modèle probabiliste de $\{U_n\}$ n'intervient pas. Nous pouvons aussi distinguer ici entre la quantification uniforme et la quantification non uniforme, que nous allons considérer ci-dessous.

Quantification uniforme La quantification est dite uniforme (ou linéaire) si tous les intervalles de quantification sont égaux ($b_{j+1} - b_j = b_{i+1} - b_i, \forall i, j \in \mathbb{N}$) et, par conséquent, les niveaux de quantifications équidistants. Cette technique a la désagréable caractéristique que pour les signaux à faibles amplitudes, le bruit de quantification peut submerger totalement le signal utile.

Définissons le pas de quantification comme étant :

$$\Delta = b_{j+1} - b_j \quad (2.8)$$

D'où, le niveau de quantification q_j pouvant être obtenu à partir de :

$$q_j = j\Delta + \frac{\Delta}{2} \quad (2.9)$$

L'opération classique de quantification consiste à attribuer un des niveaux représentatifs q_j à chacun des échantillons $u(nT)$ selon la relation suivante :

$$u(nT) \leftarrow q_j \quad \text{si} \quad i\Delta < u(nT) < (i+1)\Delta \quad (2.10)$$

Il est évident que la quantification introduit une erreur sur la valeur exacte de l'échantillon $u(nT)$. Il en résulte une distorsion de quantification. Évaluons cette distorsion.

$$\begin{aligned} D &= E[(U - V)^2] = \int_{-\Delta/2}^{+\Delta/2} f_{\Delta}(u)(u - v)^2 d(u - v) \\ &= \int_{-\Delta/2}^{+\Delta/2} f_{\Delta}(u)(u - q_j)^2 d(u - q_j) \end{aligned} \quad (2.11)$$

où $f_{\Delta}(u)$ est la densité de probabilité de U sur l'intervalle $[-\Delta/2, +\Delta/2]$. Si on suppose que la densité de probabilité est uniforme sur cet intervalle, nous avons $f_{\Delta}(u) = \frac{1}{\Delta}$ et

$$\begin{aligned} D &= E[(U - V)^2] = \int_{-\Delta/2}^{+\Delta/2} f_{\Delta}(u)(u - q_j)^2 d(u - q_j) \\ &= \int_{-\Delta/2}^{+\Delta/2} \left(\frac{1}{\Delta}\right) (u - q_j)^2 d(u - q_j) \\ &= \frac{\Delta^2}{12} \end{aligned} \quad (2.12)$$

L'expression de la distorsion de quantification ci-dessus montre clairement que le bruit de quantification est directement proportionnel au carré du pas de quantification. Il est, par conséquent, évident que si on veut minimiser l'influence du bruit de quantification il est nécessaire de maintenir Δ au petit que possible. Cependant, plus Δ est petit plus le nombre d'intervalles I_j sera grand, ce qui a aussi un impact sur le débit binaire à la sortie du codeur de source. Un compromis est donc à chercher.

Quantification non uniforme La solution au problème de la technique de quantification uniforme, noté ci-dessus, consiste à faire varier le pas de quantification en fonction des variations du signal à quantifier. C'est la quantification non uniforme. Nous pouvons assimiler que cette technique est équivalente à un schéma de *compression-extension* suivi d'une quantification uniforme.

Le rôle du compresseur est de changer la dynamique du signal $u(t)$ de façon à adapter les variations de $u(t)$ à un pas de quantification constant. L'extenseur, en jouant le rôle inverse de celui du compresseur, se charge de restituer la forme originale de $u(t)$. Deux loi de compression-extension sont à citer :

1. Norme Européenne (loi A). Elle est exprimée selon :

$$y = \begin{cases} \frac{Au}{1+(\ln A)}; & \text{pour } u \leq \frac{1}{A} \\ \frac{1+(\ln A)u}{1+(\ln A)}; & \text{pour } \frac{1}{A} \leq u \leq 1 \end{cases} \quad (2.13)$$

Notons que dans cette expression u est normalisé et que $A = 87,6$.

2. Norme américaine (Loi μ)

$$y = \frac{\ln(1 + \mu u)}{\ln(1 + \mu)} \quad (2.14)$$

2.3.1.2 Quantification optimale

Dans la seconde approche, les intervalles $\{I_j\}$ sont donnés et nous souhaitons définir $\{q_j\}$ de façon à minimiser la distorsion D . Dans ce cas le modèle probabiliste de $\{U_n\}$ est important. Évaluons la distorsion D .

$$\begin{aligned} D &= E[(U - V)^2] = \int_{-\infty}^{\infty} f(u)(u - v)^2 du \\ &= \sum_{j=1}^M \int_{I_j} f(u)(u - q_j)^2 du \\ &= \sum_{j=1}^M \int_{I_j} Pr\{u \in I_j\} f(u|u \in I_j)(u - q_j)^2 du \end{aligned} \quad (2.15)$$

Notons que :

$$f(u|u \in I_j) = \begin{cases} \frac{f(u)}{Pr\{u \in I_j\}}; & u \in I_j \\ 0; & \text{sinon} \end{cases} \quad (2.16)$$

D est une somme de termes qui sont tous positifs, donc D est minimisé en minimisant chaque terme individuellement. Nous pouvons montrer que chaque terme est minimisé si

$$q_j = E[U'] \text{ telle que } f(u') = \frac{f(u)}{Pr\{u \in I_j\}} \quad (2.17)$$

Remarque 2 *Montrer ce résultat*

On conclue donc pour les conditions nécessaires de l'optimalité (dans le sens de la distorsion minimale)

$$\begin{aligned} - b_j &= \frac{q_j + q_{j+1}}{2} \\ - q_j &= E[U'] \text{ telle que } f(u') = \frac{f(u)}{Pr\{u \in I_j\}} \end{aligned}$$

2.3.2 Quantification vectorielle

La quantification vectorielle consiste à considérer conjointement k variables aléatoires de la séquence $\{U_n\}$. Elle offre un gain considérable lorsque ces variables sont statistiquement dépendantes. Par exemple, dans le cas de deux variables U et U' , nous pouvons considérer M points représentatifs $\{(q_j, q'_j), (1 \leq j \leq M)$ dans le plan, correspondants à M régions \mathcal{R}_j . Ici, nous devons prendre en considération la *pdf* conjointe $f(u, u')$ et la procédure utilisée avec la quantification scalaire peut essentiellement être étendue à la quantification vectorielle pour définir les M régions du plan requises et les M points représentatifs (q_j, q'_j) correspondants.

Notions de théorie de l'information

3.1 Introduction

Dans le chapitre précédent, nous avons vu que la source analogique d'information est transformée, par échantillonnage puis quantification, en une source discrète d'information. On avait noté alors les M niveaux de quantification par q_1, q_2, \dots, q_M .

Appelons $Q = \{q_1, q_2, \dots, q_M\}$ l'ensemble des niveaux de quantification possibles. On peut décrire la source discrète par la séquence $\{U_n\}$ où U_n est une variable prenant ses valeurs dans Q , ce qui correspond à la valeur de l'échantillon $u(nT)$ à l'instant d'échantillonnage nT . Comme on n'a aucune connaissance, à priori, sur la valeur de l'échantillon $u(nT)$, $\{U_n\}$ est donc une séquence de variables aléatoires. A chaque instant nT , la variable aléatoire U_n peut prendre n'importe quelle valeur q_i ($i = 1, 2, \dots, M$). Soit $p_i = Pr(U_n = q_i)$ la probabilité que la variable aléatoire U_n prenne q_i à l'instant nT . La fréquence d'apparition des symboles U_n est, bien sur, égale à la fréquence d'échantillonnage, $f = \frac{1}{T}$.

Dans ce qui suit, nous allons d'abord donner quelques brefs rappels de probabilités et de variables aléatoires puis nous définissons les notions de quantité d'information d'un symbole q_i et de quantité d'information moyenne, ou entropie, d'une source d'information.

3.2 Rappels sur les probabilités

Soit \mathcal{X} un sous-ensemble de nombres réels et X une variable aléatoire prenant ses valeurs dans \mathcal{X} .

Si \mathcal{X} est fini, alors X est discrète et elle peut être décrite par sa répartition $p(x) = Pr(X = x)$ pour tout x dans \mathcal{X} .

Si \mathcal{X} est un intervalle, alors X est continue, et elle est caractérisée par sa distribution

$Pr(X \leq x_0)$ pour tout x_0 ou par sa densité de probabilité $f(x) = \frac{d}{dx}Pr(X \leq x_0)$.

Remarque 3 Une variable X est déterministe si $p(x) = 1$ pour un certain $X = x_0$, ce qui implique que pour tout $x_1 \neq x_0$, $p(X = x_1) = 0$.

Remarque 4 Une variable discrète X est équiprobable si $p(x) = 1/M$ pour tout $x \in \mathcal{X}$ où M est l'ordre de \mathcal{X} . Une variable équiprobable est aussi dite uniforme.

3.2.1 Moyenne statistique et moments

L'espérance mathématique, ou moyenne statistique de la variable aléatoire X est définie par

$$E[X] = \sum_{x \in \mathcal{X}} x \cdot p(x) \quad (\text{variable discrète}). \quad (3.1)$$

Étant donnés deux sous-ensembles \mathcal{X} et \mathcal{Y} discrets (qu'on peut appeler alphabets pour la circonstance), le produit cartésien $\mathcal{X} \times \mathcal{Y}$ est aussi un alphabet discret.

$$\mathcal{X} \times \mathcal{Y} = \{(x, y) \mid x \in \mathcal{X} \text{ et } y \in \mathcal{Y}\} \quad (3.2)$$

Si X et Y sont deux variables aléatoires prenant leur valeurs dans \mathcal{X} et \mathcal{Y} respectivement, alors les variables $X + Y$ et XY sont aussi des variables aléatoires et elles prennent leurs valeurs dans le produit cartésien $\mathcal{X} \times \mathcal{Y}$. Elles sont caractérisées par leurs probabilité conjointe notée $p(x, y)$.

La probabilité de X est reliée à la probabilité conjointe par :

$$p(x) = \sum_{y \in \mathcal{Y}} p(x, y) \quad (3.3)$$

et nous avons aussi

$$p(y) = \sum_{x \in \mathcal{X}} p(x, y) \quad (3.4)$$

La probabilité conditionnelle de Y est définie par :

$$p(y|x) = \frac{p(x, y)}{p(x)} \quad (3.5)$$

Et nous avons aussi

$$p(x|y) = \frac{p(x, y)}{p(y)} \quad (3.6)$$

ou

$$p(x, y) = p(y|x)p(x) = p(x|y)p(y) \quad (3.7)$$

Deux variables X et Y sont indépendantes si

$$p(x, y) = p(x)p(y) \quad (3.8)$$

Il s'ensuit que

$$p(y|x) = p(y) \text{ et } p(x|y) = p(x) \quad (3.9)$$

L'espérance mathématique du produit de deux variables aléatoires X et Y est égale au produit des espérances lorsque X et Y sont indépendantes

$$E[XY] = E[X]E[Y] \text{ pour } X \text{ et } Y \text{ indépendantes} \quad (3.10)$$

L'espérance mathématique de la somme de deux variables aléatoires X et Y est égale à la somme des espérances de chacune d'elles. Nous écrivons

$$E[X + Y] = E[X] + E[Y] \quad (3.11)$$

Remarque 5 Montrer les résultats des équations 3.10 et 3.11

Le n^{eme} moment de la variable X est défini par :

$$E[X^n]. \quad (3.12)$$

Le premier moment est la moyenne statistique qu'on peut noter par $\bar{X} = E[X]$. L'autocorrelation est obtenue par le second moment et la variance de X est définie par :

$$\sigma_X^2 = E[(X - \bar{X})^2] \quad (3.13)$$

Notons que

$$\sigma_X^2 = E[X^2] - \bar{X}^2 \quad (3.14)$$

σ_X est appelée la *déviatation standard*.

3.3 Introduction à la théorie de l'information

3.3.1 Quantité d'information d'un symbole

Considérons l'ensemble Q de symboles que nous avons déjà introduit. Plutôt que d'essayer de définir directement la quantité d'information que nous apporte l'apparition d'un symbole $q_i, i \in [1, M]$ ($X_n = q_i$), essayons de la caractériser. Intuitivement, la quantité d'information H_i d'un symbole q_i doit avoir les propriétés suivantes :

1. Si on est certain que q_i apparaît à un instant n quelconque, alors le fait que q_i apparaisse à cet instant n ne nous apporte aucune information ($H_i = 0$).
2. Moins on est certain de l'apparition de q_i , plus grande est la quantité d'information que nous apporte son apparition.
3. Si l'apparition de q_i est indépendante de l'apparition de q_j ($i, j \in [1, M]$), aux instants n et k respectivement, alors l'apparition conjointe de q_i et q_j nous apporte une quantité d'information égale à la somme des quantités d'information de q_i et q_j . (nous appelons une telle source une *source discrète sans mémoire* (SDSM)).

Ces hypothèses nous amènent à définir H_i comme

$$H_i = \log_2 \left(\frac{1}{p_i} \right) = -\log_2 p_i \quad (3.15)$$

En effet, H_i possède les propriétés suivantes :

- $H_i = 0$ si $p_i = 1$; $H_i \geq 0$
- $H_i > H_j$ si $p_i < p_j$
- $H_{ij} = -\log_2 p_{ij} = -\log_2 p_i p_j = -\log_2 p_i - \log_2 p_j = H_i + H_j$

3.3.2 Quantité d'information moyenne (Entropie)

Une meilleure appréciation de la quantité d'information d'une source discrète est sa quantité d'information moyenne ou *entropie*. Pour une source X l'entropie est définie par :

$$H(X) = \sum_{i=1}^M p_i H_i = -\sum_{i=1}^M p_i \log_2 p_i \quad (3.16)$$

Il est intéressant de noter que l'entropie est maximale lorsque les M symboles sont équiprobables. On a alors les relations suivantes :

$$p_i = \frac{1}{M}$$

et

$$H(X)_{max} = \sum_{i=1}^M \left(\frac{1}{M} \right) \log_2 M = \log_2 M$$

En effet, pour voir ce résultat il suffit de remarquer que :

$$H(X) - \log_2 M = \left(\sum_{i=1}^M p_i \log_2 \frac{1}{p_i} \right) - \log_2 M \quad (3.17)$$

$$= \sum_{i=1}^M p_i \left(\log_2 \left(\frac{1}{p_i} \right) - \log_2 M \right) \quad (3.18)$$

$$= \sum_{i=1}^M p_i \log_2 \left(\frac{1}{M p_i} \right) \quad (3.19)$$

$$\leq (\log_2 e) \sum_{i=1}^M p_i \left(\frac{1}{M p_i} - 1 \right) = 0 \quad (3.20)$$

où, dans la dernière équation nous avons utilisé la propriété du logarithme $\ln x \leq x - 1$.

Donc $H(X) - \log_2 M \leq 0$ et $H(X) - \log_2 M = 0$ lorsque $\log_2 \left(\frac{1}{Mp_i} \right) = 0$ d'où $p_i = \frac{1}{M}$ pour $i = 1, 2, \dots, M$. C'est donc dans le cas où les symboles sont équiprobables que l'entropie de la source est maximale.

Remarque 6 La quantité d'information s'exprime en bit/symbole.

Débit d'information Le débit d'information d'une source discrète de symboles est définie par :

$$H'(X) = H(X).f \quad (3.21)$$

L'unité du débit d'information est donc : $(\text{bit/symbole}) \times (\text{symbole/sec}) = \text{bit/sec}$.

Quantité de décision La quantité de décision est une mesure du nombre de bits qui sont nécessaires pour coder en binaire les M niveaux distincts de quantification. Elle est exprimée par :

$$D = \log_2 M \quad (3.22)$$

Redondance La redondance d'une source est la différence entre la quantité d'information maximale de la source et sa quantité d'information moyenne.

$$R = H(X)_{max} - H(X) \quad (3.23)$$

Débit de décisions Le débit de décisions est donné par :

$$D' = D.f \quad (3.24)$$

L'unité du débit de décisions est donc le bit/sec .

4.1 Introduction

Le codage de source consiste à représenter chacun des symboles $q_i \in Q$ par une série de nombres binaires.

Exemple 1 Soit Q un ensemble composé de 7 lettres ; $Q = \{a, b, c, d, e, f, g\}$. Un exemple de code source pour Q est l'association suivante :

$a \mapsto 000; b \mapsto 001; c \mapsto 010; d \mapsto 011; e \mapsto 100; f \mapsto 101; g \mapsto 110.$

On distingue deux types de codes : les codes à longueurs fixes et les codes à longueurs variables.

4.2 Codes a longueur fixe

Les codes à longueurs fixes sont les codes dans lesquels chacun des codes binaires associé à chaque symbole dans Q a une longueur constante.

Exemple 2 Le code utilisé dans l'exemple ci-dessus est un exemple de code à longueur fixe, de longueur égale à 3 (3 bits/symbole).

D'une manière générale, si on a M symboles dans l'ensemble Q alors la longueur minimale L_{min} d'un code à longueur fixe pour coder les M symboles dans Q est :

$$L_{min} = \lceil \log_2 M \rceil \tag{4.1}$$

ce qui désigne ici le plus petite entier tel que $\log_2 M \leq L_{min}$.

Exemple 3 $\log_2 6 = 2,5849625$ donc $\lceil \log_2 6 \rceil = 3.$

On a donc la relation

$$L_{min} = \bar{L} \geq \log_2 M. \quad (4.2)$$

où \bar{L} désigne la longueur moyenne.

Supposons maintenant que l'on veuille coder des blocs de n symboles à la fois. C'est-à-dire, si la sortie de la source discrète d'information est une séquence $\{U_k\}$, $k = 0, 1, 2, \dots$ alors on se propose de coder des blocs de taille n , $[u_i, \dots, u_{i+n}]$ de symboles où $u_k \in Q$ et i est un entier quelconque.

On peut regarder un bloc de n symboles comme étant la valeur d'une variable qui prend ses valeurs dans le produit cartésien $Q \times Q \times \dots \times Q = Q^n$. On peut alors distinguer M^n valeurs possibles pour la variable et donc la longueur nécessaire pour un code à longueurs fixes vérifie les relations suivantes :

$$\begin{aligned} L_n &= \lceil \log_2 M^n \rceil \\ \bar{L} &= \frac{\lceil \log_2 M^n \rceil}{n} \geq \frac{n \log_2 M}{n} = \log_2 M \\ \bar{L} &= \frac{\lceil \log_2 M^n \rceil}{n} < \frac{n \log_2 M + 1}{n} = \log_2 M + \frac{1}{n} \end{aligned}$$

où \bar{L} désigne la longueur moyenne par symbole. Donc si n est grand, la longueur moyenne est proche de la longueur nécessaire pour coder chaque symbole séparément.

Remarque 7 La méthode de codage par codes à longueur fixe ne présente aucun problème de synchronisation (?). Elle est non probabiliste. Par conséquent elle n'offre aucune possibilité de compression de données.

4.3 Codes à longueur variable

Cette méthode de codage permet la compression de donnée dans le sens où on peut associer les codes les plus courts aux symboles les plus probables et les codes les plus longs aux symboles les moins probables.

Un code à longueur variable associe à chaque symbole q ($q \in Q$) une série de bits $C(q)$ de symboles binaires appelés mot de code de longueur $l(q)$.

Exemple 4 Soit l'alphabet $Q = \{a, b, c\}$. Un exemple de code à longueur variable est :

$$C(a) = 0$$

$$C(b) = 10$$

$$C(c) = 11$$

Propriété de décodabilité unique La propriété de décodabilité unique est une propriété essentielle qu'un code à longueur variable doit avoir. Cette propriété consiste en le fait qu'étant donnée une séquence de symboles source et une synchronisation initiale, cette séquence doit pouvoir être reconstruite sans aucune ambiguïté à partir de la série binaire qui lui est associée par le codeur. On peut énoncer la définition suivante :

Définition 1 Un code C pour une source discrète est décodable de manière unique si, pour toute séquence de symboles u_1, u_2, \dots, u_n , la concaténation de ses mots de codes $C(u_1)C(u_2) \dots C(u_n)$ est différente de la concaténation des mots de codes $C(u'_1)C(u'_2) \dots C(u'_m)$ de toute autre séquence de symboles u'_1, u'_2, \dots, u'_m .

Exemple 5 Etudier la propriété de décodabilité unique pour le code C défini ci-dessus et le codes C' défini par :

$$C'(a) = 0$$

$$C'(b) = 1$$

$$C'(c) = 01$$

4.3.1 Codes sans préfixe

Une classe importante de codes qui jouissent de la propriété de décodabilité unique sont les codes sans préfixe.

Définition 2 Un préfixe pour une séquence y_1, y_2, \dots, y_n est toute sous-séquence y_1, y_2, \dots, y_m ($m < n$) de la séquence donnée.

Un code est sans préfixe si aucun mot de code n'est un préfixe pour aucun autre mot de code.

Exemple 6 Le code C donné plus haut est sans préfixe, alors que le code C' défini ci-dessous n'est pas sans préfixe.

$$Q = \{a, b, c\}$$

$$C'(a) = 0$$

$$C'(b) = 1$$

$$C'(c) = 01$$

$C'(a)$ est un préfixe pour $C'(c)$.

Remarque 8 Tout code à longueur fixe est un code sans préfixe.

4.3.2 Construction de codes sans préfixe

On utilise un arbre de décomposition pour la construction d'un code sans préfixe. Soit $Q = \{a, b, c\}$. Pour construire un code sans préfixe, on ne place jamais les symboles aux noeuds intermédiaires, comme illustré par la figure 4.1.

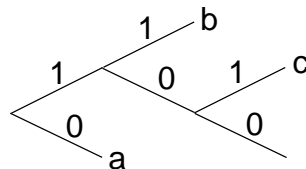


Figure 4.1 : Arbre de décomposition pour la construction de codes sans préfixe

Définition 3 Un code sans préfixe est dit plein si tous les noeuds terminaux correspondent à des mots des codes. Sinon, le code n'est pas plein.

Exemple 7 Le code de la figure 4.1 n'est pas plein car nous pouvons le réduire à celui qui est montré par la figure 4.2.

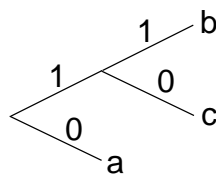


Figure 4.2 : Réduction de l'arbre de la figure 4.1

Utilisons le code \mathcal{C} et supposons que l'on veuille décoder la séquence 1010011... de gauche à droite. 1 n'est pas un mot de code ; le décodeur procède à donc à droite pour décoder 10 comme étant le symbole c . La séquence restante est 10011.... On décode encore c , puis a , puis b ... etc... La séquence décodée sera donc $ccab$...

Remarque 9 La synchronisation est d'une importance primordiale pour un code à longueur variable. Le choix du code doit donc être fait minutieusement.

Exemple 8 Le code $\{0, 10, 110, 11110\}$ est auto-synchronisable (pouvez-vous voir pourquoi?).

4.3.3 Inégalité de Kraft pour les codes sans préfixe

Un important théorème, dû à Kraft, pour les codes sans préfixe est le suivant :

Théorème 1 *pour tout code sans préfixe, pour un alphabet Q , ayant les longueurs $l(q)$, $q \in Q$, nous avons le résultat suivant :*

$$\sum_{q \in Q} 2^{-l(q)} \leq 1 \quad (4.3)$$

Réciproquement, si cette relation est satisfaite alors un code sans préfixe existe avec les longueurs $l(q)$. De plus, pour un code plein, nous avons :

$$\sum_{q \in Q} 2^{-l(q)} = 1 \quad (4.4)$$

et pour un code non-plein

$$\sum_{q \in Q} 2^{-l(q)} < 1 \quad (4.5)$$

Exemple 9 *Existe-t-il des codes ayant des longueurs 1, 2, 2 et 1, 1, 2 ?*

Revenons au modèle où l'alphabet Q est composé de M symboles q_i chacun caractérisé par sa probabilité d'apparition p_i . Supposons que l'on choisisse un ensemble de longueurs $l(q)$ satisfaisant l'inégalité de Kraft. Nous pouvons voir donc $l(q)$ comme une variable aléatoire. Si \bar{L} désigne la longueur moyenne par symbole, nous aurons :

$$\bar{L} = E[l(q) = l_i] = \sum_{i=1}^M p_i l_i \quad (4.6)$$

Notre objectif est de réduire le plus possible la longueur moyenne \bar{L} .

Théorème 2 *La longueur moyenne minimale \bar{L}_{min} pour un code sans préfixe satisfaisant l'inégalité de Kraft satisfait à la relation suivante :*

$$H \leq \bar{L}_{min} < H + 1 \quad (4.7)$$

où H est l'entropie de la source.

Si les probabilités p_i sont des puissances entières de 2 alors la relation devient

$$\bar{L}_{min} = H \quad (4.8)$$

Démonstration

$$H(X) - \bar{L} = \sum_{i=1}^M p_i \log_2 \left(\frac{1}{p_i} \right) - \sum_{i=1}^M p_i l_i \quad (4.9)$$

$$= \sum_{i=1}^M p_i \log_2 \left(\frac{1}{p_i} \right) + \sum_{i=1}^M p_i (-l_i) \quad (4.10)$$

$$= \sum_{i=1}^M p_i \log_2 \left(\frac{2^{-l_i}}{p_i} \right) \quad (4.11)$$

$$H(X) - \bar{L} \leq (\log_2 e) \sum_{i=1}^M p_i \left(\frac{2^{-l_i}}{p_i} - 1 \right) = (\log_2 e) \left(\sum_{i=0}^{N-1} 2^{-l_i} - \sum_{i=0}^{M-1} p_i \right) \leq 0 \quad (4.12)$$

Si $\frac{2^{-l_i}}{p_i} = 1$ ou $l_i = -\log_2 p_i$ alors $H(X) - \bar{L} = 0$ ce qui veut dire que $H(X) = \bar{L}_{min}$.

Pour montrer l'autre inégalité il suffit de remarquer que si l'on choisit $l_i = \lceil -\log_2 p_i \rceil$ alors nous avons :

$$\bar{L} = \sum_i p_i l_i < p_i (-\log_2 p_i + 1) \quad (4.13)$$

$$= \sum -p_i \log_2 p_i + \sum_i p_i \quad (4.14)$$

$$= H(X) + 1. \quad (4.15)$$

Le résultat du théorème est ainsi établi.

4.3.4 Algorithme de Huffman pour le codage entropique

Dans cette section, nous allons illustrer la méthode de Huffman pour construire des codes entropiques sans préfixe. L'exemple suivant, et qui est montré par la figure 4.3, illustre bien la méthode.

La technique consiste à maximiser le chemin à partir des plus petits probables jusqu'à la racine de l'arbre. Les probabilités intermédiaires sont, bien entendu, obtenues en additionnant les probabilités individuelles des arborescences correspondantes.

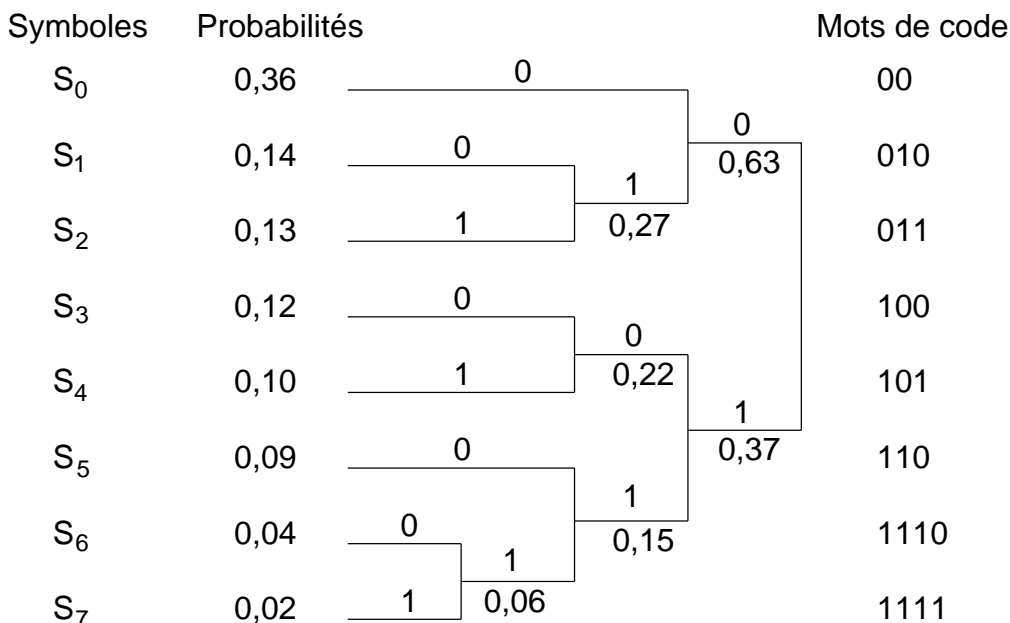


Figure 4.3 : Algorithme de Huffman

4.4 Codage de sources avec mémoire

Dans cette section nous étudions les sources avec mémoire, c’est à dire les sources pour lesquelles les symboles successifs émis sont statistiquement dépendants. Nous considérons essentiellement les sources qui sont définies en termes de chaîne (processus) de Markov. Pour représenter la mémoire de la source nous utilisons les états d’un processus de Markov et les transitions entre les états sont utilisées pour représenter la sortie de la source.

4.4.1 Sources de Markov

Définition 4 Un processus de Markov à espace d’états fini est une séquence $E_0, E_1, \dots, E_k, \dots$ de variables aléatoires discrètes, chacune prenant ses valeurs dans un alphabet fini E . Un processus de Markov est caractérisé par une distribution initiale de probabilités $q_0(e)$, $e \in E$ pour l’état initial E_0 et, pour tout $n \geq 1$ et $(e, e') \in E \times E$ nous avons

$$\begin{aligned}
 Pr(E_n = e | E_{n-1} = e') &= Pr(E_n = e | E_{n-1} = e', \dots, E_0 = e'') \\
 &= Q(e|e')
 \end{aligned}
 \tag{4.16}$$

Remarque 10 Une transition entre les états e et e' existe si $Q(e|e') > 0$.

Définition 5 Une source de Markov est une séquence de variables aléatoires discrètes X_1, X_2, \dots , chacune prenant ses valeurs dans un alphabet commun \mathcal{X} et qui est basée sur un processus de Markov E_0, E_1, \dots comme suit :

Chaque transition ($e \rightarrow e'$) est étiquetée par un symbole de \mathcal{X} et chaque symbole de \mathcal{X} ne pouvant apparaître que sur une branche sortante de chaque état e . Il y a donc autant de branches sortantes de chaque état que de symboles dans l'alphabet \mathcal{X} .

Exemple 10 Considérons une source binaire ($X = \{0, 1\}$) dont la structure peut être schématisée comme indiqué sur la figure 4.4.

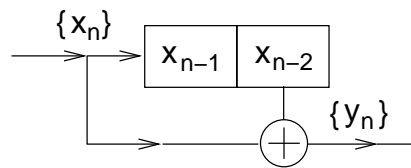


Figure 4.4 : Exemple de source de Markov

Comme nous pouvons le constater la sortie y_n dépend de y_{n-1} et de y_{n-2} . On a donc une mémoire de deux bits et nous pouvons utiliser un processus de Markov à quatre états pour modéliser la source.

Remarque 11 Tracer le diagramme d'états pour cette source.

Définition 6 1. On dit qu'un état e est accessible à partir d'un état e' dans un processus de Markov s'il existe un chemin de e vers e' , c'est à dire $\exists n > 0$ tel que $Pr(E_n = e | E_0 = e') > 0$.

2. La période d'un état e est le PGCD de l'ensemble des entiers n tels que $Pr(E_n = e | E_0 = e') > 0$.

3. Un processus de Markov est dit ergodique si tous les états sont accessibles à partir de tous les autres états et si tous les états sont apériodiques (de période 1).

Pour un processus ergodique nous pouvons retenir :

$$Pr(E_n = e | E_0 = e') = q(e) \text{ pour tous } e, e' \in E$$

$$q(e) = \sum_{e' \in E} q(e')Q(e|e'); e \in E \quad (4.17)$$

$$\sum_{e' \in E} q(e) = 1$$

4.4.2 Codage de sources de Markov

Pour tout $e \in E$, et en se basant sur la distribution des probabilités $p(x, e)$, nous pouvons choisir un code sans préfixe ayant un ensemble de longueurs $\{l(x, e)\}$ satisfaisant l'inégalité de Kraft :

$$\sum_x 2^{-l(x, e)} \leq 1 \quad (4.18)$$

Le code ainsi choisi peut être généré par codage entropique (par exemple, par l'Algorithme de Huffman) et sa longueur moyenne $\bar{L}(e)$ satisfait au théorème du codage de source

$$H(X|e) \leq \bar{L}(e) < H(X|e) + 1 \quad (4.19)$$

où $H(X|e) = \sum_{x \in \mathcal{X}} -p(x|e) \log_2 p(x|e)$.

La longueur moyenne \bar{L} du code vérifie alors la relation :

$$H(X|E) \leq \bar{L} < H(X|E) + 1 \quad (4.20)$$

où $\bar{L} = \sum_{e \in E} q(e) \bar{L}(e)$ et $H(X|E) = \sum_{e \in E} q(e) H(X|e)$.

Remarque 12 *Montrer que*

$$H(XE) = H(E) + H(X|E) \quad (4.21)$$

Nous pouvons noter que

$$H(XE) \leq H(E) + H(X) \quad (4.22)$$

D'où nous déduisons que

$$H(X|E) \leq H(X) \quad (4.23)$$

Remarque 13 *Commenter ce dernier résultat.*

5.1 Introduction

Les données numériques obtenues après codage de source (éventuellement aussi codage de canal) sont destinées à être normalement transmises à travers des canaux de transmission qui sont implémentés en utilisant des media physiques de différentes natures. La caractéristique commune à tous ces medias et qui est intéressante pour nous ici, du point de vue transmission d'information, est qu'ils sont toujours affectés par du bruit qui déforme (ou altère) les signaux transmis. Il est donc nécessaire d'étudier (modéliser) et de comprendre comment ce bruit affecte les signaux transmis. Ceci est l'objet de ce chapitre. Nous allons ainsi découvrir que le modèle le plus usuel pour caractériser un bruit additif est le bruit blanc gaussien Additif (AWGN : Additive White Gaussian Noise).

5.2 Bruit additif

Le schéma le plus simple pour décrire un canal de transmission avec bruit additif est comme montré par la figure 5.1, où $X(t)$ est le signal d'entrée, $Y(t)$ le signal de sortie et $Z(t)$ est un bruit additif. On écrit dans ce cas :

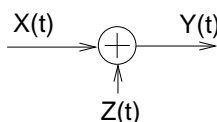


Figure 5.1 : Modèle de canal avec bruit additif

$$Y(t) = X(t) + Z(t) \tag{5.1}$$

Le plus souvent, cependant, le canal de transmission ayant une bande passante limitée se comporte aussi comme un filtre ayant une réponse impulsionnelle $h(t)$. Un schéma plus précis pour un tel canal de transmission serait donc comme montré sur la figure 5.2 et on écrit dans ce cas

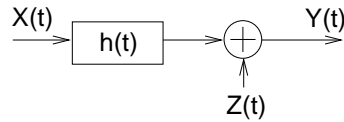


Figure 5.2 : *Modèle de canal de transmission à bande passante limitée*

$$Y(t) = X(t) * h(t) + Z(t) \quad (5.2)$$

ou

$$Y(t) = \int_{-\infty}^{\infty} X(t - \tau)h(\tau)d\tau + Z(t) \quad (5.3)$$

Dans la suite, nous nous intéresserons à la modélisation du bruit $Z(t)$.

5.3 Modélisation du bruit

Cette section introduira la notion de processus stochastique et en particulier celle de processus stochastique stationnaire. La modèle le plus usuel et le plus convenable pour caractériser le bruit dans les systèmes de communication est celui dont la distribution de probabilités est conjointement gaussienne.

5.3.1 Processus stochastique

Définition 7 *Un processus stochastique $\{N(t); -\infty < t < \infty\}$ est une collection de variables aléatoires, chacune définie pour chaque valeur de t . C'est-à-dire, pour tout $t_0 \in \mathbb{R}$, $N(t_0)$ est une variable aléatoire.*

Un processus stochastique est complètement défini si on connaît sa distribution de probabilités. Pour tout entier $k > 0$ et pour tous les instants t_1, t_2, \dots, t_k il doit être possible de trouver la fonction de distribution conjointe

$$F(n_1, n_2, \dots, n_k) = Pr\{N(t_1) \leq n_1, N(t_2) \leq n_2, \dots, N(t_k) \leq n_k\} \quad (5.4)$$

ou, de manière équivalente, la densité de probabilité conjointe, si elle existe

$$f(n_1, n_2, \dots, n_k) = \frac{\partial^k F(n_1, n_2, \dots, n_k)}{\partial n_1 \partial n_2 \dots \partial n_k} \quad (5.5)$$

Pour un processus gaussien (que nous allons définir ci-après) il est possible de chercher sa distribution de probabilité. Ceci se fait en utilisant la distribution de probabilité d'un processus gaussien normalisé. En effet, nous allons voir que la densité de probabilité d'un processus gaussien ne dépend que de la fonction de covariance que nous définissons dans la suite.

Etant donné donc un processus stochastique $\{N(t)\}$, la moyenne statistique du processus à l'instant t_0 est définie par

$$N(t_0) = E[N(t_0)] \quad (5.6)$$

La covariance aux instants t et τ est donnée par

$$K_N(t, \tau) = E\{[N(t) - \bar{N}(t)][N(\tau) - \bar{N}(\tau)]\} \quad (5.7)$$

Si le processus a une moyenne nulle à tout instant t alors la covariance se simplifie à

$$K_N(t, \tau) = E\{N(t)N(\tau)\} \quad (5.8)$$

5.3.1.1 Processus gaussien

Définissons d'abord une variable aléatoire gaussienne normalisée. Une telle variable est une variable aléatoire dont la distribution est donnée par :

$$f(n) = \frac{1}{\sqrt{2\pi}} e^{-\frac{n^2}{2}} \quad (5.9)$$

L'allure de cette fonction est représentée par la figure 5.3. Par symétrie, on voit bien que la moyenne est nulle. La variance est égale à 1 et ceci peut être retrouvé en utilisant une table d'intégrales.

En utilisant une variable aléatoire gaussienne *normalisée*, nous pouvons maintenant définir une variable aléatoire gaussienne en général. Une variable aléatoire gaussienne Z de moyenne \bar{Z} et de variance σ^2 , est donnée en fonction d'une variable normalisée N par :

$$Z - \bar{Z} = \sigma N \quad (5.10)$$

Et sa densité de probabilité est donnée par :

$$f_Z(z) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(z-\bar{Z})^2}{2\sigma^2}} \quad (5.11)$$

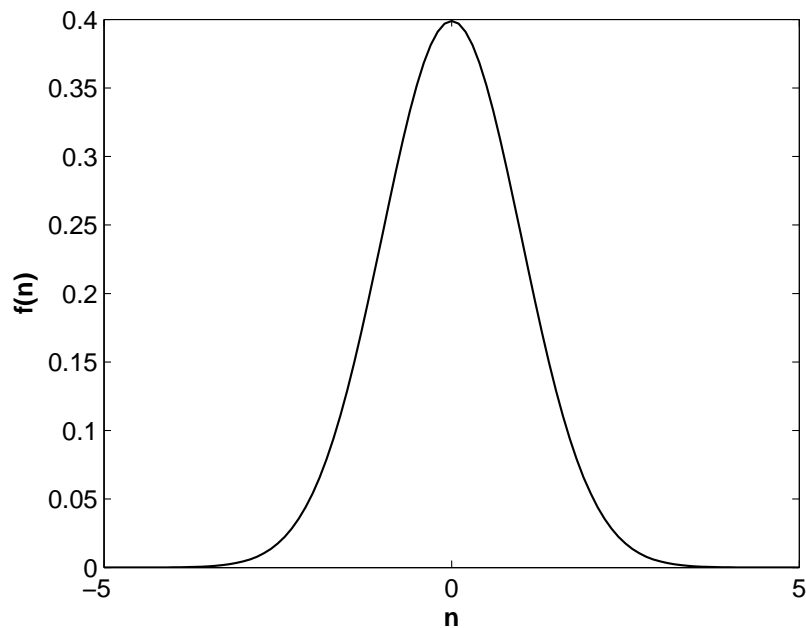


Figure 5.3 : Distribution gaussienne ou normale

L'importance de la distribution gaussienne est justifiée par le théorème central aux limites de la théorie des probabilités (?).

Définition 8 Un ensemble de k variables aléatoires Z_1, Z_2, \dots, Z_k est un ensemble de variables conjointement gaussiennes si chacune des variables Z_l peut être écrite sous forme d'une combinaison linéaire de variables gaussiennes normalisées.

$$Z_l = \sum_{j=1}^m a_{lj} N_j \quad (5.12)$$

où $a_{lj} \in \mathbb{R}$ et N_j sont des variables gaussiennes normalisées.

De manière plus compacte, on peut écrire :

$$\mathbf{Z} = \mathbf{A}\mathbf{N} \quad (5.13)$$

Avec :

$$\mathbf{Z} = [Z_1, Z_2, \dots, Z_k]^T$$

$$\mathbf{N} = [N_1, N_2, \dots, N_m]^T$$

et

$\mathbf{A} = [a_{lj}]$ une matrice de coefficients de dimension $k \times m$.

Définition 9 Un processus stochastique est un processus gaussien si pour tout $k > 0$ et tous les instants t_1, t_2, \dots, t_k l'ensemble des variables t_1, t_2, \dots, t_k est un ensemble de variables aléatoires conjointement gaussiennes.

5.3.1.2 Processus stochastique stationnaire

Définition 10 Un processus stochastique est stationnaire si et seulement si $E[N(t)] = E[N(0)]$ et $K_N(t, t + \tau) = K_N(0, \tau)$ pour tous $t, \tau \in \mathbb{R}$.

5.3.1.3 Densité de probabilité d'un ensemble de variables conjointement gaussiennes

Prenons un ensemble de variables aléatoires Z_1, Z_2, \dots, Z_k conjointement gaussiennes de moyennes nulles. On peut écrire (comme déjà présenté)

$$\mathbf{Z} = \mathbf{A}\mathbf{N} \quad (5.14)$$

La covariance entre chaque paire de variable aléatoire Z_i et Z_j peut être écrite sous forme d'une matrice de covariance.

$$\mathbf{K}_Z = E[\mathbf{Z}\mathbf{Z}^T] \quad (5.15)$$

Si les m variables aléatoires N_1, N_2, \dots, N_m sont indépendantes alors on peut supposer qu'il n'y a aucune corrélation entre elles. Donc :

$$E[N_l][N_j] = 0 \text{ si } l \neq j \quad (5.16)$$

et

$$E[N_l][N_j] = 1 \text{ si } l = j \quad (5.17)$$

Par conséquent :

$$K_N = E[\mathbf{N}\mathbf{N}^T] = \mathbf{I}_m \quad (5.18)$$

et

$$\mathbf{K}_Z = E[\mathbf{A}\mathbf{N}\mathbf{N}^T\mathbf{A}^T] = \mathbf{A}E[\mathbf{N}\mathbf{N}^T]\mathbf{A}^T = \mathbf{A}\mathbf{A}^T \quad (5.19)$$

Le calcul de la densité de probabilités conjointe des variables N_1, N_2, \dots, N_m est direct puisque celles-ci sont indépendantes.

$$f(\mathbf{n}) = \frac{1}{(2\pi)^{\frac{m}{2}}} e^{-\frac{n_1^2 + n_2^2 + \dots + n_m^2}{2}} \quad (5.20)$$

Remarque 14 *Pouvez-vous voir ce résultat ?*

En utilisant cette fonction, on peut montrer que la densité de probabilités conjointe des variables Z_1, Z_2, \dots, Z_k peut être obtenue comme :

$$f(\mathbf{z}) = \frac{1}{(2\pi)^{\frac{k}{2}} \sqrt{\det(\mathbf{K}_Z)}} e^{-\frac{1}{2} \mathbf{z}^T \mathbf{K}_Z^{-1} \mathbf{z}} \quad (5.21)$$

5.3.1.4 Bruit blanc gaussien

Un bruit blanc gaussien est un processus stochastique dont les variables aléatoires à des instants différents t_1 et t_2 sont complètement décorrélées.

$$E[N(t)N(t + \tau)] = \frac{N_0}{2} \delta(\tau) \quad (5.22)$$

où $\delta(\tau)$ est la fonction de Dirac.

La densité spectral de puissance est la transformée de Fourier de la fonction de covariance (auto-corrélation). Elle est donc donnée par :

$$S(f) = \frac{N_0}{2} \quad (5.23)$$

Pour un bruit blanc gaussien.

5.4 Rapport signal sur bruit

Le rapport signal sur bruit (noté S/N ou SNR) est une mesure importante qui apparaît dans la plupart des calculs en télécommunications. Elle compare la puissance du signal reçu à celle du bruit.

On suppose que le signal transmis est une séquence de symboles. Soit T la durée d'un symbole. La bande passante nominale du signal est définie comme étant :

$$W = \frac{1}{T} \quad (5.24)$$

Dans la bande passante nominale, la puissance du bruit sera égale à $(\frac{N_0}{2} \times W) \times 2 = N_0 W$, où $\frac{N_0}{2}$ est la densité spectrale de puissance bilatérale du bruit. Si P est la puissance du signal, alors le rapport signal sur bruit est :

$$S/N = SNR = \frac{P}{N_0 W} \quad (5.25)$$

5.5 Capacité d'un canal de transmission

Pour un canal AWGN, Shannon a montré un théorème fondamental relatif à la capacité, C , d'un canal de transmission. Celle-ci est définie comme étant le nombre maximum de bits par seconde (ou de symboles par seconde) qu'on peut transmettre en garantissant une probabilité d'erreur arbitrairement petite. La capacité est donnée par :

$$C = W \log_2 \left(1 + \frac{P}{N_0 W} \right) \quad (5.26)$$

En d'autres termes, le théorème de Shannon nous assure que pour un débit source R ,

- si $R \leq C$ alors la probabilité d'erreur peut être arbitrairement petite.
- si $R > C$ alors la probabilité d'erreur est incontrôlable.

Examinons l'allure de la capacité C en fonction de W . Celle-ci est représentée par la figure 5.4.

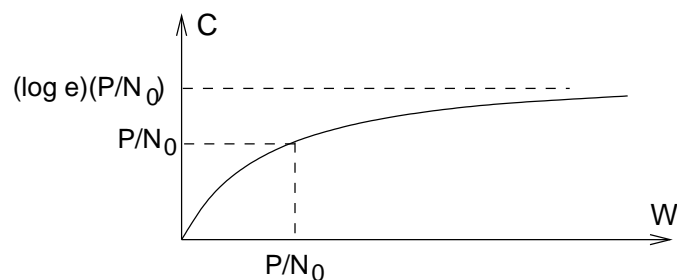


Figure 5.4 : Allure de la capacité du canal en fonction de la bande passante W .

On peut réécrire :

$$C = W \left(\frac{1}{\ln 2} \right) \ln \left(1 + \frac{P}{N_0 W} \right) \quad (5.27)$$

$$= \left(\frac{P}{N_0} \log_2 e \right) \left(\frac{N_0 W}{P} \right) \ln \left(1 + \frac{P}{N_0 W} \right) \quad (5.28)$$

Remarque 15 Commenter cette courbe et faire les observations relatives à la bande passante nominale W .

5.5.1 Efficacité spectrale

Revenons au théorème de Shannon. Ainsi, pour garantir une probabilité d'erreur aussi petite que l'on désire, le débit R qu'on doit utiliser doit être de telle sorte que :

$$C \leq W \log_2 \left(1 + \frac{P}{N_0 W} \right) \quad (5.29)$$

On définit l'efficacité spectrale par

$$\rho = \frac{R}{W} \quad (5.30)$$

L'efficacité spectrale est donc le nombre de bits par seconde qu'on peut transmettre dans chaque Hertz de bande passante. On déduit :

$$\rho \leq \log_2 \left(1 + \frac{S}{N} \right) \quad (5.31)$$

Soit E_b l'énergie d'un bit. Puisque P est la puissance du signal, nous avons :

$$E_b = \frac{P}{R}, \quad \text{où } R \text{ est le débit binaire.} \quad (5.32)$$

D'où

$$E_b = \frac{PN_0}{RN_0} = \frac{PN_0W}{RN_0W} \quad (5.33)$$

Donc

$$\frac{E_b}{N_0} = \frac{PW}{RN_0W} = \frac{SNR}{\rho} \quad \text{et} \quad \rho = \frac{SNR}{E_b/N_0} \quad (5.34)$$

En remplaçant dans l'inégalité de ρ , on arrive à :

$$\frac{E_b}{N_0} > \frac{SNR}{\log(1 + SNR)} \quad (5.35)$$

pour garantir une transmission fiable.

6.1 Introduction

Ce chapitre concerne la détection de l'information numérique transmise à travers un canal bruité. Considérons une source numérique de symboles extraits d'un certain alphabet $\mathcal{A} = \{\alpha_0, \alpha_1, \dots, \alpha_{m-1}\}$ ayant m symboles. Pour chaque symbole $\alpha_i \in \mathcal{A}$ délivré par la source, celui-ci est transmis à travers un canal bruité, et à la réception nous pouvons observer un nombre réel y correspondant à la somme de la valeur du symbole et de la valeur du bruit. La détection consiste à décider lequel des symboles α_i a vraisemblablement produit l'observation y .

6.2 Principes de la détection

Les symboles $\alpha_i \in \mathcal{A}$, $i \in \{0, 1, \dots, m-1\}$ sont chacun caractérisés par leurs probabilités d'apparition $p_i = Pr\{\alpha_i\}$.

Les probabilités p_i sont appelées les probabilités *à priori* des symboles de source.

On suppose aussi qu'on a des informations sur le canal de transmission. C'est-à-dire qu'on connaît ses probabilités de transitions $p(y|\alpha_i)$, $\alpha_i \in \mathcal{A}$.

Pour déterminer l'hypothèse la plus vraisemblable (c'est-à-dire effectuer la détection) nous avons besoin de calculer les probabilités $p(\alpha_i|y)$ pour tout $\alpha_i \in \mathcal{A}$ et la décision sur l'hypothèse la plus vraisemblable sera alors :

$$H(y) = \arg \max_{\alpha_i} p(\alpha_i|y) \quad (6.1)$$

Les probabilités $p(\alpha_i|y)$ sont appelées les probabilités *à posteriori*.

Remarque 16 Puisque y est réelle et continue, on utilise souvent les densités de probabilités $f(y|\alpha_i)$ au lieu des probabilités $p(y|\alpha_i)$.

6.3 Détection en binaire

Considérons une source binaire pouvant délivrer les symboles binaire 0 et 1. L'alphabet source est dans ce cas $\mathcal{A} = \{0, 1\}$. Les probabilités à priori sont notées p_0 et p_1 . On calcule les probabilités à posteriori :

$$p(y, \alpha_i) = p(y|\alpha_i)p(\alpha_i) = p(\alpha_i|y)p(y) \quad (6.2)$$

ou

$$p_i p(y|\alpha_i) = p(\alpha_i|y)p(y) \quad (6.3)$$

donc

$$p(\alpha_i|y) = \frac{p_i p(y|\alpha_i)}{p(y)} \quad (6.4)$$

Ou encore, en utilisant les densités de probabilité,

$$p(\alpha_i|y) = \frac{p_i f(y|\alpha_i)}{f(y)} \quad (6.5)$$

$p(y)$ peut être calculées à partir des probabilités de transition. Nous avons :

$$p(y) = \sum_i p(y, \alpha_i) = \sum_i p(y|\alpha_i)p_i \quad (6.6)$$

Plus précisément,

$$p(y) = p_0 p(y|0) + p_1 p(y|1) \quad (6.7)$$

En termes de densités de probabilités, on aura :

$$f(y) = p_0 f(y|0) + p_1 f(y|1) \quad (6.8)$$

Finalement, les probabilités a posteriori sont :

$$p(0|y) = \frac{p_0 f(y|0)}{f(y)} \quad \text{et} \quad p(1|y) = \frac{p_1 f(y|1)}{f(y)} \quad (6.9)$$

La décision sur l'hypothèse la plus vraisemblable se fait en comparant ces probabilités à posteriori. C'est à dire :

$$H = \begin{cases} 0; & \text{si } \frac{p_0 f(y|0)}{f(y)} \geq \frac{p_1 f(y|1)}{f(y)} \\ 1; & \text{sinon} \end{cases} \quad (6.10)$$

On peut réécrire ce résultat comme :

$$\frac{f(y|0)}{f(y|1)} \stackrel{\geq H=0}{< H=1} \frac{p_1}{p_0} = \eta \quad (6.11)$$

Si on pose $\Lambda(y) = \frac{f(y|0)}{f(y|1)}$ alors la décision est :

$$\Lambda(y) \begin{cases} \geq & H=0 \\ < & H=1 \end{cases} \eta \quad (6.12)$$

η est appelé le seuil de décision et $\Lambda(y)$, qui dépend de y est appelé le rapport de vraisemblance.

Le rapport de vraisemblance $\Lambda(y)$ est une fonction de la variable aléatoire y . Il est donc lui-même une variable aléatoire et on voit donc que l'on peut faire des erreurs de décision. En fait, on peut voir qu'on peut faire deux sortes d'erreurs selon que l'hypothèse vraie est $H = 0$ ou $H = 1$. Les probabilités d'erreurs correspondantes sont les suivantes :

$$Pr(e|H = 0) = Pr(\Lambda(y) < \eta | H = 0) \quad (6.13)$$

$$Pr(e|H = 1) = Pr(\Lambda(y) \geq \eta | H = 1) \quad (6.14)$$

La probabilité d'erreur moyenne est alors :

$$Pr(e) = p_0 Pr(e|H = 0) + p_1 Pr(e|H = 1) \quad (6.15)$$

6.4 Détection sur canal AWGN

On prend une source binaire et aux bits 0 et 1, on fait correspondre les niveaux de signal respectifs $+a$ et $-a$ ($a \in \mathbb{R}$) (codage BPSK) et les signaux sont transmis à travers un canal AWGN. Le schéma de principe est illustré sur la figure 6.1. On s'intéresse au

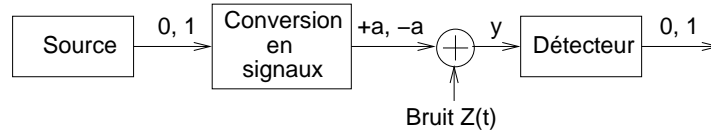


Figure 6.1 : Détection sur canal AWGN en utilisant une modulation BPSK

processus de décision et au calcul de la probabilité d'erreur. On suppose que le bruit $Z(t)$ est un bruit blanc gaussien de moyenne nulle et de variance $\frac{N_0}{2}$. Sa distribution est donc :

$$f_Z(z) = \frac{1}{\sqrt{2\pi N_0/2}} e^{-\frac{z^2}{N_0}} \quad (6.16)$$

Selon que l'hypothèse est $-a$ ou $+a$, l'observation y a les distributions suivantes :

$$f(y|0) = \frac{1}{\sqrt{\pi N_0}} e^{-\frac{(y-a)^2}{N_0}} \quad (6.17)$$

ou

$$f(y|1) = \frac{1}{\sqrt{\pi N_0}} e^{-\frac{(y+a)^2}{N_0}} \quad (6.18)$$

Le rapport de vraisemblance $\Lambda(y)$ est alors calculé comme étant :

$$\Lambda(y) = e^{\frac{4ay}{N_0}} \quad (6.19)$$

Donc la décision sur l'hypothèse vraisemblable se fait comme :

$$\Lambda(y) = \begin{cases} \geq H=0 \\ < H=1 \end{cases} \eta = \frac{p_1}{p_0} \quad (6.20)$$

ou mieux, on peut utiliser le $LLR(y)$ (Log Likelihood Ratio) défini par

$$LLR(y) = \ln \Lambda(y). \quad (6.21)$$

Nous écrivons donc :

$$LLR(y) = \begin{cases} \geq H=0 \\ < H=1 \end{cases} \ln(\eta) \quad (6.22)$$

Ce qui équivaut à :

$$y \begin{cases} \geq H=0 \\ < H=1 \end{cases} \frac{N_0 \ln(\eta)}{4a} \quad (6.23)$$

Traçons sur un même axe, les deux distributions conditionnées sur les deux hypothèses $H = 0$ et $H = 1$. Le résultat est montré sur la figure 6.2.

Nous pouvons maintenant calculer, comme auparavant, les probabilités d'erreur conditionnelles puis la probabilité d'erreur moyenne. Lorsque $H = 1$, la probabilité d'erreur est la valeur de la surface hachurée sur la figure 6.2. En effet, on fait une erreur de décision lorsque la valeur du bruit dépasse $a + \frac{N_0}{4a} \ln(\eta)$. Puisque la variance du bruit est $N_0/2$, une erreur est faite lorsque

$$\frac{Z}{\sqrt{N_0/2}} > \frac{a}{\sqrt{N_0/2}} + \frac{N_0 \ln(\eta)}{4a \sqrt{N_0/2}}$$

Donc la probabilité d'erreur lorsque $H = 1$ est :

$$Pr(e|H = 1) = Q \left(\frac{a}{\sqrt{N_0/2}} + \frac{\sqrt{N_0/2} \ln(\eta)}{2a} \right) \quad (6.24)$$

où Q est la fonction erreur complémentaire qui est définie par :

$$Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \quad (6.25)$$

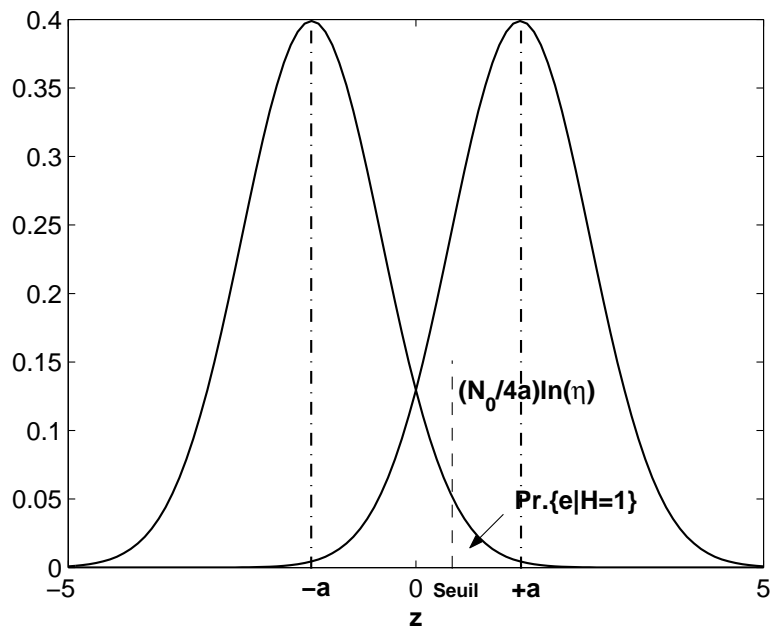


Figure 6.2 : Distributions normales conditionnées sur les hypothèses $+a$ et $-a$

Remarque 17 Montrer que

$$Pr(e|H = 0) = Q\left(\frac{a}{\sqrt{N_0/2}} - \frac{\sqrt{N_0/2} \ln(\eta)}{2a}\right) \quad (6.26)$$

Dans le cas où $p_0 = p_1$ on retrouve $Pr(e|H = 1) = Pr(e|H = 0)$ et la probabilité d'erreur moyenne se réduit à :

$$Pr(e) = Q\left(\frac{a}{\sqrt{N_0/2}}\right) \quad (6.27)$$

Bibliographie

- [1] R. G. Gallager, "Principles of digital communications," Mar. 2005. [Online]. Available : <http://www.mit.edu/opencourseware>
- [2] P. H. Hwei, *Communications analogiques et numériques Cours et problèmes*. McGraw-Hill Inc., 1994.
- [3] G. Baudoin et coll., *Radiocommunications numériques / 2*. Dunod, Paris, 2002.
- [4] P. H. Hwei, *Signaux et Communications*. Mc. Graw-Hill, Inc., 1994.
- [5] J. G. Proakis, *Digital Communications*. Mc. Graw-Hill, Inc., 1989.