Communications Numériques et Théorie de l'Information

Maurice Charbit

14 mai 2007

Table des matières

Chapitre 1

Représentation en enveloppe complexe

1.1 Rappels de théorie du signal

Dans le cas des signaux déterministes à temps continu, on distingue les signaux d'énergie finie servant le plus souvent à modéliser les signaux de durée finie ou à décroissance rapide et les signaux de puissance finie plus particulièrement les mélanges de sinusoïdes de la forme :

$$x(t) = \sum_{k=1}^{P} A_k \cos(2\pi f_k t + \phi_k)$$

Energie et Puissance

L'*énergie* d'un signal $x(t)$, fonction complexe de la variable réelle t, est la quantité E définie par :

$$E = \int_{-\infty}^{+\infty} |x(t)|^2 dt \tag{1.1}$$

La *puissance* d'un signal $x(t)$, fonction complexe de la variable réelle t, est la quantité P définie par :

$$P = \lim_{T \to +\infty} \frac{1}{T} \int_{-T/2}^{+T/2} |x(t)|^2 dt \tag{1.2}$$

Si $x(t)$ est périodique de période T de la forme :

$$x(t) = \sum_{k=1}^{P} A_k \cos(2\pi kt/T + \phi_k)$$

la puissance est donnée par :

$$P = \frac{1}{T} \int_{-T/2}^{+T/2} |x(t)|^2 dt = \sum_{k=1}^{P} A_k^2/2$$

Représentation fréquentielle des signaux

Lorsqu'on applique le signal complexe $e^{2j\pi f_0 t}$ à l'entrée d'un filtre linéaire, le signal en sortie a pour expression $Ae^{2j\pi f_0 t}$ où A est une constante ne dépendant que de la valeur f_0. C'est une des raisons de l'importance de la décomposition d'un signal en une somme d'exponentielles complexes. Cette décomposition porte le nom de *représentation de Fourier* ou représentation *fréquentielle* ou plus simplement spectre.

Pour un signal $x(t)$ périodique de période T et de puissance finie, on a les formules de Fourier suivantes[1] :

$$\begin{cases} x(t) = \sum_{n=-\infty}^{+\infty} X_n e^{2j\pi nt/T} \\ X_n = \frac{1}{T} \int_{-T/2}^{T/2} x(t) e^{-2j\pi nt/T} dt \end{cases}$$

[1]La convergence dans le développement de $x(t)$ est moyenne quadratique. Il peut ne pas y avoir convergence uniforme. Ce phénomène n'a aucune conséquence pour nous dans la suite.

Les coefficients X_n s'appellent les coefficients de Fourier du signal $x(t)$. Du fait que ces formules établissent une correspondance bijective entre la fonction $x(t)$ et la suite X_n de ses coefficients de Fourier, il n'y a pas plus d'information dans l'une ou l'autre de ces représentations. Toutefois en traitement du signal, elles ont chacune leur intérêt. De façon imagée, on peut dire que ce sont deux manières de "voir le même phénomène sous des angles différents".

Pour un signal $x(t)$ d'énergie finie, les formules de transformations de Fourier directe et inverse sont :

$$\begin{cases} X(f) = \int_{-\infty}^{+\infty} x(t)e^{-2j\pi ft}dt \\ x(t) = \int_{-\infty}^{+\infty} X(f)e^{2j\pi ft}df \end{cases} \tag{1.3}$$

La variable f s'appelle la *fréquence*. Son unité est le *Hertz* (en abrégé : Hz).

Exemple 1.1 *Soit le signal rectangle $x(t) = 1$ pour $t \in (-T/2, T/2)$ et 0 sinon. Un calcul immédiat donne pour sa transformée de Fourier :*

$$X(f) = \frac{\sin(\pi fT)}{\pi f} = T\text{sinc}(fT) \tag{1.4}$$

où la fonction $\text{sinc}(x) = \sin(\pi x)/(\pi x)$ *s'appelle la fonction sinus-cardinal (car elle s'annule pour les valeurs entières de la variable).*

Le tableau ci-dessous donne les principales propriétés de la transformation de Fourier.

Propriétés	$x(t)$	$X(f)$
Similitude	$x(at)$	$\frac{1}{\|a\|}X(f/\|a\|)$
	$x^*(-t)$	$X^*(f)$
Linéarité	$ax(t) + by(t)$	$aX(f) + bY(f)$
Translation	$x(t - t_0)$	$X(f)\exp(-2j\pi f_0 t)$
Modulation	$x(t)\exp(2j\pi ft_0)$	$X(f - f_0)$
Convolution	$x(t) \star y(t)$	$X(f)Y(f)$
Produit	$x(t)y(t)$	$X(f) \star Y(f)$
Dérivations	$d^n x(t)/dt^n$	$(2j\pi f)^n X(f)$
	$(-2j\pi t)^n x(t)$	$d^n X(f)/df^n$
Parité, conjugaison	réelle paire	réelle paire
	réelle impaire	imaginaire impaire
	imaginaire paire	imaginaire paire
	imaginaire impaire	réelle impaire
	complexe paire	complexe paire
	complexe impaire	complexe impaire
	réelle	$X(f) = X^*(-f)$ $\text{Re}(X(f)), \|X(f)\|$ paires $\text{Im}(X(f)), \arg(X(f))$ impaires

Formule de Parseval

En utilisant la propriété de convolution des transformées de Fourier, on obtient :

$$\int_{-\infty}^{+\infty} x(t)y^*(t)dt = \int_{-\infty}^{+\infty} X(f)Y^*(f)df$$

Filtrage linéaire

Soit le signal $x(t)$ d'énergie finie et $h(t)$ une fonction de module sommable ($\int_{\mathbb{R}} |h(t)|dt < +\infty$). Le signal :

$$y(t) = \int_{-\infty}^{+\infty} x(u)h(t-u)du = \int_{-\infty}^{+\infty} x(t-u)h(u)du \tag{1.5}$$

définit l'opération dite de *convolution* qui est notée de façon plus concise $y(t) = x(t) \star h(t)$. Le système, qui associe au signal d'entrée $x(t)$ le signal de sortie $y(t)$, s'appelle un filtre linéaire et la fonction $h(t)$, qui caractérise le filtre, s'appelle sa réponse impulsionnelle. La condition $\int_{\mathbb{R}} |h(t)|dt < +\infty$ est équivalente à la propriété dite de stabilité qui assure qu'à toute entrée bornée correspond une sortie bornée. Pour obtenir l'équation de filtrage dans le domaine des fréquences, il suffit de se rappeler que la transformée de Fourier d'un produit de convolution est un produit simple. Par conséquent l'expression 1.5 donne en passant aux transformées de Fourier :

$$Y(f) = H(f)X(f)$$

Dans ce contexte $H(f)$ s'appelle le gain complexe du filtre.

Par linéarité, si le signal à l'entrée est un mélange harmonique de la forme $x(t) = \sum_{k=1}^{P} \alpha_k e^{2j\pi f_k t}$ où α_k est une suite de valeurs complexes et f_k une suite de fréquences réelles, la sortie a pour expression $y(t) = \sum_{k=1}^{P} \alpha_k \mathcal{F}(e^{2j\pi f_k t})$ où $\mathcal{F}$ désigne l'opération de filtrage. Il suffit donc de déterminer l'expression de la sortie lorsque l'entrée est de la forme $x(t) = e^{2j\pi f_0 t}$. Notons que, dans ce cas, $x(t)$ n'est pas d'énergie finie. Toutefois en portant $x(t) = \exp(2j\pi f_0 t)$ dans 1.5, on obtient comme expression pour le signal en sortie :

$$y(t) = \int_{-\infty}^{+\infty} e^{2j\pi f_0(t-u)} h(u)du = H(f_0)e^{2j\pi f_0 t}$$

Une autre façon d'écrire est $\mathcal{F}(e^{2j\pi f_k t}) = H(f_0)e^{2j\pi f_k t}$. On dit que les exponentielles complexes $e^{2j\pi f_0 t}$ sont les fonctions propres des filtres linéaires.

1.2 Echantillonnage

L'échantillonnage est une opération qui consiste à prélever sur un signal à temps continu une suite de valeurs, prises en une suite d'instants t_n, $n \in \mathbb{Z}$. Dans la suite nous n'envisagerons que l'échantillonnage dit régulier où $t_n = nT$. L'intérêt porté aux problèmes de l'échantillonnage tient dans le développement des techniques numériques de traitement du signal. La question fondamentale est de savoir s'il est possible de reconstruire $x(t)$ à partir des échantillons $x(nT)$. A première vue il existe une infinité de fonctions qui passent par les valeurs $x(nT)$ aux instants nT. Toutefois le théorème d'échantillonnage montre que, pour les signaux à bande limitée, la reconstruction est possible.

1.3 Cas des signaux passe-bas

Théorème 1.1 (Formule de Poisson) *Soit $x(t)$ un signal de module intégrable ($x(t) \in L^1(\mathbb{R})$) et dont la transformée de Fourier $X(f)$ est elle-même de module intégrable. On a alors pour tout $T > 0$:*

$$\sum_{n=-\infty}^{+\infty} X(f - \frac{n}{T}) = T \sum_{k=-\infty}^{+\infty} x(kT)e^{-2j\pi kfT} \tag{1.6}$$

En effet le premier membre de 2.41 est une fonction de f de période $1/T$. Elle est donc développable en série de Fourier sous la forme $\sum_k X_k e^{-2j\pi kfT}$, où X_k est donné par :

$$X_k = T \int_{-1/2T}^{1/2T} \left(\sum_{n=-\infty}^{+\infty} X(f - \frac{n}{T}) \right) e^{2j\pi kfT} df = T \sum_{n=-\infty}^{+\infty} \int_{-1/2T}^{1/2T} X(f - \frac{n}{T}) e^{2j\pi kfT} df$$

En faisant le changement de variable $u = f - n/T$, il vient :

$$X_k = T \sum_{n=-\infty}^{+\infty} \int_{-1/2T-n/T}^{1/2T-n/T} X(u)e^{2j\pi kuT} df = T \int_{-\infty}^{+\infty} X(u)e^{2j\pi kuT} df = Tx(kT)$$

qui est le résultat annoncé.

Théorème 1.2 (Théorème d'échantillonnage) *Soit un signal réel $x(t)$ de module intégrable ($x(t) \in L^1(\mathbb{R})$), à bande limitée B ($X(f) = 0$ pour $|f| > B$) et soit $F_e = 1/T$ une fréquence d'échantillonnage. On suppose que $\sum_{\mathbb{Z}} |x(nT)| < +\infty$.*

Si $F_e \geq 2B$, $x(t)$ peut être reconstruit de manière unique à partir de la suite d'échantillons $x(nT)$, suivant la formule dite d'interpolation :

$$x(t) = \sum_{n=-\infty}^{+\infty} x(nT)h(t-nT) \tag{1.7}$$

où

$$h(t) = \frac{\sin(2\pi Bt)}{\pi F_e t} \tag{1.8}$$

Si $F_e < 2B$, la reconstruction est impossible. La fréquence minimale $2B$ s'appelle la fréquence de Nyquist.

Cela signifie que pour un signal qui a de l'énergie dans les fréquences élevées et donc des variations rapides, il faut prendre une fréquence d'échantillonnage élevée. En pratique ce résultat est appliqué, de façon intuitive, lors du relevé d'une courbe point par point : dans les parties à variations rapides (hautes fréquences), on augmente la fréquence d'échantillonnage en prenant un plus grand nombre de points.

Le problème de l'échantillonnage, tel qu'il est posé ici, consiste à montrer que, pour une certaine classe de signaux $x(t)$, il est possible de faire coïncider $x(t)$ avec :

$$\tilde{x}(t) = \sum_{n=-\infty}^{+\infty} x(nT)h(t-nT) \tag{1.9}$$

pour une certaine fonction $h(t)$ à déterminer. La relation 1.9 est une équation de convolution semblable à celle rencontrée dans le cas du filtrage linéaire, sauf qu'ici l'entrée est la suite $x(nT)$ à temps discret et la sortie le signal $\tilde{x}(t)$ à temps continu.

Afin de comparer $\tilde{x}(t)$ et $x(t)$ nous allons passer en fréquence. Pour cela notons $H(f)$ la transformée de Fourier de $h(t)$. Alors $h(t-nT)$ a pour transformée de Fourier $H(f)e^{-2j\pi nfT}$. On en déduit que $\tilde{x}(t)$ a pour transformée de Fourier :

$$\tilde{X}(f) = \sum_{n=-\infty}^{+\infty} x(nT)H(f)e^{-2j\pi nfT}$$

En sortant $H(f)$ du signe somme et en utilisant la formule de Poisson, on obtient :

$$\tilde{X}(f) = \frac{1}{T}H(f) \sum_{n=-\infty}^{+\infty} X(f - \frac{n}{T})$$

Cette expression fait dire que l'opération d'échantillonnage en temps a pour effet, en fréquence, de *périodiser* le spectre du signal avec une période égale à la fréquence d'échantillonnage $F_e = 1/T$. Il est à noter que le résultat est vrai même si $X(f)$ n'est pas à bande limitée. Toutefois quand $X(f)$ est à bande limitée, il est possible de choisir $H(f)$ de façon à ce que cette expression coïncide avec $X(f)$.

Supposons que $X(f) = 0$ pour $|f| > B$. Deux cas sont possibles :

$F_e = 1/T < 2B$ **:** il y a recouvrement des différentes courbes obtenues par périodisation de $X(f)$. On dit alors qu'il y a *repliement de spectre* (en anglais *aliasing*). L'origine de ce terme s'explique de la façon suivante : la partie de $X(f-n/T)$ qui s'ajoute à $X(f)$ dans l'intervalle $(-1/2T, 1/2T)$ est la même que la partie de $X(f)$ qui se trouve au delà de n/T. Tout se passe comme si on empilait dans l'intervalle $(-1/2T, 1/2T)$, après repliement, les deux extrémités de $X(f)$. La conséquence du repliement de spectre est l'impossibilité de reconstruire $X(f)$ à partir de $\tilde{X}(f)$ et, par là même, $x(t)$ à partir des échantillons $x(nT)$.

$F_e = 1/T \geq 2B$ **:** en choisissant $H(f) = T\text{rect}_{2B}(f)$, il vient $X(f) = \tilde{X}(f)$ et donc $x(t) = \tilde{x}(t)$. La transformée de Fourier inverse de $H(f) = T\text{rect}_{2B}(f)$ a pour expression $h(t) = T\sin(2\pi Bt)/\pi t$. En portant dans 1.9, on obtient la formule d'interpolation :

$$x(t) = \sum_n x(nT)\frac{\sin(2\pi B(t-nT))}{\pi F_e(t-nT)}$$

La formule d'interpolation montre que le signal réel $x(t)$ est reconstruit de façon unique à partir de la suite de ses échantillons $x(nT)$. Mais cette opération n'est *pas causale* puisque la reconstruction de $x(t)$ au temps t, nécessite de connaître la suite $x(nT)$ au delà de t. Toutefois comme la fonction $h(t)$ décroît rapidement quand t tend vers $-\infty$, il est possible de réaliser une bonne approximation causale, en acceptant un retard fini. Cela revient à dire que $x(t)$ est calculé, de façon approchée, avec quelques échantillons situés au delà de t.

FIG. 1.1 – *Périodisation du spectre par échantillonnage*

Cas des signaux complexes à bande limitée

Pour un signal complexe dont le spectre est nul à l'extérieur d'une bande B_c, c'est-à-dire $X(f) = 0$ pour $f \notin B_c$, le calcul de la transformée de Fourier de $\tilde{x}(t)$ est en tout point identique à celui fait précédemment. On en déduit que la fréquence minimale d'échantillonnage, qui s'obtient en exprimant simplement la condition de non-repliement, a pour expression $F_e = 1/T \geq B_c$. En fait on peut dire que, dans les cas réel et complexe, la fréquence de Nyquist est égale à la largeur du support en fréquence de la transformée de Fourier de $x(t)$.

1.4 Cas des signaux réels passe-bande

Considérons à présent un signal réel $x(t)$ dont la transformée de Fourier est nulle en dehors des deux intervalles de fréquence définis par :

$$f_m \leq |f| \leq f_M$$

FIG. 1.2 – *Spectre d'un signal à bande étroite*

Rappelons que puisque $x(t)$ est réel, $X(f)$ possède la symétrie hermitienne. L'application brutale du théorème d'échantillonnage conduit à prendre comme fréquence de Nyquist la valeur $2f_M$. Pourtant il est possible d'échantillonner à une cadence bien plus faible, si l'on met à profit le fait que le spectre est nul dans l'intervalle $(-f_m, f_m)$.

Cherchons les conditions sur F_e pour que le spectre, une fois périodisé, soit constitué de bandes *disjointes*.

On voit graphiquement qu'il suffit de choisir deux valeurs k et F_e, telles que la k-ième et la $(k+1)$-ième translatées de la partie de $X(f)$ dans les fréquences négatives ne recouvrent pas la partie de $X(f)$ dans les

FIG. 1.3 – *Périodisation du spectre pour un signal à bande étroite*

fréquences positives. Ce qui s'écrit :

$$\begin{aligned} -f_m + kF_e &< f_m \\ -f_M + (k+1)F_e &> f_M \end{aligned}$$

Par conséquent F_e doit être choisie dans des plages de valeurs de la forme :

$$\frac{2f_M}{k+1} < F_e < \frac{2f_m}{k} \tag{1.10}$$

où k est un entier tel que $(2f_M/k+1) < 2f_m/k$, c'est-à-dire $k \leq f_m/(f_M - f_m)$. Plus la valeur choisie de k est grande, plus la plage de fréquences possibles d'échantillonnage est située dans les fréquences basses. Par conséquent la plus petite fréquence d'échantillonnage qui assure le non repliement du spectre est donc donnée par $2f_M/(k_0+1)$ où k_0 est la *partie entière* de $f_m/(f_M - f_m)$. Les fréquences F_e d'échantillonnage permises sont regroupées dans le tableau ci-desous.

	Plage pour F_e		
k_0	$2f_M/(k_0+1)$	$\leq F_e \leq$	$2f_m/k_0$
$\vdots$		$\vdots$	
k	$2f_M/(k+1)$	$\leq F_e \leq$	$2f_m/k$
$\vdots$		$\vdots$	
0	$2f_M$	$\leq F_e <$	$+\infty$

Remarquons que, plus la fréquence d'échantillonnage est choisie petite, plus la plage de fréquences à laquelle elle appartient est étroite.

Formule d'interpolation

Pour établir la formule de reconstruction, le calcul est en tout point analogue à celui fait pour un signal passe-bas. Mais il faut prendre, pour faire coïncider $\tilde{x}(t)$ avec $x(t)$, le filtre passe-bande réel, défini par :

$$H(f) = T\left(\text{rect}_{\Delta f}(f - f_0) + \text{rect}_{\Delta f}(f + f_0)\right)$$

où $\Delta f = f_M - f_m$ et $f_0 = (f_M + f_m)/2$. Et donc $h(t) = 2T\cos(2\pi f_0 t)\sin(\pi \Delta f t)/\pi t$. Il suffit alors d'utiliser l'expression $x(t) = \sum_n x(nT)h(t - nT)$ pour obtenir la formule d'interpolation.

1.5 Cas des signaux passe-bas de bande infinie

En pratique, lorsque la fréquence d'échantillonnage est imposée, le phénomène de repliement de spectre ne peut être évité. Il y a donc perte d'information sur le signal à échantillonner. Le problème est de limiter autant que possible cette perte. Pour cela on choisit de filtrer *préalablement* le signal avant l'opération d'échantillonnage proprement dite, suivant le schéma représenté figure 1.4.

FIG. 1.4 – *Préfiltrage du signal avant échantillonnage*

A priori le signal $x_2(t)$ reconstruit doit contenir toutes les fréquences compatibles avec la condition de non-repliement à la fréquence d'échantillonnage $F_e = 1/T$: il faut donc supposer que la bande $B = F_e/2$. Dans ce cas le filtre $H(f)$ a pour gain complexe $H(f) = T\text{rect}_{F_e}(f)$.

Afin de déterminer au mieux le filtre $G(f)$, nous allons minimiser l'écart quadratique :

$$\epsilon^2 = \int_{-\infty}^{+\infty} |x(t) - x_2(t)|^2 dt$$

entre le signal original $x(t)$ et le signal $x_2(t)$ obtenu à partir des échantillons $x_1(nT)$. Avec des notations évidentes, en utilisant la formule de Parseval, on a encore :

$$\epsilon^2 = \int_{-\infty}^{+\infty} |X(f) - X_2(f)|^2 df$$

Déterminons l'expression de $X_2(f)$. Il vient en utilisant la formule de Poisson :

$$X_2(f) = \sum_n x_1(nT)H(f)e^{-2j\pi nfT} = \frac{1}{T}\sum_n X_1(f - n/T)H(f)$$

Comme $H(f) = T\text{rect}_{F_e}(f)$ et que $X_1(f) = X(f)G(f)$, on en déduit que :

$$X_2(f) = \text{rect}_{F_e}(f)\sum_n X(f - n/T)G(f - n/T) \tag{1.11}$$

et donc que :

$$\epsilon^2 = \int_{|f|<F_e/2} |X(f) - X_2(f)|^2 df + \int_{|f|>F_e/2} |X(f)|^2 df \tag{1.12}$$

Comme tous les termes sont positifs et que le second terme du membre droit de l'équation (1.11) ne dépend pas du choix de $G(f)$, le minimum est obtenu en prenant $G(f) = \text{rect}_{F_e}(f)$: en effet, dans ce cas et d'après 1.11, $X_2(f) = X(f)\text{rect}_{F_e}(f)$, ce qui annule complètement le premier terme de l'équation (1.11).

Ce résultat est important, puisqu'il indique que l'on doit faire précéder l'opération d'échantillonnage d'un filtrage passe-bas idéal dans la bande $(-F_e/2, F_e/2)$, appelé filtrage *anti-repliement*. Évidemment il y a perte d'information et ce que l'on peut reconstruire, au mieux, est le signal $x_1(t)$. Ce qui est hors de la bande $(-F_e/2, F_e/2)$ est perdu.

Exemple 1.2 (Signal MIC en téléphonie numérique) *le signal téléphonique est eéchantillonné à la fréquence de* 8000 *échantillons/s. Pour éviter le repliement, on effectue un filtrage du signal dans la bande* $(0 - 3400Hz)$ *légèrement plus étroite que le minimum requis de* $4000Hz$*. Chaque échantillon est ensuite codé sur* 8 *bits. On obtient ainsi un débit de* $64kbits/s$*. Cette suite est désignée par le terme de MIC (pour Modulation par Impulsion et Codage).*

1.6 Reconstruction pratique

La conversion du signal analogique à partir de la suite numérique nécessite une bonne approximation causale d'un filtre passe-bas idéal. En pratique la façon la plus simple de procéder consiste à partir d'un *bloqueur d'ordre 0*, qui bloque pendant toute la durée T entre deux échantillons la valeur numérique appliquée à l'entrée. Le signal obtenu en sortie du bloqueur d'ordre 0 a donc pour expression :

$$x_0(t) = \sum_n x(nT)h_0(t - nT)$$

où $h_0(t) = \text{rect}(t - T/2)$.

FIG. 1.5 – *Reconstruction par bloqueur d'ordre 0*

Le signal $x_0(t)$ est un signal en escalier dont les marches ont pour amplitude $x((n+1)T) - x(nT)$. En prenant la transformée de Fourier de $x_0(t)$ et en utilisant la formule de Poisson, on obtient :

$$X_0(f) = \sum_n H_0(f)X(f - n/T) \quad \text{où} \quad H_0(f) = \frac{\sin(\pi fT)}{\pi f} e^{-j\pi fT}$$

FIG. 1.6 – *Spectre en sortie du bloqueur d'ordre 0*

En observant $|X_0(f)|$ représenté à la figure 1.6, on voit apparaître 2 formes de distorsion entre le signal de départ $x(t)$ et le signal $x_0(t)$ en sortie du bloqueur d'ordre 0.

1. distorsion dans la bande utile $(-1/2T, 1/2T)$ du signal. Un remède est de réaliser avant échantillonnage une compensation par le filtre de gain $\pi fT/\sin(\pi fT)$.
2. distorsion hors de la bande utile $(-1/2T, 1/2T)$. Cette distorsion peut être génante : ainsi, en acoustique, si elle se trouve dans la bande audible, il faut alors utiliser un filtre passe-bas de fréquence de coupure B.

Le calcul précédent montre que les lobes de la fonction en sinus-cardinal ont pour largeur $1/T$ où T représente la durée de la fonction de reconstruction du bloqueur mais aussi la période d'échantillonnage du signal. D'où l'idée de faire précéder le bloqueur d'une opération d'interpolation. Cette opération est possible puisque le signal vérifie les conditions du théorème d'échantillonnage. Dans ce cas l'énergie hors de la bande utile est située essentiellement autour de la fréquence $1/T$. Ainsi en audio, en choisissant le facteur d'interpolation suffisamment grand, la bande de fréquence autour de $1/T$ peut même être hors de la bande audible et l'écoute ne nécessite alors aucun filtre supplémentaire en sortie du bloqueur.

1.7 Représentation en enveloppe complexe ou représentation en phase et quadrature

Dans ce chapitre nous nous intéressons à des signaux *réels* de la forme :

$$x(t) = p(t)\cos(2\pi f_p t) - q(t)\sin(2\pi f_p t) \tag{1.13}$$

Comme nous le verrons ces signaux interviennent couramment dans les systèmes de communication. Dans ce contexte les signaux $p(t)$ et $q(t)$ dépendent du message à transmettre et la fréquence f_p est dite *fréquence*

porteuse. Dans le cas où $f_p \gg B$, où B désigne la bande en fréquence de $p(t)$ et $q(t)$, le signal $x(t)$ peut être vu comme une "sinusoïde" de fréquence f_p, dont l'amplitude et la phase varient lentement et son spectre est alors à bande "étroite" autour de f_p.

Pour de tels signaux, il est d'usage d'utiliser la représentation dite en *enveloppe complexe* dont la justification est contenue dans la remarque suivante : le signal $x(t)$ étant *réel*, la partie du spectre dans les fréquences négatives se déduit (par symétrie hermitienne) de la partie du spectre dans les fréquences positives. On peut donc, de façon bijective, associer à $x(t)$ un signal *complexe* $x_b(t)$ dont le spectre s'obtient en translatant de $(-f_0)$ la partie du spectre de $x(t)$ située dans les fréquences positives. A priori la fréquence f_0 peut être choisie arbitrairement. Cependant pour les signaux donnés par l'expression (1.13), il est souvent pratique de prendre $f_0 = f_p$. Le processus complexe $x_b(t) = p_x(t) + jq_x(t)$, associé au signal réel $x(t)$, s'appelle son enveloppe complexe par rapport à f_0. Elle est un outil essentiel pour décrire, en communication, une opération de modulation.

Remarquons ici que la représentation en enveloppe complexe ne nécessite pas que le signal soit à bande étroite. Elle est seulement liée au fait que le signal est réel et centré.

1.7.1 Enveloppe complexe d'un signal

Pour définir l'enveloppe complexe il est nécessaire d'introduire au préalable la notion de signal analytique.

Définition 1.1 *On appelle signal analytique associé au signal $x(t)$ réel, centré, le signal complexe $z_x(t)$ obtenu à partir de $x(t)$ par un filtrage linéaire de gain complexe $2 \times \mathbb{1}_{(0,+\infty)}(f)$.*

Propriétés 1.1 *Le signal analytique $z(t)$ associé à $x(t)$ vérifie :*

$$x(t) = Ré\{z_x(t)\} \tag{1.14}$$

En effet, soit $x(t)$ un signal déterministe, réel, appliqué à l'entrée du filtre de gain complexe $2 \times \mathbb{1}_{(0,+\infty)}(f)$ et $z_x(t)$ le signal en sortie. On note $X(f)$ et $Z_x(f)$ leurs transformées de Fourier respectives. D'après la relation de filtrage, exprimée en fréquence, $Z_x(f) = 2 \times \mathbb{1}_{(0,+\infty)}(f)X(f)$.

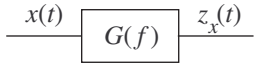

FIG. 1.7 – *Filtrage analytique : gain $G(f) = 2 \times \mathbb{1}pf \in (0, +\infty)$*

Déterminons la transformée de Fourier de Ré $\{z_x(t)\} = (z_x(t) + z_x^*(t))/2$. Il vient $\frac{1}{2}(Z_x(f) + Z_x^*(f)) = \mathbb{1}_{(0,+\infty)}(f)X(f) + \mathbb{1}^*_{(0,+\infty)}(-f)X^*(-f) = X(f)$, où on a utilisé que le signal $x(t)$ est réel et donc que $X(f) = X^*(-f)$. Par conséquent Ré $\{z_x(t)\} = x(t)$. Le signal complexe en sortie du filtre analytique a donc comme partie réelle le signal d'entrée.

On montre de la même façon que la transformée de Fourier de la partie imaginaire de $z_x(t)$ a pour expression $-j\text{signe}(f)X(f)$. On dit alors que Imag $(z_x(t))$ est la *transformée de Hilbert* du signal $x(t)$. Nous la notons $\hat{x}(t)$.

On note que, si $Z_x(f)$ est nul pour les fréquences négatives, alors $Z_x(f) = Z_x(f)\mathbb{1}_{(0,+\infty)}(f)$ et les parties réelle et imaginaire de $z_x(t)$ forment une paire de transformées de Hilbert. On en déduit une règle simple pour déterminer le signal analytique associé à un signal réel $x(t)$: *si $x(t)$ est tel que $x(t) = Ré\{z_x(t)\}$ et tel que $Z_x(f) = 0$ pour $f < 0$, alors $z_x(t)$ est le signal analytique associé à $s(t)$.*

Définition 1.2 *On appelle enveloppe complexe par rapport à la fréquence f_0 du signal $x(t)$ réel, le signal :*

$$x_b(t) = z_x(t)\exp(-2j\pi f_0 t)$$

où $z_x(t)$ désigne le signal analytique associé à $x(t)$. Si on note $X_b(f)$ et $Z_x(f)$ les transformées de Fourier respectivement de $x_b(t)$ et de $z_x(t)$, on a

$$X_b(f) = Z_x(f + f_0) = 2X^+(f + f_0) \tag{1.15}$$

où $X^+(f)$ désigne la partie de $X(f)$ située dans les fréquences positives.

L'expression 1.15 fournit une *règle simple de construction* de la transformée de Fourier de l'enveloppe complexe à partir de la transformée de Fourier du signal. Inversement partant de la transformée de Fourier de l'enveloppe complexe, on déduit la transformée de Fourier du signal réel (1) en translatant de $+f_0$ la transformée de Fourier de l'enveloppe complexe, (2) en complétant par symétrie hermitienne autour de $-f_0$ et (3) en divisant par 2.

Propriétés 1.2 *L'enveloppe complexe $x_b(t)$ de $x(t)$ par rapport à la fréquence f_0 vérifie :*

$$x(t) = Ré\{x_b(t)\exp(2j\pi f_0 t)\} \tag{1.16}$$

Il suffit d'appliquer la propriété 1.14.

Définition 1.3 *Pour un signal réel, on appelle respectivement composante en phase et composante quadrature (sous-entendu par rapport à la fréquence f_0), les parties réelle et imaginaire de son enveloppe complexe par rapport à la fréquence f_0.*

En notant $x_b(t) = p_x(t) + jq_x(t)$ et en utilisant l'expression 1.16, il vient :

$$x(t) = p_x(t)\cos(2\pi f_0 t) - q_x(t)\sin(2\pi f_0 t)$$

Propriétés 1.3 *En notant $P_x(f)$ et $Q_x(f)$ les transformées de Fourier respectives de $p_x(t)$ et de $q_x(t)$, on a :*

$$P_x(f) = \frac{1}{2}(X_b(f) + X_b^*(-f)) = X^+(f+f_0) + X^-(f-f_0) \tag{1.17}$$

$$jQ_x(f) = \frac{1}{2}(X_b(f) - X_b^*(-f)) = X^+(f+f_0) - X^-(f-f_0) \tag{1.18}$$

où $X_b(f) = Z_x(f+f_0)$ avec $Z_x(f) = 2\mathbb{1}_{(0,+\infty)}(f)X(f)$. et où $X^+(f)$ et $X^-(f)$ désignent les parties de $X(f)$ situées respectivement dans les fréquences positives et négatives.

Pour montrer cette propriété, il suffit de noter que $2p_x(t) = x_b(t) + x_b^*(t)$ et que $2jq_x(t) = x_b(t) - x_b^*(t)$ et d'effectuer les transformations de Fourier des deux membres (on rappelle que la transformée de Fourier de $g^*(t)$ est $G^*(-f)$).

On en déduit (voir figure 1.8) une *construction graphique simple* des transformées de Fourier des composantes en phase et quadrature par rapport à f_0 : (1) on translate de $(+f_0)$ la partie de $X(f)$ situées dans les fréquences négatives, (2) on translate de $(-f_0)$ la partie de $X(f)$ situées dans les fréquences positives, puis on effectue la somme pour obtenir $P_x(f)$ et la différence pour obtenir $Q_x(f)$ (notons ici que la somme et la différence portent sur des fonctions complexes).

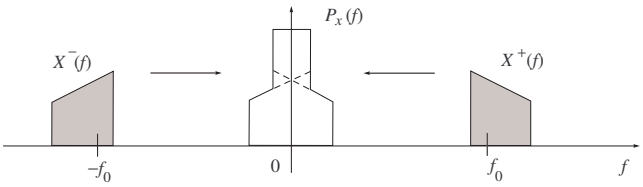

FIG. 1.8 – *Construction du spectre de la composante en phase.*

Définition 1.4 *Pour un signal réel, on appelle respectivement enveloppe et phase instantanée (sous-entendu par rapport à la fréquence f_0), le module et la phase de son enveloppe complexe par rapport à la fréquence f_0. On appelle fréquence instantanée :*

$$f_i(t) = \frac{1}{2\pi}\frac{d\phi_x(t)}{dt}$$

En notant $x_b(t) = a_x(t)e^{j\phi_x(t)}$ et en utilisant l'expression 1.16, il vient :

$$x(t) = a_x(t)\cos(2\pi f_0 t + \phi_x(t)) \tag{1.19}$$

On note que l'expression de l'enveloppe complexe dépend du choix de f_0. Toutefois quand il n'y aura pas d'ambiguïté, nous parlerons de l'enveloppe complexe d'un signal, sans indiquer explicitement f_0. Par contre l'enveloppe $a_x(t) = |x_b(t)|$, qui est aussi égale au module $|z_x(t)|$ du signal analytique, *ne dépend pas du choix de* f_0.

Notons enfin que les transformées de Fourier de $a_x(t)$ et de $\phi_x(t)$ ne s'expriment pas simplement à partir de $X(f)$, du fait du caractère non linéaire qui lie $a_x(t)$ et $\phi_x(t)$ à $x(t)$. Toutefois en règle générale si $x(t)$ est à bande étroite, $a_x(t)$ et $\phi_x(t)$ sont eux-même à bande étroite (par rapport à f_0). Dans ce cas la relation 1.19 permet d'interpréter $x(t)$ comme une "sinusoïde" de fréquence f_0 dont l'amplitude et la phase sont "lentement" variables. De là le nom de signal quasi-monochromatique qu'ils portent aussi dans la littérature.

Deux traitements jouent un rôle important en communication : le filtrage et la démodulation synchrone. Voyons à présent comment ces traitements peuvent être faits au moyen des enveloppes complexes : on dit que le calcul est fait en *en bande de base*.

1.7.2 Filtrage équivalent en bande de base

Soit le signal réel $x(t)$ dont le spectre est nul hors d'une bande B autour de f_0. On suppose que $f_0 > B$. On note $x_b(t)$ l'enveloppe complexe de $x(t)$ par rapport à f_0. Le signal $x(t)$ est filtré par un filtre réel dont le gain $H(f)$ est supposé, sans perte de généralités, nul hors de la bande de $x(t)$. On note $y(t)$ le signal en sortie. On montre aisément que $y(t)$ a pour enveloppe complexe par rapport à f_0 :

$$y_b(t) = x_b(t) \star h_b(t) \Leftrightarrow Y_b(f) = X_b(f) \star H_b(f)$$

où

$$H_b(f) = H^+(f + f_0) \tag{1.20}$$

Le filtre de gain $H_b(f)$ est appelé *le filtre équivalent en bande de base*. Attention contrairement à la formule 1.15, il n'y a pas ici le facteur 2. Le filtre équivalent en bande de base s'obtient en translatant de f_0 la partie de $H(f)$ située dans les fréquences positives. Pour obtenir la réponse impulsionnelle il suffit, si besoin est, de prendre la transformée de Fourier inverse. Evidemment le filtre équivalent en bande de base n'a aucune raison d'avoir une réponse impulsionnelle réelle. En pratique on fait tout les calculs sur les enveloppes complexes à partir des filtres équivalents en bande de base. Si on veut revenir au signal $y(t)$ il suffit alorsd'écrire :

$$y(t) = \text{Ré}\left\{y_b(t)e^{2j\pi f_0 t}\right\}$$

1.7.3 Démodulation synchrone

Considérons le schéma représenté figure 1.9. Les signaux $x(t)$ et $y(t)$ ont des spectres nuls en dehors d'une bande de largeur de fréquence B localisée autour de f_0. Le signal $x(t) \times y(t)$ est filtré par un filtre passe-bas de largeur B qui élimine les composantes autour de la fréquence $2f_0$. En pratique, si $B \ll f_0$, cette opération de filtrage est simple. On note $x_b(t)$ et $y_b(t)$ les enveloppes complexes respectives de $x(t)$ et $y(t)$ par rapport à f_0. On peut alors écrire :

$$\begin{aligned} 2x(t)y(t) &= 2(p_x(t)\cos(2\pi f_0 t) - q_x(t)\sin(2\pi f_0 t)) \\ &\quad \times(p_y(t)\cos(2\pi f_0 t) - q_y(t)\sin(2\pi f_0 t)) \\ &= p_x(t)p_y(t) + q_x(t)q_y(t) + \text{“}2f_0''\end{aligned}$$

où le terme noté "$2f_0$" indique toutes les composantes, telles que $p_x(t)p_y(t)\cos(4\pi f_0 t)$, situées autour de la fréquence $2f_0$. Notons ici que les signaux de la forme $p_x(t)p_y(t)$ sont de bande $2B$. En effet la transformée de Fourier de $p_x(t)p_y(t)$ a pour expression $P_x(f) \star P_y(f)$ et la convolution entre deux fonctions à supports bornés a pour support la somme des supports.

En supposant à présent que le filtre passe-bas annule ou, pour le moins, rende négligeables les composantes de $2x(t)y(t)$ autour de la fréquence $2f_0$, il s'en suit que le signal en sortie a pour expression :

$$s(t) = \frac{1}{2}\,\text{Ré}\left\{x_b(t)y_b^*(t)\right\} \tag{1.21}$$

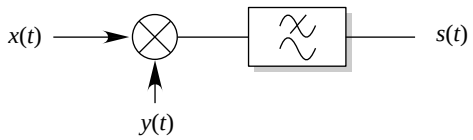

FIG. 1.9 – *Détecteur comparateur.*

Appliquons ce résultat aux cas suivants :

Détection synchrone : Si $y(t) = 2\cos(2\pi f_0 t)$ alors $y_b(t) = 2$ et par conséquent $s(t) = p_x(t)$. De même en prenant $y(t) = -2\sin(2\pi f_0 t)$ on obtient $s(t) = q_x(t)$. Nous avons représenté figure 1.10 le schéma d'un détecteur synchrone. Il récupère sur la voie dite en phase, associée à $\cos(2\pi f_0 t)$, la partie réelle de l'enveloppe complexe par rapport à f_0 et sur la voie dite en quadrature, associée à $\sin(2\pi f_0 t)$, la partie imaginaire de l'enveloppe complexe par rapport à f_0. Il est à la base des dispositifs de réception en communications numériques sur fréquence porteuse. Comme le montre l'item suivant, il est important que les deux oscillateurs soient respectivement en phase et en quadrature avec f_0.

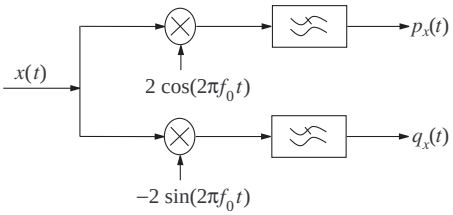

FIG. 1.10 – *Détecteur synchrone. La voie où l'oscillateur local est $2\cos(2\pi f_0 t)$ s'appelle la voie en phase et celle où l'oscillateur local est $-2\sin(2\pi f_0 t)$ s'appelle la voie en quadrature.*

Détection "désynchronisée" : supposons que, sur la voie en phase, on ait $y(t) = 2\cos(2\pi f_0 t + \phi)$ au lieu de $y(t) = 2\cos(2\pi f_0 t)$. On a alors $y_b(t) = 2\exp(j\phi)$ et donc :

$$s(t) = p_x(t)\cos(\phi) + q_x(t)\sin(\phi) \tag{1.22}$$

Le signal sur la voie dite en phase contient une part non négligeable de la composante en quadrature : on dit qu'il y a *diaphonie*.

Comparateur de phase : $x(t) = A_x\cos(2\pi f_0 t + \phi_x(t))$ et $y(t) = A_y\sin(2\pi f_0 t + \phi_y(t))$ auxquels correspondent respectivement $x_b(t) = A_x\exp(j\phi_x(t))$ et $y_b(t) = -jA_y\exp(j\phi_y(t))$. Par conséquent on a :

$$s(t) = \frac{1}{2}A_x A_y \sin(\phi_y(t) - \phi_x(t))$$

1.7.4 Enveloppe complexe d'un processus aléatoire du second ordre

Processus aléatoire stationnaire au second ordre

Un processus aléatoire $x(t)$ est dit *stationnaire au second ordre* si sa moyenne $m_x = \mathbb{E}\{x(t)\}$ et sa fonction d'autocovariance $R_{xx}(\tau) = \mathbb{E}\{(x(t+\tau) - m_x)(x(t) - m_x)\}$ sont indépendantes de t. Sa densité spectrale de puissance $S_{xx}(f)$ est la transformée de Fourier de $R_{xx}(\tau)$, ce qui s'écrit :

$$R_{xx}(\tau) = \int_{\mathbb{R}} S_{xx}(f)e^{2j\pi f\tau}df$$

On montre que $S_{xx}(f)$ est positive. On rappelle que la puissance est donnée par $\mathbb{E}\left\{|x(t)|^2\right\} = \int_{\mathbb{R}} S_{xx}(f)df + |m_x|^2 = R_{xx}(0) + |m_x|^2$.

Si $x(t)$ est mis à l'entrée d'un filtre linéaire stable de gain en fréquence $H(f)$, alors la sortie $y(t)$ est un processus aléatoire stationnaire au second ordre, de moyenne $m_y = m_x H(0)$ et de densité spectrale de puissance :

$$S_{yy}(f) = |H(f)|^2 S_{xx}(f)$$

Composantes en phase et quadrature d'un processus stationnaire au second ordre

Considérons un processus aléatoire $x(t)$ *réel, centré*, stationnaire au second ordre de d.s.p. $S_{xx}(f)$. Rappelons que $S_{xx}(f)$ est positive et paire. Choisissons arbitrairement une fréquence f_0. Si la d.s.p. de $x(t)$ est localisée dans une bande B on prendra souvent pour f_0 la fréquence située au milieu de la bande d'occupation spectrale. Partant du signal analytique $z_x(t)$, obtenu par filtrage de $x(t)$ par le filtre de gain complexe $2\mathbb{1}_{(0,+\infty)}(f)$, puis de l'enveloppe complexe $x_b(t) = z_x(t)e^{-2j\pi f_0 t}$, on a :

$$x(t) = \text{Ré}\left\{x_b(t)\exp(2j\pi f_0 t)\right\}$$

En posant $x_b(t) = p(t) + jq(t)$, on obtient :

$$x(t) = p(t)\cos(2\pi f_0 t) - q(t)\sin(2\pi f_0 t)$$

Alors les processus $x_b(t)$, $p(t)$ et $q(t)$ ont les propriétés suivantes :

1. $x_b(t)$ est un processus aléatoire (complexe) *stationnaire au second ordre* appelé *enveloppe complexe* de $x(t)$ (sous-entendu par rapport à la fréquence f_0). $x_b(t)$ est centré et vérifie, pour tout couple d'instants (t_1, t_2) :
$$\mathbb{E}\left\{x_b(t_1)x_b(t_2)\right\} = 0 \tag{1.23}$$
2. $p(t)$ et $q(t)$ sont deux processus aléatoires, centrés, *stationnaires au second ordre* et d'*intercorrélation stationnaire*. Ils sont appelés respectivement *composante en phase et composante en quadrature* de $x(t)$ (sous-entendu par rapport à la fréquence f_0).
3. on note $R_{pp}(\tau) = \mathbb{E}\left\{p(t+\tau)p(t)\right\}$ et $R_{qq}(\tau) = \mathbb{E}\left\{q(t+\tau)q(t)\right\}$. Alors pour tout τ, $R_{pp}(\tau) = R_{qq}(\tau)$,
4. on note $R_{pq}(\tau) = \mathbb{E}\left\{p(t+\tau)q(t)\right\}$. Alors pour tout τ, $R_{pq}(\tau) = -R_{pq}(-\tau)$. On en déduit que, pour tout t, $\mathbb{E}\left\{p(t)q(t)\right\} = 0$.
5. $\mathbb{E}\left\{x^2(t)\right\} = \mathbb{E}\left\{p^2(t)\right\} = \mathbb{E}\left\{q^2(t)\right\} = \frac{1}{2}\mathbb{E}\left\{|x_b(t)|^2\right\}$,
6. $x_b(t)$ a comme d.s.p. :
$$S_{x_b}(f) = 4S^+_{xx}(f + f_0) \tag{1.24}$$
où $S^+_{xx}(f)$ désigne la partie de $S_{xx}(f)$ situées dans les fréquences positives. On retiendra que l'on passe du spectre du signal à celui de son enveloppe complexe (1) en translatant de $-f_0$ la partie de $S_{xx}(f)$ située dans les fréquences positives et (2) en multipliant par 4. Inversement pour passer du spectre $S_{x_b}(f)$ de l'enveloppe complexe au spectre $S_{xx}(f)$ du signal, (1) on translate de $+f_0$ la fonction positive $S_{x_b}(f)$, (2) on complète par symétrie paire autour de $-f_0$ puis (3) on divise par 4, ce qui s'écrit :
$$S_{xx}(f) = \frac{1}{4}\left(S_{x_b}(f - f_0) + S_{x_b}(-f - f_0)\right) \tag{1.25}$$
7. $p(t)$ et $q(t)$ ont la même densité spectrale de puissance qui a pour expression :
$$S_{pp}(f) = S_{qq}(f) = S^-_{xx}(f - f_0) + S^+_{xx}(f + f_0) \tag{1.26}$$
La construction s'obtient graphiquement à partir de $S_{xx}(f)$ comme cela est représenté figure1.11 : on translate de $(+f_0)$ la partie de $S_{xx}(f)$ situées dans les fréquences négatives, de $(-f_0)$ la partie de $S_{xx}(f)$ situées dans les fréquences positives, puis on effectue la somme (notons ici que toutes ces fonctions sont non seulement réelles mais positives).

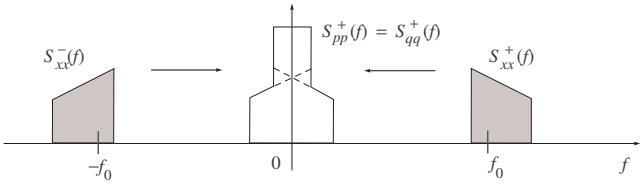

FIG. 1.11 – *Construction des d.s.p. des composantes en phase et en quadrature.*

8. Si on note respectivement $S_{pq}(f)$ et $S_{qp}(f)$ les transformées de Fourier de $R_{pq}(\tau)$ et de $R_{qp}(\tau)$, on a :

$$S_{pq}(f) = S_{qp}(-f) = \frac{1}{4j}\left(S_{x_b}(-f) - S_{x_b}(f)\right) = \frac{1}{j}\left(S^{-}_{xx}(f-f_0) - S^{+}_{xx}(f+f_0)\right) \tag{1.27}$$

   Le procédé graphique est le suivant : on translate de $(+f_0)$ la partie de $S_{xx}(f)$ situées dans les fréquences négatives, de $(-f_0)$ la partie de $S_{xx}(f)$ situées dans les fréquences positives, puis on effectue la différence et on divise par j.

9. D'après (1.27), si la partie de $S_{xx}(f)$ située dans les fréquences positives est symétrique autour de la fréquence f_0 alors pour tout couple (t_1, t_2), $\mathbb{E}\{p(t_1)q(t_2)\} = 0$, et donc $p(t_1)$ et $q(t_2)$ sont décorrélés pour tout couple d'instants t_1, t_2,

10. Si $x(t)$ est gaussien, $p(t)$ et $q(t)$ sont gaussiens (rappelons que, dans le cas gaussien, la décorrélation entraîne aussi l'indépendance).

11. Pour deux processus réels $x(t)$ et $y(t)$:

$$\mathbb{E}\{x(t)y(t)\} = \frac{1}{2}\,\mathrm{Ré}\left\{\mathbb{E}\{x_b(t)y_b^*(t)\}\right\}$$

   Il suffit de poser $v(t) = x(t) + y(t)$, de noter que $v_b(t) = x_b(t) + y_b(t)$ et d'appliquer le point 5 à $v(t)$.

Remarque En général deux p.a. SSL ne sont pas les composantes en phase et quadrature d'un p.a. SSL. Il faut en plus qu'ils vérifient l'équation (1.23). Considérons, en effet, le processus $x(t) = p(t)\cos(2\pi f_0 t) - q(t)\sin(2\pi f_0 t)$ où $p(t)$ et $q(t)$ sont deux p.a. SSL, centrés et d'intercovariance stationnaire. On a $\mathbb{E}\{x(t)\} = 0$. Calculons la fonction d'autocovariance de $x(t)$. Il vient :

$$\begin{aligned}
&\mathbb{E}\{x(t+\tau)x(t)\} \\
&\quad = R_{pp}(\tau)\cos(2\pi f_0(t+\tau))\cos(2\pi f_0 t) + R_{qq}(\tau)\sin(2\pi f_0(t+\tau))\sin(2\pi f_0 t) \\
&\quad -R_{pq}(\tau)\cos(2\pi f_0(t+\tau))\sin(2\pi f_0 t) - R_{pq}(-\tau)\cos(2\pi f_0 t)\sin(2\pi f_0(t+\tau)) \\
&\quad = \frac{R_{pp}(\tau)+R_{qq}(\tau)}{2}\cos(2\pi f_0\tau) + \frac{R_{pp}(\tau)-R_{qq}(\tau)}{2}\cos(2\pi f_0(2t+\tau)) \\
&\quad + \frac{R_{pq}(\tau)-R_{pq}(-\tau)}{2}\sin(2\pi f_0\tau) + \frac{R_{pq}(\tau)+R_{pq}(-\tau)}{2}\sin(2\pi f_0(2t+\tau))
\end{aligned}$$

Pour que $x(t)$ soit stationnaire, à savoir $\mathbb{E}\{x(t+\tau)x(t)\}$ indépendant de t, il faut et il suffit que $p(t)$ et $q(t)$ vérifient *à la fois* $R_{pp}(\tau) - R_{qq}(\tau) = 0$ et $R_{pq}(\tau) + R_{pq}(-\tau) = 0$. Si on pose $x_b(t) = p(t) + jq(t)$, ces deux conditions sont équivalentes à l'unique condition suivante portant sur $x_b(t)$:

$$\mathbb{E}\{x_b(t+\tau)x_b(t)\} = \mathbb{E}\{(p(t+\tau) + jq(t+\tau))(p(t) + jq(t))\} = 0$$

Exemple 1.3 (Composantes d'un bruit blanc rectangulaire) *On considère un filtre réel passe-bande idéal, de gain en fréquence $H(f) = 1$ dans les bandes de largeur B centrées autour de f_0 et de $-f_0$. Le signal à l'entrée est un bruit blanc, centré, de densité spectrale de puissance $N_0/2$. On note $b(t)$ le signal en sortie et $p(t)$ et $q(t)$ ses composantes en phase et quadrature.*

1. *Déterminer l'expression de la densité spectrale de puissance de $b(t)$.*
2. *On note respectivement $p(t)$, $q(t)$ et $x_b(t) = p(t) + jq(t)$ la composante en phase, la composante en quadrature et l'enveloppe complexe de $b(t)$ par rapport à f_0. Déterminer les expressions des densités spectrales de puissance et des puissances de $p(t)$, $q(t)$ et $x_b(t)$.*
3. *Même question en considérant les composantes en phase et en quadrature de $b(t)$ par rapport à $f_1 = f_0 - B$.*

Réponse :

1. La densité spectrale de puissance de $b(t)$ a pour expression :

$$S_{bb}(f) = \frac{N_0}{2}\,|H(f)|^2 = \frac{N_0}{2} \quad \text{si} \quad f_0 - B/2 < |f| < f_0 + B/2$$

2. D'après les résultats de la section 1.7.4, les composantes en phase et quadrature par rapport à f_0 ont donc pour densité spectrale :

$$S_{pp}(f) = S_{qq}(f) = N_0 \mathbb{1}(f \in (-B/2, B/2))$$

On en déduit que $\mathbb{E}\left\{p^2(t)\right\} = \mathbb{E}\left\{q^2(t)\right\} = N_0 B$.
Pour l'enveloppe complexe $x_b(t) = p(t) + jq(t)$ on a, d'après la propriété 9 de la section 1.7.4, $\mathbb{E}\{x_b(t+\tau)x^*(t)\} = \mathbb{E}\{p(t+\tau)p(t)\} + \mathbb{E}\{q(t+\tau)q(t)\}$ où on a utilisé que $p(t_2)$ et $q(t_2)$ sont non corrélées. On en déduit que :

$$S_{x_b}(f) = 2N_0 \mathbb{1}(f \in (-B/2, B/2))$$

et que $\mathbb{E}\left\{|x_b(t)|^2\right\} = 2N_0 B$.

3. Les composantes en phase et quadrature de $b(t)$ *par rapport à* $(f_0 - B/2)$ ont pour densité spectrale :

$$S_{pp}(f) = S_{qq}(f) = \frac{N_0}{2} \mathbb{1}(f \in (-B, B))$$

et on a $\mathbb{E}\left\{p^2(t)\right\} = \mathbb{E}\left\{q^2(t)\right\} = N_0 B$. Pour l'enveloppe complexe $x_b(t) = p(t) + jq(t)$, on note tout d'abord que la puissance est donnée par $\mathbb{E}\left\{|x_b(t)|^2\right\} = \mathbb{E}\{(p(t) + jq(t))(p(t) - jq(t))\} = 2N_0 B + 0 + 0$ car $p(t)$ et $q(t)$ ne sont pas corrélées. L'expression de la densité spectrale de puissance de $x_b(t)$ est donnée par (1.27) :

$$S_{pq}(f) = S_{qp}(-f) = \frac{1}{j}(\mathbb{1}(f \in (-B, 0)) - \mathbb{1}(f \in (0, B))$$

On remarque que $p(t_1)$ et $q(t_2)$ sont à présent corrélés pour $t_1 \neq t_2$. L'expression du spectre de $x_b(t) = p(t) + jq(t)$ s'obtient à partir de l'expression (1.24) qui donne :

$$S_{x_b}(f) = 2N_0 \mathbb{1}(f \in (0, B))$$

On en déduit que $\mathbb{E}\left\{|x_b(t)|^2\right\} = 2N_0 B$.

1.8 Eléments de décision statistique

1.8.1 Position du problème

Dans un problème de décision statistique, l'observation Y est modélisée par une variable aléatoire, à valeurs dans un espace $\mathcal{Y}$, typiquement $\mathbb{R}^K$, $\mathbb{C}^K$, $\{0,1\}^K$. Les M états possibles, sur lesquels portent la décision, sont modélisés par une variable aléatoire X à valeurs dans un ensemble $\mathcal{X} = \{x_1, x_2, \cdots, x_M\}$ de cardinalité finie M. On suppose que la loi de probabilité de X est connue. Typiquement si X est uniforme sur $\mathcal{X}$ alors on a, pour tout $x \in \mathcal{X}$, $\mathbb{P}\{X = x\} = 1/M$. On suppose enfin que la loi de probabilité de Y conditionnellement à X est connue. Si la loi de Y est discrète, elle est caractérisée par la donnée des probabilités de la forme $\mathbb{P}\{Y = y | X = x\}$ quand y décrit $\mathcal{Y}$. Dans le cas contraire, nous supposerons que la loi de Y possède une densité de probabilité que nous notons $p_{Y|X=x}(y)$.

A partir de l'observation Y, on veut inférer sur la valeur de X. On ne connaît pas X mais, en revanche, on connaît sa distribution de probabilité pour toute valeur de Y. On peut donc utiliser cette connaissance pour construire une "bonne" fonction de décision. On appelle *fonction de décision* une application de $\mathcal{Y}$ dans $\mathcal{X}$ de la forme :

$$g : y \in \mathcal{Y} \mapsto x_m \in \mathcal{X} = \{x_1, x_2, \cdots, x_M\}$$

La donnée de cette fonction est équivalente à la donnée d'une partition de l'ensemble $\mathcal{Y}$ en M domaines $\{\Lambda_1, \cdots, \Lambda_M\}$ tels que pour $y \in \Lambda_m$ alors $g(y) = x_m$. Dans le contexte de la décision statistique, le *singleton* $\{x_m\}$ est appelé une *hypothèse* et le problème ainsi posé est aussi appelé un *test d'hypothèses*.

La détermination de la fonction g ou ce qui est équivalent de la M-partition de $\mathcal{Y}$ se fait en optimisant un critère pertinent pour le problème considéré. Dans le cas de la réception optimale d'un signal numérique en présence de bruit, le critère choisi est la *maximisation de la probabilité de décision correcte* dont l'expression est :

$$P_c = \sum_{m=1}^{M} \mathbb{P}\{X = x_m\} \underbrace{\mathbb{P}\{\Lambda_m | X = x_m\}}_{P_{c|m}} \tag{1.28}$$

De façon équivalente on peut minimiser la probabilité d'erreur $P_e = 1 - P_c$ qui s'écrit :

$$P_e = \sum_{m=1}^{M} \mathbb{P}\{X = x_m\} \mathbb{P}\left\{\overline{\Lambda}_m | X = x_m\right\} \tag{1.29}$$

où $\overline{\Lambda}_m$ désigne l'ensemble complémentaire de Λ_m dans le domaine d'observation $\mathcal{Y}$. Si la loi de probabilité de Y est discrète, on a :

$$P_{c|m} = \mathbb{P}\{Y \in \Lambda_m | X = x_m\} = \sum_{y \in \Lambda_m} \mathbb{P}\{Y = y | X = x_m\}$$

qui donne :

$$P_c = \sum_{m=1}^{M} \sum_{y \in \Lambda_m} \mathbb{P}\{X = x_m\} \mathbb{P}\{Y = y | X = x_m\}$$

Si la loi possède une densité de probabilité, on a :

$$P_{c|m} = \mathbb{P}\{Y \in \Lambda_m | X = x_m\} = \int_{\Lambda_m} p_{Y|X=x_m}(y) dy$$

qui donne :

$$P_c = \sum_{m=1}^{M} \int_{\Lambda_m} \mathbb{P}\{X = x_m\} p_{Y|X=x_m}(y) dy$$

Dans les deux cas, la maximisation de P_c est simple : il suffit de mettre dans la région Λ_m les points $y \in \mathcal{Y}$ qui maximisent l'intégrande. Explicitons la solution suivant que la loi de Y est discrète ou possède une densité :

1. Si la loi de Y est discrète, les régions de décision optimales sont données par :

$$\begin{aligned} \Lambda_m = \{y \in \mathcal{Y} : \mathbb{P}\{X = x_m\} \mathbb{P}\{Y = y | X = x_m\} \\ > \mathbb{P}\{X = x_\ell\} \mathbb{P}\{Y = y | X = x_\ell\} \ \ \forall \ell \neq m\} \end{aligned} \tag{1.30}$$

   Dans le cas particulier où $\mathbb{P}\{X = x_m\} = 1/M$, on obtient comme règle de décision la règle dite du *maximum de vraisemblance* où les régions de décision ou régions de vraisemblance maximale sont définies par :

$$\Lambda_m = \{y \in \mathcal{Y} : \mathbb{P}\{Y = y | X = x_m\} > \mathbb{P}\{Y = y | X = x_\ell\} \ \ \forall \ell \neq m\} \tag{1.31}$$

   En fait la famille $\{\Lambda_m\}$, décrite par (1.30) et (1.31), ne forme pas à proprement parler une partition de $\mathcal{Y}$. Pour avoir une partition, il suffit de ranger les points d'une frontière dans l'une quelconque des régions adjacentes à la frontière. On vérifie alors aisément que cela ne change pas la valeur de la probabilité P_c.

   La règle de décision peut encore s'écrire :

$$\hat{x}_m = \arg \max_{x_m \in \mathcal{X}} \mathbb{P}\{Y = y | X = x_m\} \tag{1.32}$$

2. Si la loi de Y possède une densité, les régions de décision sont données par :

$$\begin{aligned} \Lambda_m = \{y \in \mathcal{Y} : \mathbb{P}\{X = x_m\} p_{Y|X=x_m}(y) \\ > \mathbb{P}\{X =_\ell\} p_{Y|X=x_\ell}(y) \ \ \forall \ell \neq m\} \end{aligned} \tag{1.33}$$

   Dans le cas où $\mathbb{P}\{X = x_m\} = 1/M$, on obtient la règle dite du *maximum de vraisemblance* caractérisée par les régions de décision :

$$\Lambda_m = \{y \in \mathcal{Y} : p_{Y|X=x_m}(y) > p_{Y|X=x_\ell}(y) \ \ \forall \ell \neq m\} \tag{1.34}$$

   La règle de décision peut encore s'écrire :

$$\hat{x}_m = \arg \max_{x_m \in \mathcal{X}} p_{Y|X=x_m}(y) \tag{1.35}$$

1.8.2 Exemple de deux observations gaussiennes de dimension K

Le cas où l'observation est conditionnellement gaussienne est fondamental : il représente, en communications, la situation où le bruit sur le canal est AGB.

Commençons par le cas le plus simple où la dimension de l'observation est $K = 1$.

Exemple 1.4 (Test à 2 hypothèses, observation gaussienne scalaire) *On a $Y = X + W$ où X est à valeurs dans $\{x_1, x_2\}$ supposés équiprobables. Sans perte de généralités, on peut poser que $x_2 > x_1$. On peut aussi poser le problème sous la forme du test à deux hypothèses suivant :*

$$\begin{cases} H_1 : & Y = x_1 + W \\ H_2 : & Y = x_2 + W \end{cases}$$

où x_1 et x_2 sont réelles et où W est une variable aléatoire gaussienne, centrée de variance σ^2. On pose $\rho = (x_2 - x_1)/2\sigma$.

1. *Ecrire la loi de Y conditionnellement aux deux hypothèses.*
2. *En déduire la règle de décision optimale.*
3. *Déterminer en fonction de ρ, l'expression de la probabilité d'erreur.*

Exemple corrigé 1 *1. De l'hypothèse d'indépendance de X et W, on déduit que l'observation a pour densité de probabilité conditionnellement à X :*

$$\begin{aligned} &- \textit{ sous } H_1 : \ p_{Y|X=x_1}(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(y-x_1)^2}{2\sigma^2}\right) \\ &- \textit{ sous } H_2 : \ p_{Y|X=x_2}(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(y-x_2)^2}{2\sigma^2}\right) \end{aligned}$$

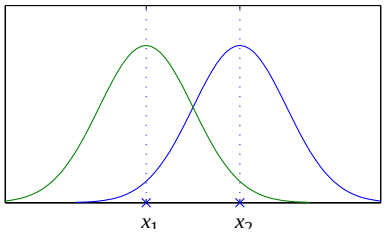

FIG. 1.12 – *Densités de probabilité sous les deux hypothèses*

2. *La région de décision associée à l'hypothèse H_1 et qui minimise la probabilité d'erreur, est donnée par* (1.34) *qui s'écrit :*

$$\Lambda_1 = \{y \in \mathbb{R} : p_{Y|X=x_1}(y) \geq p_{Y|X=x_2}(y)\}$$

En remplaçant $p_{Y|X=x_1}(y)$ et $p_{Y|X=x_2}(y)$ par leurs expressions et en simplifiant on obtient :

$$\Lambda_1 = \{y \in \mathbb{R} : y \leq (x_1 + x_2)/2\}, \ \textit{ et } \ \Lambda_2 = \mathbb{R} - \Lambda_1$$

L'organe de décision compare simplement l'observation au seuil $s = (x_1 + x_2)/2$: si $Y < s$, il décide l'hypothèse H_1 et sinon il décide l'hypothèse H_2.

3. *Comme les deux hypothèses sont supposées de probabilité $1/2$, la probabilité moyenne d'erreur est donnée par :*

$$P_e = \frac{1}{2}\int_{\Lambda_1} p_{Y|X=x_2}(y)dy + \frac{1}{2}\int_{\Lambda_2} p_{Y|X=x_1}(y)dy$$

Par raison de symétrie on a aussi :

$$P_e = \int_s^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(y-x_1)^2}{2\sigma^2}\right) dy$$

En posant $t = (y - x_1)/\sigma$, *il vient :*

$$P_e = \int_{(x_2-x_1)/2\sigma}^{+\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt = Q(\rho)$$

où on a posé $\rho = (x_2 - x_1)/2\sigma$ *et :*

$$Q(x) = \int_x^{+\infty} \frac{1}{\sqrt{2\pi}}\, e^{-t^2/2} dt \tag{1.36}$$

On note que $(x_2 - x_1)$ *représente la distance euclidienne entre les deux points représentatifs des deux hypothèses.*

Nous avons représenté figure 1.13 la forme de $Q(x)$ en fonction de x exprimé en dB (l'abscisse représente $20\log_{10}(x)$). On voit que $Q(x)$ est monotone décroissante. Par conséquent, dans l'exemple 1.4, plus le terme ρ est grand, plus le système discriminera facilement les deux hypothèses. Ce résultat, qui est très général, est satisfaisant puisque ρ s'interprète comme un rapport signal sur bruit.

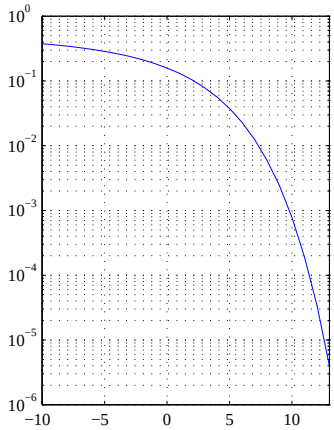

FIG. 1.13 – *Fonction* $Q(x)$ *définie par la formule* (1.36). *L'abscisse exprimée en dB représente* $20\log_{10}(x)$.

Dans la littérature on utilise aussi, pour déterminer les probabilités d'erreur sur les canaux gaussiens, la fonction :

$$\mathrm{erfc}(x) = \int_x^{+\infty} \frac{2}{\sqrt{\pi}}\, e^{-t^2} dt \tag{1.37}$$

Ces deux fonctions sont reliées par :

$$Q(x) = \frac{1}{2}\,\mathrm{erfc}\left(\frac{x}{\sqrt{2}}\right) \Leftrightarrow \mathrm{erfc}(x) = 2Q(x\sqrt{2}) \tag{1.38}$$

Pour $x \gg 1$, la valeur approchée suivante de $Q(x)$ peut être utile :

$$Q(x) \approx \frac{1}{2}\, e^{-x^2/2} \tag{1.39}$$

$M = 2$ et K est quelconque

On considère à présent, le cas où les 2 signaux sont de dimension K et où les deux observations possibles sont gaussiennes. Ce calcul généralise celui de l'exemple 1.4. La densité de probabilité de la loi d'observation s'écrit pour $m = 1, 2$:

$$p_{Y|X=x_m}(y) = \frac{1}{(2\pi\sigma^2)^{K/2}} \exp\left(-\frac{1}{2\sigma^2}\, d_E^2(y, x_m)\right) \tag{1.40}$$

On suppose que X prend, de façon équiprobable, soit la valeur x_1 soit la valeur x_2, où x_1 et x_2 appartiennent à $\mathbb{R}^K$.

La région de vraisemblance maximale associée à l'hypothèse 1 est l'ensemble $\Lambda_1 = \{y \in \mathbb{R}^K : d_E(y, x_1) \leq d_E(y, x_2)\}$ et celle associée à l'hypothèse 2 est $\Lambda_2 = \mathbb{R}^K - \Lambda_1$. Il s'ensuit que la frontière qui sépare les deux régions, est l'hyperplan médiateur du segment x_1x_2.

Règle de décision

La règle de décision revient à trouver x_m qui minimise $d_E(y, x_m)$. En utilisant l'expression (**??**), on obtient :

$$d_E^2(y, x_m) = \|y\|^2 + \|x_m\|^2 - 2\langle y, x_m \rangle$$

Comme le premier terme ne dépend pas de m, il peut être omis dans la comparaison et la règle de décision se réduit à trouver x_m qui maximise :

$$L(y, x_m) = \langle y, x_m \rangle - \frac{1}{2}\|x_m\|^2 \tag{1.41}$$

L'expression (1.41) montre que la règle de décision optimale s'exprime *linéairement* en fonction de l'observation. Ce résultat général est une conséquence directe du caractère gaussien du bruit additif. Le second terme $\|x_m\|^2$ représente l'énergie du signal. Si les deux signaux sont d'égale énergie, il peut alors être omis dans la comparaison.

Expression de la probabilité d'erreur

Par raison de symétrie et du fait que les deux hypothèses sont équiprobables, la probabilité d'erreur s'écrit :

$$\begin{aligned} P_e &= \int_{\Lambda_2} p_{Y|X=x_1}(y)dy \\ &= \underbrace{\int \cdots \int}_{\Lambda_2} \frac{1}{(2\pi\sigma^2)^{K/2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{k=1}^{K}(y_k - x_{1,k})^2\right) dy_1 \cdots dy_K \end{aligned}$$

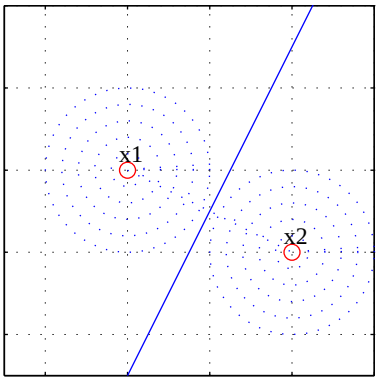

FIG. 1.14 – *Régions de décision pour $K = 2$. La médiatrice du segment x_1x_2 en trait plein est la séparatrice des deux régions de vraisemblance maximale. Les cercles en pointillé représentent les courbes d'égales densités de probabilité.*

Effectuons le changement de variables qui consiste à rapporter l'espace au repère, caractérisé par la droite qui porte x_1x_2 et par le hyperplan-médiateur du segment x_1x_2 (voir figure 1.14 où nous avons représenté le cas où $K = 2$). Notons u la variable le long de x_1x_2. Après simplification on peut écrire :

$$\begin{aligned} P_e &= \mathbb{P}\{U > 0 | X = x_1\} = \int_0^{+\infty} \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(u+d/2)^2}{2\sigma^2}} du \\ &= Q\left(\frac{d_E(x_1, x_2)}{2\sigma}\right) \end{aligned} \tag{1.42}$$

où la fonction $Q(x)$ est donnée par (1.36). L'expression (1.42) montre que les performances dépendent du rapport $d_E(x_1, x_2)/2\sigma$ qui s'interprète comme un rapport signal sur bruit. Plus la distance euclidienne entre les deux points x_1 et x_2 est grande par rapport à σ, plus faible est P_e.

Dans le cas où les deux signaux sont de même énergie $E = \|x_1\|^2 = \|x_2\|^2$, l'expression de la distance est :

$$d_E^2(x_1, x_2) = 2E(1 - \cos(\theta)) = 4E \sin^2(\theta/2) \tag{1.43}$$

où θ désigne l'angle entre les deux vecteurs-signaux défini par :

$$\cos(\theta) = \frac{\langle x_1, x_2 \rangle}{\|x_1\| \, \|x_2\|} \tag{1.44}$$

On en déduit que :

$$P_e = Q\left(\sqrt{\frac{E}{\sigma^2} \sin^2(\theta/2)}\right) \tag{1.45}$$

dont on tire pour des *signaux orthogonaux* , où $\theta = \pi/2$:

$$P_e = Q\left(\sqrt{\frac{E}{2\sigma^2}}\right) \tag{1.46}$$

et pour des *signaux antipodaux*, où $\theta = \pi$:

$$P_e = Q\left(\sqrt{\frac{E}{\sigma^2}}\right) \tag{1.47}$$

Dans tous les cas, l'énergie moyenne est $E = 0.5(\|x_1\|^2 + \|x_2\|^2)$. On peut alors se demander quelle est la disposition géométrique qui minimise E pour une distance $d_E(x_1, x_2)$ fixée, c'est-à-dire pour un niveau fixé de la probabilité d'erreur, ce qui revient à résoudre :

$$\min_{\|x_1 - x_2\|^2 = \text{constante}} \left(\|x_1\|^2 + \|x_2\|^2\right)$$

En notant que $\|x_1\|^2 + \|x_2\|^2 = \|x_1 - x_2\|^2 + 2\langle x_1, x_2 \rangle$, on déduit que le minimum est obtenu pour $x_1 = -x_2$. On dit dans ce cas que les signaux sont *antipodaux*.

Pour une même probabilité d'erreur, il faut une énergie moyenne deux fois plus grande pour des signaux orthogonaux que pour des signaux antipodaux. Cela correspond à une différence de 3 dB en faveur des signaux antipodaux. De façon générale, on peut mesurer l'écart par rapport au cas des deux signaux antipodaux par la valeur :

$$\rho_{dB} = 10 \log_{10}(\sin^2(\theta/2)) \tag{1.48}$$

1.8.3 Cas de M observations gaussiennes de dimension K

Les résultats précédents se généralisent sans difficulté à un test portant sur M signaux appartenant à un espace de dimension K. La densité de probabilité de la loi d'observation s'écrit pour $m = 1, \cdots, M$:

$$p_{Y|X=x_m}(y) = \frac{1}{(2\pi\sigma^2)^{K/2}} \exp\left(-\frac{1}{2\sigma^2} \, d_E^2(y, x_m)\right) \tag{1.49}$$

où x_m appartient à $\mathbb{R}^K$.

La région de vraisemblance maximale associée à l'hypothèse m est l'ensemble $\Lambda_m = \{y \in \mathbb{R}^K : d_E(y, x_m) \leq d_E(y, x_\ell), \ \forall \ell \neq m\}$. Et donc les frontières qui séparent les régions sont construites à partir des plans-médiateur des segments $x_m x_\ell$.

La règle de décision revient à trouver x_m qui minimise $d_E(y, x_m)$. En utilisant l'expression (**??**), on obtient :

$$d_E^2(y, x_m) = \|y\|^2 + \|x_m\|^2 - 2\langle y, x_m \rangle$$

Comme le premier terme ne dépend pas de m, il peut être omis dans la comparaison et la règle de décision se réduit à trouver x_m qui maximise :

$$L(y, x_m) = \langle y, x_m \rangle - \frac{1}{2}\|x_m\|^2 \tag{1.50}$$

L'expression (1.50) montre que la règle de décision optimale a une *expression linéaire* en fonction de l'observation.

La détermination de l'expression de la probabilité d'erreur est en général compliquée. On se contente souvent de bornes supérieures ou inférieures ou tout simplement de simulations.

De façon très générale, la probabilité d'erreur, dans le cas AGB, dépend essentiellement des interdistances entre les points-signaux dans l'espace d'observation tandis que l'énergie moyenne dépend de la position de l'origine par rapport à ces mêmes points. Intuitivement pour minimiser l'énergie moyenne pour des interdistances fixées, il faut placer l'origine au "centre" de l'ensemble constitué des points représentatifs des signaux dans l'espace des signaux.

1.9 Statistique suffisante sur un canal soumis à un bruit AGB

Les résultats obtenus paragraphe 1.8.3, s'appliquent au problème de la détection d'un signal dans un bruit. On considère l'observation $Y(t) = X(t) + W(t)$ où $W(t)$ est un bruit AGB de densité spectrale $N_0/2$ et où $X(t)$ est un processus aléatoire à valeurs dans l'ensemble fini de M signaux déterministes $\{x_1(t), \cdots, x_M(t)\}$. On suppose que $X(t)$ et $W(t)$ sont indépendants.

L'espace engendré $\mathcal{S}$ par $\{x_1(t), \cdots, x_M(t)\}$ est appelé l'*espace des signaux*. Il est de dimension K avec $K \leq M$.

Le problème de la détection du signal $X(t)$ dans le bruit $W(t)$, revient à trouver une fonction qui associe une valeur du paramètre $m \in \{1, \cdots, M\}$ à l'observation $Y(t)$, tout en minimisant la probabilité d'erreur. Le résultat suivant que nous énonçons sans démonstration est fondamental : si le bruit $W(t)$ est AGB, alors la projection orthogonale de l'observation $Y(t)$ dans l'espace $\mathcal{S}$ est une *statistique suffisante* pour le paramètre m. En anglais on parle de *sufficient statistic*. Une statistique suffisante est aussi appelée *statistique exhaustive*. Sans en entrer dans le détail de la démonstration, ce résultat utilise le fait que les composantes du bruit $W(t)$ hors de $\mathcal{S}$, ne sont pas corrélées aux composantes dans $\mathcal{S}$ et, étant gaussiennes, sont statistiquement indépendantes les unes des autres. Il s'ensuit que la loi de l'observation conditionnellement à sa projection orthogonale dans $\mathcal{S}$ ne dépend plus du paramètre m, ce qui est la définition d'une statistique suffisante.

On montre en statistique [**?**] que pour construire un "bon" estimateur, il suffit de ne s'intéresser qu'aux fonctions d'une statistique suffisante. De façon plus imagée, les composantes du bruit $W(t)$ hors de $\mathcal{S}$ ne contiennent aucune information utile sur le paramètre m.

Par conséquent pour prendre, à partir de l'observation $Y(t)$, une décision optimale sur le paramètre m, il suffit de ne considérer que la projection de $Y(t)$ sur $\mathcal{S}$ qui est de *dimension finie*. Ce résultat est fondamental puisqu'il permet de remplacer l'objet très "compliqué" qu'est le processus aléatoire $Y(t)$ par une suite *finie* de K variables aléatoires. Nous pouvons à présent déterminer les lois de probabilité de ces K variables aléatoires conditionnellement à m.

1.9.1 Représentation réelle des signaux

$Y(t) = X(t) + W(t)$ où $W(t)$ est un bruit AGB de densité spectrale $N_0/2$. $X(t)$ est un processus aléatoire à valeurs dans l'ensemble fini des M signaux déterministes $\{x_1(t), \cdots, x_M(t)\}$. On suppose que $X(t)$ et $W(t)$ sont indépendants. On note $\mathcal{S}$ l'espace engendré par $\{x_1(t), \cdots, x_M(t)\}$ et K sa dimension, $K \leq M$.

Partant de la propriété **??**, la loi conditionnelle des composantes dans $\mathcal{S}$ de $Y(t)$, sachant $X(t) = x_m(t)$, se détermine en remplaçant le signal aléatoire $X(t)$ par le signal déterministe $x_m(t)$.

Notons $\{\phi_k(t)\}$ une base orthonormée de $\mathcal{S}$, $x_{m,k}$ les composantes de $x_m(t)$ par rapport à cette base et $\mathbf{x}_m = [x_{m,1}, \cdots, x_{m,K}]^T$ le vecteur de $\mathbb{R}^K$ qui lui est associé. De même, on note $W_k = \int W(t)\phi_k^*(t)dt$ les projections de $W(t)$ sur cette base et $\mathbf{W} = [W_1, \cdots, W_K]^T$ le vecteur aléatoire associé. Enfin on note $\mathbf{Y}$ le vecteur de composantes $Y_k = \int Y(t)\phi_k^*(t)dt$. Dans ces conditions, $\mathbf{Y} = \mathbf{x}_m + \mathbf{W}$ et $\mathbb{E}\{\mathbf{Y}\} = \mathbf{x}_m$ puisque $\mathbb{E}\{\mathbf{W}\} = 0$. En utilisant (**??**), on déduit que la matrice de covariance de $\mathbf{Y}$ a pour expression $\mathbb{E}\left\{(\mathbf{Y} - \mathbf{x}_m)(\mathbf{Y} - \mathbf{x}_m)^T\right\} = 0.5\, N_0 \mathbf{I}_K$ où $\mathbf{I}_k$ est la matrice identité de dimension K.

Du fait que le caractère gaussien se conserve par transformation linéaire, les composantes $\{W_k\}$ forment un vecteur gaussien. Il s'ensuit que la densité de probabilité de la loi de $\mathbf{Y}$ conditionnellement à $X(t) = x_m(t)$

a pour expression :

$$\begin{aligned} p_{\mathbf{Y}|m}(y_1,\cdots,y_K) &= \frac{1}{(\pi N_0)^{K/2}}\exp\left(-\frac{1}{N_0}\sum_{k=1}^{K}(y_k-x_{m,k})^2\right) \\ &= \frac{1}{(\pi N_0)^{K/2}}\exp\left(-\frac{d_E^2(\mathbf{y},\mathbf{x}_m)}{N_0}\right) \end{aligned} \tag{1.51}$$

où la distance euclidienne entre $\mathbf{x}$ et $\mathbf{y}$, éléments de $\mathbb{R}^K$, est donnée d'après (**??**) par les expressions suivantes, qui sont équivalentes :

$$\begin{aligned} d_E^2(\mathbf{y},\mathbf{x}) &= (\mathbf{x}-\mathbf{y})^T(\mathbf{x}-\mathbf{y}) = \sum_{k=1}^{K}(y_k-x_k)^2 \\ &= \int_{-\infty}^{+\infty}|y(t)-x(t)|^2dt = \|x\|^2+\|y\|^2-2\langle x,y\rangle \end{aligned} \tag{1.52}$$

On peut alors appliquer les résultats précédents. En portant (**??**) et (**??**) dans (1.50), on obtient la règle de décision optimale. Elle consiste à trouver l'indice m qui rend maximum l'expression :

$$L(y,x_m) = \int_{-\infty}^{+\infty}Y(t)x_m(t)dt - \frac{1}{2}\int_{-\infty}^{+\infty}|x_m(t)|^2dt \tag{1.53}$$

Dans le cas où $M=2$ et où les deux hypothèses sont équiprobables, la probabilité d'erreur est donnée par (1.42). En utilisant (1.52) et en supposant que les signaux sont de même énergie E, on obtient $d_E^2(x_1,x_2) = 2E - 2\langle x_1,x_2\rangle = 2E(1-\cos(\theta))$ où θ est défini par (1.44). En utilisant (**??**), à savoir $\sigma^2 = N_0/2$, et en remplaçant dans (1.42), on obtient :

$$P_e = Q\left(\sqrt{\frac{d_E^2(x_1,x_2)}{4\sigma^2}}\right) = Q\left(\sqrt{\frac{2E\sin^2(\theta/2)}{N_0}}\right) \tag{1.54}$$

Pour M quelconque, l'expression de la probabilité d'erreur n'a pas une expression simple même si la disposition géométrique des signaux dans $\mathcal{S}$ est simple.

1.9.2 Représentation en phase et quadrature

On considère à nouveau un signal *réel* $Y(t) = X(t) + W(t)$ mais on s'intéresse à présent à la loi conditionnelle associée à sa représentation *en phase et quadrature* (par rapport à une fréquence arbitraire f_0). On suppose que $W(t)$ est un bruit additif, gaussien, blanc dans une bande grande devant celle du processus $X(t)$. On suppose que $X(t)$ est une processus aléatoire à valeurs dans $\{x_1(t),\cdots,x_M(t)\}$ et que $X(t)$ et $W(t)$ sont indépendants.

Partant de la propriété **??** il suffit, pour déterminer la loi de l'observation conditionnellement à chacun des M signaux possibles, de fixer $X(t) = x_m(t)$ où $m \in \{1,\cdots,M\}$.

Le signal réel $x_m(t)$ est représenté au moyen de sa composante en phase $x_{bm}^r(t)$ et de sa composante en quadrature $x_{bm}^i(t)$. On peut donc écrire que $x_m(t) = x_{bm}^r(t)\cos(2\pi f_0 t) - x_{bm}^r(t)\sin(2\pi f_0 t)$. On suppose que les M signaux en phase $x_{bm}^r(t)$ (respectivement les M signaux en quadrature $x_{bm}^i(t)$) engendrent un espace de dimension K où $K \leq M$. Rapportés respectivement à une base de chacun de ces espaces, les signaux $x_m(t)$ sont décrits par leurs composantes $(x_{bm,1}^r,\cdots,x_{bm,K}^r)$ et $(x_{bm,1}^i,\cdots,x_{bm,K}^i)$. Une autre façon est de dire que $x_m(t)$ appartient à un espace de dimension $2K$ qui peut être décrit par le vecteur à $2K$ composantes $\mathbf{x}_{bm} = (x_{bm,1}^r, x_{bm,1}^i,\cdots,x_{bm,K}^r,x_{bm,K}^i)^T$ obtenu en concaténant les composantes de $x_{bm}^r(t)$ et celles de $x_{bm}^i(t)$. Avec des notations évidentes on a :

$$\begin{aligned} \underbrace{(Y_{b,1}^r,Y_{b,1}^i,\cdots,Y_{b,K}^r,Y_{b,K}^i)}_{\mathbf{Y}_b} &= \underbrace{(x_{bm,1}^r,x_{bm,1}^i,\cdots,x_{bm,K}^r,x_{bm,K}^i)}_{\mathbf{x}_{bm}}\cdots \\ &\quad + \underbrace{(W_{b,1}^r,W_{b,1}^i,\cdots,W_{b,K}^r,W_{b,K}^i)}_{\mathbf{W}_b} \end{aligned} \tag{1.55}$$

Si $W(t)$ est un processus gaussien, les composantes de $W(t)$ sont elles-mêmes, par linéarité, des variables aléatoires conjointement gaussiennes. Leur moyenne est nulle. Pour avoir leur loi, il reste à déterminer leurs

covariances. Un calcul analogue à celui effectué pour établir l'expression (**??**) montre que, pour un bruit blanc de bande très grande devant celle des signaux, on a :

$$\mathbb{E}\{W_{bk}^r W_{b\ell}^r\} = \mathbb{E}\{W_{bk}^i W_{b\ell}^i\} = N_0\delta_{k,\ell} \quad (1.56)$$

$$\mathbb{E}\{W_{bk}^r W_{b\ell}^i\} = 0 \quad (1.57)$$

où nous avons utilisé d'une part l'expression (**??**), donnant la densité spectrale des composantes en phase et quadrature de $W(t)$ qui est égale à N_0, et d'autre part le fait que ces composantes ne sont pas corrélées (voir item 9 des propriétés **??**).

Ainsi, partant de l'expression (**??**), on obtient pour la densité de probabilité du vecteur gaussien $\mathbf{Y}_b$ conditionnellement à $X(t) = x_m(t)$ l'expression :

$$p_{\mathbf{Y}_b|m}(\mathbf{y}_b) = \frac{1}{(2\pi N_0)^K}\ \exp\left(-\frac{d_E^2(\mathbf{y}_b, \mathbf{x}_{bm})}{2N_0}\right) \quad (1.58)$$

où $\mathbf{y}_b = (y_{b,1}^r, y_{b,1}^i, \cdots, y_{b,K}^r, y_{b,K}^i)$ et où $d_E^2(\mathbf{y}_b, \mathbf{x}_b) = \int |y_b(t) - x_b(t)|^2 dt$ est donné par (1.59). Rappelons qu'on peut exprimer la distance et le produit scalaire de deux éléments x_b et y_b, appartenant à cet espace de dimension $2K$, à partir de leurs enveloppes complexes respectives. En effet en utilisant (**??**), on obtient pour la distance les expressions suivantes qui sont équivalentes :

$$\begin{aligned} d_E^2(x_b, y_b) &= \sum_{k=1}^{K} (x_{b,k}^r - y_{b,k}^r)^2 + (x_{b,k}^i - y_{b,k}^i)^2 \\ &= \int_{-\infty}^{+\infty} |x_b^r(t) - y_b^r(t)|^2 dt + \int_{-\infty}^{+\infty} |x_b^i(t) - y_b^i(t)|^2 dt \\ &= \int_{-\infty}^{+\infty} |x_b(t) - y_b(t)|^2 dt = 2\int_{-\infty}^{+\infty} |x(t) - y(t)|^2 dt \end{aligned} \quad (1.59)$$

et pour le produit scalaire, les expressions :

$$\begin{aligned} \langle x_b, y_b\rangle &= \sum_{k=1}^{K} \left(x_{b,k}^r y_{b,k}^r + x_{b,k}^i y_{b,k}^i\right) = \sum_{k=1}^{K} \text{Ré}\left\{(x_{b,k}^r + jx_{b,k}^i)(y_{b,k}^r - jy_{b,k}^i)\right\} \\ &= \text{Ré}\left\{\int_{-\infty}^{+\infty} x_b(t) y_b^*(t) dt\right\} = 2\int_{-\infty}^{+\infty} x(t) y(t) dt \end{aligned} \quad (1.60)$$

On peut à nouveau utiliser les résultats du paragraphe 1.8.2. En portant (1.59) et (1.60) dans (1.50), on obtient la règle de décision optimale qui consiste à trouver l'indice m qui rend maximum l'expression :

$$L(y_b, x_{bm}) = \text{Ré}\left\{\int Y_b(t) x_{bm}^*(t) dt\right\} - \frac{1}{2}\int |x_{bm}(t)|^2 dt \quad (1.61)$$

Dans le cas où $M = 2$ et où les deux hypothèses sont équiprobables, l'expression de la probabilité d'erreur est fournie par (1.42). On obtient :

$$P_e = Q\left(\sqrt{\frac{d_E^2(x_{b1}, x_{b2})}{4\sigma^2}}\right) \quad (1.62)$$

où :

$$\begin{aligned} d_E^2(x_{b1}, x_{b2}) &= \int_{-\infty}^{+\infty} |x_{b1}(t) - x_{b2}(t)|^2 dt = 2\int_{-\infty}^{+\infty} |x_1(t) - x_2(t)|^2 dt \\ &= 2(\|x_1\|^2 + \|x_2\|^2 - 2\langle x_1, x_2\rangle) \end{aligned}$$

Ce qui donne pour des signaux de même énergie $E = \int |x_m(t)|^2 dt$, $d_E^2(x_{b1}, x_{b2}) = 4E(1 - \cos(\theta))$ où θ est défini par (1.44). En utilisant (1.56) à savoir $\sigma^2 = N_0$, on a :

$$P_e = Q\left(\sqrt{\frac{2E\sin^2(\theta/2)}{N_0}}\right) \quad (1.63)$$

Evidemment, les formules (1.54) et (1.63) sont identiques : ce sont les mêmes signaux que l'on utilise ! Ce qui diffère est la façon de faire les calculs *sans* ou *avec* la représentation en phase et quadrature.

Exemple 1.5 *On utilise les notations précédentes. On considère* $Y(t) = X(t) + W(t)$. $W(t)$ *est un bruit gaussien, de densité spectrale* $N_0/2$. $X(t)$ *est un processus aléatoire à valeurs dans* $\{x_m(t) = \alpha_m \cos(2\pi f_0 t + \phi_m), m \in \{1,2\}\}$ *où* $\alpha_m \in \mathbb{R}^+$ *et* $\phi_m \in (0, 2\pi)$. *On suppose que les deux signaux possibles sont de même probabilité* $1/2$. *On suppose que* $W(t)$ *et* $X(t)$ *sont indépendants.*

1. *Déterminer, en fonction de* α_m *et* ϕ_m, *la décomposition en phase et en quadrature de* $x_m(t)$ *par rapport à* f_0.
2. *Déterminer l'espace des signaux de la décomposition en phase et en quadrature de* $x_m(t)$ *ainsi que les composantes dans une base orthonormée.*
3. *Déterminer l'énergie de* $x_{bm}(t)$. *En déduire l'énergie moyenne de* $X(t)$.
4. *Déterminer la distance entre* $x_{b1}(t)$ *et* $x_{b2}(t)$.
5. *Donner les propriétés statistiques des projections de* $W(t)$ *sur la base précédente.*
6. *En déduire la règle optimale de décision.*
7. *En déduire l'expression de la probabilité d'erreur.*
8. *Appliquer les résultats précédents lorsque* $\alpha_m = \alpha$ *(indépendant de* m*).*

Exemple corrigé 2

1. *L'enveloppe complexe par rapport à* f_0 *s'écrit* $x_{bm}(t) = \alpha_m e^{j\phi_m} = a_m + jb_m$ *où* $a_m = \alpha_m \cos(\phi_m)$ *et* $b_m = \alpha_m \sin(\phi_m)$.
2. *En concaténant les composantes en phase et en quadrature, on obtient le couple* (a_m, b_m) *où* $m = 1, 2$. *Par conséquent, l'espace des signaux est* $\mathbb{R}^2$. *En le rapportant à la base orthonormée* $(1,0)$ *et* $(0,1)$, *on obtient comme points représentatifs des deux signaux, les points de coordonnées* (a_1, b_1) *et* (a_2, b_2).
3. $E_{bm} = a_m^2 + b_m^2 = \alpha_m^2$. *Si les deux signaux sont équiprobables, l'énergie moyenne de l'enveloppe complexe est* $E_b = 0.5(\alpha_1^2 + \alpha_2^2)$. *D'après* (**??**) *l'énergie moyenne du signal transmis est* $E = E_b/2$.
4. $d_E^2(x_{b1}, x_{b2}) = (a_1 - a_2)^2 + (b_1 - b_2)^2$.
5. *Dans la base précédente, les composantes en phase et quadrature de* $W(t)$ *sont deux variables gaussiennes centrées, indépendantes et de même variance* N_0.
6. *On note respectivement* $P_Y(t) = Ré\{Y_b(t)\}$ *et* $Q_Y(t) = Imag(Y_b(t))$ *les composantes en phase et en quadrature de* $Y(t)$ *par rapport à* f_0. *En pratique elles s'obtiennent, à partir de l'observation* $Y(t)$, *par le système schématisé figure 1.10. En portant* $x_{bm}(t) = a_m + jb_m$ *dans* (1.61), *on obtient la règle optimale de décision qui consiste à choisir la valeur de* m *qui maximise :*

$$L(Y_b, m) = a_m \int_{-\infty}^{+\infty} P_Y(t)dt + b_m \int_{-\infty}^{+\infty} Q_Y(t)dt - \frac{1}{2}\,\alpha_m^2$$

7. *D'après* (1.62) *et en utilisant le résultat de la question 4, on a :*

$$P_e = Q\left(\sqrt{\frac{(a_1 - a_2)^2 + (b_1 - b_2)^2}{4N_0}}\right)$$

8. *Si* $\alpha_m = \alpha$ *indépendant de* m, *la règle optimale de décision se simplifie et s'écrit :*

$$L(Y_b, m) = \cos(\phi_m) \underbrace{\int_{-\infty}^{+\infty} P_Y(t)dt}_{u_r} + \sin(\phi_m) \underbrace{\int_{-\infty}^{+\infty} Q_Y(t)dt}_{u_i}$$

Après quelques simplifications, cela revient à comparer $\mathrm{arctang}(\mathrm{u_i}/\mathrm{u_r})$ *à* $(\phi_1 + \phi_2)/2$. $d_E^2(x_{b1}, x_{b2}) = 4\alpha^2 \sin^2(\theta/2)$ *où* $\theta = \phi_1 - \phi_2$. *Si* $\alpha_m = \alpha$, *l'énergie moyenne du signal transmis est* $E = \alpha^2/2$. *Et donc* $d_E^2(x_{b1}, x_{b2}) = 8E \sin^2(\theta/2)$. *En utilisant* $\sigma^2 = N_0$, *on retrouve l'expression* (1.54) *:*

$$P_e = Q\left(\sqrt{\frac{2E \sin^2(\theta/2)}{N_0}}\right)$$

1.9.3 Résumé sur la détection d'un signal dans un bruit AGB

On retiendra que :

– si on utilise la représentation réelle des signaux observés, alors la règle de décision du maximum de vraisemblance consiste à déterminer la valeur de m qui maximise la fonction $L(y,m)$ définie par :

$$L(y,m) = \int_{-\infty}^{+\infty} Y(t)x_m(t)dt - \frac{1}{2}\int_{-\infty}^{+\infty} x_m^2(t)dt \tag{1.64}$$

– si on utilise la représentation en enveloppe complexe par rapport à une fréquence f_0 choisie arbitrairement, alors la règle de décision du maximum de vraisemblance consiste à déterminer la valeur de m qui maximise la fonction $L_b(y,m)$ définie par :

$$L_b(y,m) = \text{Ré}\left\{\int_{-\infty}^{+\infty} Y_b(t)x_{bm}^*(t)dt\right\} - \frac{1}{2}\int_{-\infty}^{+\infty} |x_{bm}(t)|^2 dt \tag{1.65}$$

Nous avons déjà souligné que, grâce aux caractères additif, gaussien et blanc, la règle de décision optimale avait une expression linéaire en fonction de l'observation.

La détermination de l'expression de la probabilité d'erreur est en général compliquée. On se contente souvent de bornes supérieures ou inférieures ou tout simplement de simulations.

1.10 Exercices

Exercice 1.1 *Soit le signal $x(t) = m(t)\cos(2\pi f_0 t) - n(t)\sin(2\pi f_0 t)$ où $m(t)$ et $n(t)$ désignent deux signaux réels, dont les spectres sont nuls en dehors de la bande de fréquence $(-B, B)$. On suppose d'autre part que $f_0 > B$.*

1. *Déterminer l'expression du signal analytique associé à $x(t)$.*
2. *Déterminer l'expression de l'enveloppe complexe du signal $x(t)$ par rapport à la fréquence f_0.*

Exercice 1.2 (Retard de phase, retard de groupe) *Soit $x(t) = a(t)\cos(2\pi f_0 t)$, un signal à bande étroite, de largeur de bande B. Ce signal traverse un canal passe-bande, dont le gain complexe est de module constant $K > 0$ et à phase $\Phi(f)$ dans la bande de fréquence occupée de $x(t)$. On approche la phase $\Phi(f)$ par un développement limité au premier ordre. Montrer que le signal en sortie du canal peut s'écrire :*

$$y(t) = Ka(t - t_g)\cos(2\pi f_0(t - t_\phi))$$

où t_g et t_ϕ désignent deux quantités, appelées respectivement retard de groupe et retard de phase et définies par :

$$t_g = -\frac{1}{2\pi}\left.\frac{d\Phi(f)}{df}\right|_{f=f_0} \quad et \quad t_\phi = -\frac{1}{2\pi}\left.\frac{\Phi(f)}{f}\right|_{f=f_0}$$

Exercice 1.3 (Mesure de puissance avec un détecteur d'enveloppe) *Soit $x(t)$ un processus aléatoire réel, stationnaire au second ordre, centré. Ce signal est appliqué à l'entrée d'un système qui détermine l'enveloppe $a(t)$ (on rappelle que la notion d'enveloppe n'est pas liée au choix d'une fréquence) puis estime la puissance par $P = (\mathbb{E}\{a(t)\})^2/2$.*

Déterminer les expressions de $P_x = \mathbb{E}\left\{x^2(t)\right\}$ et de P dans les trois cas suivants :

1. *$x(t) = A\cos(2\pi f_0 t + \Phi)$, où Φ désigne une variable aléatoire de loi uniforme sur $(0, 2\pi)$ et A une constante positive.*
2. *$x(t) = A_1\cos(2\pi f_1 t + \Phi_1) + A_2\cos(2\pi f_2 t + \Phi_2)$, où Φ_1 et Φ_2 désignent deux variables aléatoires, indépendantes et de loi uniforme sur $(0, 2\pi)$ et A_1 et A_2 deux constantes positives.*
3. *$x(t)$ est gaussien, centré, de puissance $P_x = \mathbb{E}\left\{x(t)^2\right\}$.*

Indication : si U et V sont deux variables aléatoires gaussiennes, indépendantes, centrées de même variance σ^2, alors $R = \sqrt{U^2 + V^2}$ suit une loi dite de Raylegh de densité :

$$p_R(r) = \frac{r}{\sigma^2}\exp\left(-r^2/2\sigma^2\right)\mathbb{1}_{(0,+\infty)}(r)$$

Chapitre 2

Communications numériques

2.1 Introduction

En communications numériques, la source émet un message discret : on entend par là que le message prend ses valeurs dans un ensemble dénombrable, le plus souvent fini, de valeurs. Typiquement une suite de K bits ne peut prendre que 2^K valeurs. Afin d'être transmise, la suite des données d'information émise par la source est associée, par le modulateur, à un signal qui subit à travers le canal des perturbations. Dans notre présentation, nous supposons que le canal agit comme un filtre idéal de bande en fréquence *limitée* et est le siège d'un bruit additif, gaussien, blanc. A la réception le but du destinataire est de retrouver avec le minimum d'erreurs la suite des symboles émis et *non pas* le signal transmis.

2.2 Modulation numérique

2.2.1 Message numérique et signal numérique

- Nous supposons que le *message numérique*, produit par la source, est une suite de variables aléatoires $\{d_k\}$ à valeurs dans l'alphabet $\{0, 1\}$, indépendantes et identiquement distribuées (i.i.d.) suivant une loi uniforme, ce qui s'écrit $\mathbb{P}\{d_k = 1\} = \mathbb{P}\{d_k = 0\} = 1/2$.
- Le modulateur associe, de façon bijective, à chaque message numérique un signal numérique à temps continu $x(t)$.

L'hypothèse que la source peut être considérée comme binaire, i.i.d. et uniforme, trouve sa justification en théorie de l'information, plus précisément avec le théorème de codage de source, dont nous verrons un énoncé dans le chapitre 4. Limitons nous à en donner ici une justification intuitive. Si on représente les caractères d'un texte, écrit en français, par un code comportant 8 bits (comme le fait le code ASCII), ce qui permet de coder au total 256 caractères, la suite binaire obtenue ne sera pas uniforme. Cela tient au fait que les probabilités d'apparition des différents caractères ne sont pas égales. On sait, par exemple, que les caractères /e/, /s/, /a/ sont beaucoup plus fréquents que les caractères /w/ ou /y/. On a donc intérêt à coder de façon plus courte les caractères les plus fréquents. De façon plus précise, C. Shannon a montré que, sous des hypothèses larges, il existe un codage qui associe, à des suites de n caractères, des mots-code d'autant plus courts que ces suites sont plus probables et tel que la suite de 0 et de 1 obtenue est, quand n tend vers l'infini, asymptotiquement i.i.d. et uniforme.

Comme nous l'avons dit l'opération de *modulation* consiste à associer, de façon bijective, à la suite des symboles $\{d_k\}$ un signal $x(t)$. Parmi toutes les façons de procéder, l'exemple suivant est fondamental : le signal numérique associé au message numérique constitué d'une suite de K bits a pour expression :

$$x(t) = \sum_{k=1}^{K} a_k h(t - kT_b)$$

où $h(t - kT_b)$ représente une même impulsion $h(t)$ décalée de kT_b et modulée par le symbole :

$$a_k = 2d_k - 1$$

où a_k prend ses valeurs dans $\{-1, +1\}$. Notons qu'on ne suppose pas que $h(t)$ soit de durée inférieure à T_b ni même de durée finie. En temps réel, le temps T_b désigne l'intervalle de temps entre deux bits émis. Il lui

correspond un débit binaire mesuré en *bits/s* donné par :

$$D_b = \frac{1}{T_b} \text{ en } bits/s$$

La modulation M-aire représente une généralisation de cet exemple.

2.2.2 Transmission M-aire en bande de base

On considère un alphabet $\mathcal{A}$ *fini* à $M = 2^m$ symboles. Typiquement on prend une suite *centrée et dont les amplitudes régulièrement espacées*. Par conséquent cette suite, à une constante multiplicative près, est la suite des nombres impairs :

$$\mathcal{A} = \{-(M-1), -(M-3), \cdots, -1, +1, \cdots, +(M-3), +(M-1)\}$$

On voit que ce choix traduit simplement le fait que (1) les écarts d'amplitudes entre 2 symboles voisins sont les mêmes (on ne privilégie aucun symbole, sauf peut-être les deux extrêmes) et que (2), sous l'hypothèse que les symboles sont équiprobables, l'ensemble $\mathcal{A}$ doit être centré de façon à minimiser l'énergie moyenne.

On choisit ensuite un "codage" qui associe, de façon bijective, à toute suite de m bits du message numérique un symbole a_k de l'alphabet $\mathcal{A}$. Partant de la suite d_k, le modulateur fournit le signal numérique dit en *bande de base*[1] :

$$x(t) = \sum_k a_k h(t - kT) \tag{2.1}$$

On parle alors de Modulation par Impulsion en Amplitude (en abrégé MIA, en anglais PAM pour Pulse Amplitude Modulation).

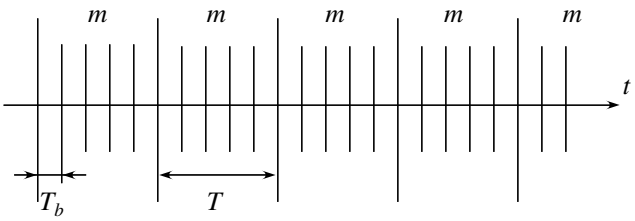

FIG. 2.1 – *Période bit et période symbole*

A présent $T = mT_b = \log_2(M)T_b$. On en déduit que la vitesse de modulation ou *débit symbole*, qui se mesure en *bauds*, a pour expression en fonction du débit binaire et de M :

$$R = \frac{1}{T} = \frac{D_b}{\log_2(M)} \text{ en } bauds$$

C'est précisément pour réduire R, en augmentant la taille M de $\mathcal{A}$, que l'on a introduit les modulations M-aires. La raison est que l'on espère réduire ainsi l'occupation en fréquence du signal numérique et donc de pouvoir transmettre sur un canal de bande plus étroite.

Codage de Gray

Parmi tous les codages possibles, qui associent m bits de la source à une amplitude de l'ensemble $\mathcal{A}$, le codage de Gray est tel que deux symboles d'amplitudes adjacentes sont associés à deux suites de m bits qui ne varient que d'*un seul bit*. Ce choix minimise, comme nous le verrons plus loin, le taux d'erreur par élément binaire (TEEB). Le tableau 2.1 indique un codage de Gray dans le cas où $M = 8$.

2.2.3 Modulation numérique sur fréquence porteuse

La modulation sur porteuse a pour but d'engendrer un signal passe-bande, c'est-à-dire un signal *réel* dont le spectre se situe autour de la fréquence f_0 et de largeur $B < f_0$. Comme nous l'avons expliqué, on a l'habitude, dans le domaine des communications, de représenter ces signaux par leur enveloppe complexe ou

[1] Dans la littérature on utilise aussi l'expression *code en ligne* pour désigner les modulations numériques en bande de base.

message	symbole
100	-7
101	-5
111	-3
110	-1
010	$+1$
011	$+3$
001	$+5$
000	$+7$

TAB. 2.1 – *Codage de Gray*

encore par leurs composantes en phase et en quadrature par rapport à la fréquence f_0. On peut alors écrire pour le signal modulé :

$$x(t) = x_p(t)\cos(2\pi f_0 t) - x_q(t)\sin(2\pi f_0 t)$$

L'enveloppe complexe de $x(t)$ par rapport à f_0 est alors donnée par :

$$x_b(t) = x_p(t) + jx_q(t)$$

et on a :

$$x(t) = \text{Ré}\left\{x_b(t)e^{2j\pi f_0 t}\right\} \tag{2.2}$$

En admettant que le signal $x(t)$ est stationnaire au second ordre de spectre (ou d.s.p.) $S_x(f)$, nous avons vu page 17 équation (1.25) que le spectre $S_x(f)$ du signal $x(t)$ se déduit de $S_{x_b}(f)$ par la relation :

$$S_x(f) = \frac{1}{4}\left(S_{x_b}(f - f_0) + S_{x_b}(-f - f_0)\right)$$

Graphiquement, $S_x(f)$ s'obtient par translation de $S_{x_b}(f)$ de $+f_0$ suivie d'une symétrisation puis d'une division par 4. Attention rappelons ici que, contrairement à $S_x(f)$, $S_{x_b}(f)$ est une fonction *positive mais pas nécessairement paire.*

MDP-M

La modulation par déplacement de phase (en abrégé MDP-M , en anglais PSK pour *Phase Shift Keying*)à M états consiste à émettre pendant le temps T de durée d'un symbole, une impulsion sinusoïdale, dont la phase prend l'une des M valeurs de l'alphabet :

$$\Phi = \left\{0, \frac{2\pi}{M}, \cdots, \frac{2\pi(M-1)}{M}\right\}$$

Ainsi au k-ième symbole est associé, dans l'intervalle de temps $(kT, (k+1)T)$ le signal :

$$x(t) = A\cos(2\pi f_0 t + \phi_k)$$

où $\phi_k \in \Phi$. Il s'en suit que l'enveloppe complexe de $x(t)$ a pour expression :

$$x_b(t) = A\sum_k a_k \text{rect}_T(t - kT) \text{ avec } a_k = \exp(j\phi_k) \tag{2.3}$$

Les valeurs de l'alphabet Φ peuvent être représentées par des points régulièrement répartis sur un cercle de rayon unité. La figure ainsi obtenue s'appelle une *constellation.* Si $M = 2^m$, à chaque point de la constellation est associée une suite de m bits. Nous avons représenté figure 2.2 le cas $M = 4$. Notons que le choix de la constellation est tel qu'à deux symboles voisins sont associés deux mots de code qui ne varient que par un bit : on retrouve un codage de Gray.

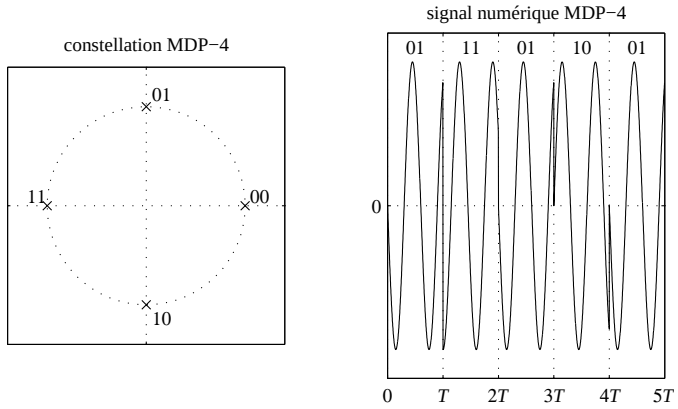

FIG. 2.2 – *Modulation MDP-4*

MAQ

L'expression (2.3) de la modulation MDP-M se généralise sans difficulté en prenant pour l'enveloppe complexe du signal numérique :

$$x_b(t) = A \sum_k a_k h(t - kT) \tag{2.4}$$

où a_k est, à présent, une suite de valeurs prises dans un alphabet complexe de M valeurs et $h(t)$ une impulsion de forme quelconque. Evidemment on a :

$$x(t) = \text{Ré}\left\{x_b(t)e^{2j\pi f_0 t}\right\} = x_b^r(t)\cos(2\pi f_0 t) - x_b^i(t)\sin(2\pi f_0 t)$$

où $x_b(t) = x_b^r(t) + jx_b^i(t)$. D'où le nom de Modulation par Amplitude en Quadrature (en abrégé MAQ, en anglais QAM pour Quadrature Amplitude Modulation). La figure 2.3 représente la constellation et un exemple de signal d'une modulation MAQ-16.

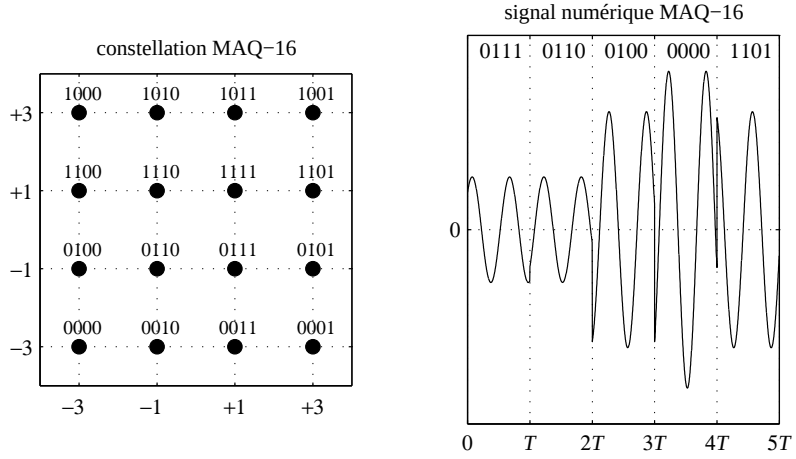

FIG. 2.3 – *Modulation MAQ-16*

Remarque : on verra que les propriétés spectrales et les performances en présence de bruit sont essentiellement déterminées par la forme de l'expression (2.4) qui est la même que l'expression (2.1) donnée pour la modulation en bande de base. On peut donc traiter simultanément, dans un même formalisme, les deux types de modulation en travaillant uniquement sur l'expression $A\sum_k a_k h(t - kT)$ où les symboles a_k sont complexes, voire réels si la modulation est en bande de base. Toutefois, par souci pédagogique, nous considérerons tout d'abord le cas des modulations en bande de base et nous généraliserons ensuite les résultats au cas des modulations sur fréquence porteuse.

2.2.4 Limite fondamentale : formule de Shannon

Le concepteur du modulateur est conduit à considérer des alphabets de taille $M > 2$ de façon, pense-t-il, à réduire l'occupation spectrale de $x(t)$. On voit en effet que, pour un débit binaire fixé, le fait d'augmenter M, augmente T et, par conséquent, le signal $x(t)$ contient a priori moins de hautes fréquences. Toutefois, pour un rapport signal/bruit donné, cela va rapprocher les niveaux significatifs des symboles adjacents et donc provoquer, pour ce type de modulation, une augmentation de la probabilité d'erreur P_e.

Théoriquement le concepteur dispose d'une bande en fréquence B et d'un rapport signal sur bruit (RSB) et il cherche à transmettre le débit binaire maximum avec une probabilité d'erreur la plus faible possible.

Pour une bande en fréquence B donnée, la rapidité $R = 1/T$ est par conséquent "fixée" et est de l'ordre de B. Si on veut alors, sans augmenter le RSB, diminuer la probabilité d'erreur P_e, il faudra diminuer M et, comme $R = D_b/\log_2(M)$ est "fixée" par le choix de B, il faudra réduire le débit D_b.

On a longtemps cru que ce raisonnement était juste et que le seul moyen de réduire la probabilité d'erreur pour une bande de fréquence B et un RSB donnés, était de réduire le débit. Les résultats, obtenus en 1948 par C. Shannon, ont montré que le bruit ne constituait pas une limite aux transmissions sûres, mais une *limite au débit*. En voici un énoncé (nous y reviendrons plus en détails au chapitre 4) :

Théorème 2.1 (Canal gaussien - Shannon (1948)) *Soit un canal de bande en fréquence B soumis à un bruit additif gaussien blanc et soit un rapport signal sur bruit RSB. On appelle capacité du canal gaussien la quantité mesurée en bits/s et définie par :*

$$C = B\log_2(1 + RSB) \text{ (bits/s)}$$

Alors, si le débit binaire de la source $D_b < C$, il existe un ensemble (modulateur/démodulateur) asymptotiquement sans erreur.

Exemple : sur le canal téléphonique $B \approx 3000$ Hz $(300 - 3400$ Hz). Pour un RSB de 30 dB, $C = 30000$ bits/s.

2.2.5 Paramètres

- L'efficacité spectrale exprimée en *bits/s/Hz* est définie par :
$$\eta = \frac{D_b}{B} \text{ (bits/s/Hz)}$$
où D_b désigne le débit binaire et B la bande de fréquence du canal.
- Le rapport signal sur bruit défini par :
$$\rho = \frac{E_b}{N_0}$$
où E_b désigne la quantité d'énergie par bit, exprimée en nombre de Joules par bit, et $N_0/2$ la densité spectrale du bruit additif, blanc sur le canal, exprimée en W/Hz. On en déduit que la puissance moyenne du signal est donnée par $P_s = E_b D_b$ et que la puissance du bruit dans la bande B et donnée par $P_b = N_0 B$. On en déduit le rapport signal sur bruit en puissance :
$$\frac{P_s}{P_b} = \frac{E_b}{N_0}\eta = \rho\eta$$
- La probabilité d'erreur par symbole définie par $P_e = \mathbb{P}\{\hat{a}_k \neq a_k\}$ où $\hat{a}_k$ désigne la valeur choisie par le récepteur et a_k le symbole émis.
- On considère aussi le taux d'erreur par éléments binaire (TEEB). Dans le cas où le rapport signal sur bruit est grand, nous verrons qu'une expression approchée du TEEB est :
$$\text{TEEB} \approx \frac{P_e}{\log_2(M)}$$

Ce qui est remarquable est que, au-dessous d'une certaine valeur du débit, et donc de l'efficacité, il est possible de rendre P_e aussi faible que l'on veut. Ainsi, sur le canal gaussien sans mémoire, la courbe :

$$\eta = \log_2(1 + \rho\eta) \text{ (bits/s/Hz)} \Leftrightarrow \rho = \frac{2^\eta - 1}{\eta} \tag{2.5}$$

donne une limite fondamentale aux transmissions sûres. Nous avons représenté figure 2.4 la courbe donnant ρ en dB en fonction de η en bits/s/Hz. Pour les points situés au dessus de la courbe, il existe un système de communication dont la probabilité d'erreurs peut être rendue aussi faible que l'on veut.

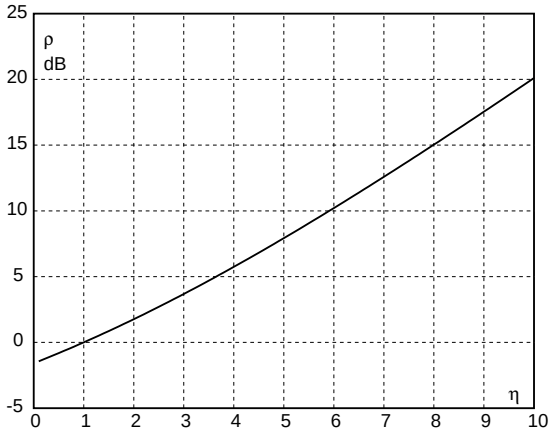

FIG. 2.4 – *Limite fondamentale de transmission sur le canal AGB*

2.2.6 Spectre des signaux numériques

Dans ce paragraphe, nous donnons les propriétés du second ordre (moyenne et densité spectrale de puissance) du signal numérique pour une modulation en bande de base et l'appliquons à quelques cas pratiques importants. Puis, à partir de l'expression de l'enveloppe complexe, nous en déduisons le spectre des modulations MDP.

Reprenons l'expression (2.1) définissant le signal numérique en bande de base ou encore l'expression (2.3) définissant l'enveloppe complexe d'une modulation sur fréquence porteuse :

$$x(t) = \sum_k a_k h(t - kT) \tag{2.6}$$

On note $H(f)$ la transformée de Fourier de $h(t)$. On considère que les symboles a_k forment une suite aléatoire, à valeurs dans un alphabet $\mathcal{A}$ de taille M, et que cette suite est *stationnaire au second ordre*, ce qui signifie que la moyenne $\mathbb{E}\{a_k\} = m_a$ est indépendante de k et que la fonction d'autocovariance :

$$\mathbb{E}\{(a(n+k) - m_a)(a^*(n) - m_a^*)\} = R_a(k)$$

ne dépend que de l'écart de temps k. On en déduit que :

$$\mathbb{E}\{a(n+k)a^*(n)\} = R_a(k) + |m_a|^2$$

On suppose de plus que $R_a(k)$ est de module sommable.

Stationnarisation de $x(t)$

En toute rigueur, l'expression (2.6) ne définit pas un processus aléatoire stationnaire au second ordre même si la suite a_k est supposée stationnaire au second ordre. En effet, on vérifie aisément que les propriétés statistiques de $x(t)$ dépendent de l'instant t ; il suffit de considérer, par exemple des instants tels que $t = nT$. Pour obtenir un processus stationnaire il faut alors considérer que $x(t)$ a pour expression :

$$x(t) = \sum_k a_k h(t - kT + U)$$

où U désigne une variable aléatoire uniforme sur $(0, T)$ et *indépendante* des variables aléatoires $\{a_k\}$. Evidemment cette façon de faire conduit à des propriétés qui ignorent les conditions de phase de $x(t)$. Ce qui n'est pas surprenant dans la mesure où la notion de spectre, à laquelle nous nous intéressons ici, est une grandeur statistique qui ne prend pas en compte la phase des signaux (se souvenir par exemple que la d.s.p. en sortie d'un filtre a pour expression $|H(f)|^2 S_x(f)$ qui ne dépend pas de la phase de $H(f)$).

Moyenne de $x(t)$

Montrons que :

$$m_x = \mathbb{E}\{x(t)\} = \frac{H(0)}{T}\, m_a$$

En effet en introduisant la variable U uniforme sur $(0,T)$ indépendante de a_k, on a :

$$\mathbb{E}\{x(t)\} = \sum_k \mathbb{E}\{a_k\}\,\mathbb{E}\{h(t-kT+U)\} = m_a \sum_k \int_0^T h(t-kT+u)\frac{1}{T}\, du$$

en effectuant le changement de variable $\theta = t - kT + U$, on obtient le résultat annoncé.

Spectre de $x(t)$

On rappelle que le spectre est la transformée de Fourier de la fonction d'autocovariance de $x(t)$, à savoir $R_x(\tau) = \mathbb{E}\{(x(t+\tau)-m_x)(x(t)-m_x)^*\} = \mathbb{E}\{x(t+\tau)x^*(t)\} - |m_x|^2$. En adjoignant au spectre la composante $|m_x|^2$, liée à la moyenne $m_x = \mathbb{E}\{x(t)\}$, ce qui revient à prendre la transformée de Fourier de $\mathbb{E}\{x(t+\tau)x^*(t)\}$, on montre annexe 2.6.1 que la transformée de Fourier de $\mathbb{E}\{x(t+\tau)x^*(t)\}$ a pour expression :

$$S_x(f) = \underbrace{\frac{1}{T}\,|H(f)|^2 \sum_k R_a(k) e^{-2j\pi f kT}}_{S_x^c(f)} + \underbrace{\frac{1}{T^2}\,|m_a|^2 \sum_k |H(k/T)|^2 \delta(f-k/T)}_{S_x^d(f)} \tag{2.7}$$

Le terme $S_x^d(f)$, dans l'expression (2.7), est constitué d'une suite de mesures de Dirac concentrées aux fréquences multiples de $1/T$. En particulier, on retrouve en $k=0$ la composante $|m_a|^2|H(0)|^2/T^2$ due à la composante continue liée à la présence d'une moyenne m_x non nulle. On en déduit qu'il suffit que la suite a_k soit centrée pour que le terme $S_x^d(f)$ soit nul. Par contre, si m_a et $H(k/T)$ sont tous deux différents de 0, le signal comporte une composante sinusoïdale située à la fréquence k/T et donc en synchronisme avec la vitesse de modulation. Dans ce cas on peut récupérer, à la réception, l'information sur le rythme des symboles, en isolant cette raie par filtrage linéaire. On retiendra qu'une condition nécessaire pour avoir des raies à la fréquence symbole est que la moyenne m_a soit non nulle. Remarquons toutefois qu'un spectre peut ne pas comporter de raies aux fréquences multiples de $1/T$ mais que, par une transformation *qui n'est pas un filtrage linéaire*, comme par exemple une élévation au carré, il puisse donner un signal comportant des raies à un multiple de la fréquence $1/T$. Cette remarque est mise à profit dans certains dispositifs non linéaires de récupération de fréquence.

Dans le cas particulier où les symboles ne sont pas corrélés et sont centrés, $R_a(k) = 0$ pour $k \neq 0$ et on a :

$$S_x(f) = \frac{1}{T}\, R_a(0)|H(f)|^2 \tag{2.8}$$

$$R_x(\tau) = \frac{1}{T}\, R_a(0) h(\tau) \star h^*(-\tau) \tag{2.9}$$

$$R_x(0) = \frac{1}{T}\, R_a(0) \int_{\mathbb{R}} |h(t)|^2 dt = \frac{1}{T}\, R_a(0) \int_{\mathbb{R}} |H(f)|^2 df \tag{2.10}$$

Voyons à présent quelques exemples.

Signal binaire NRZ

Le signal binaire *Non Retour à Zéro* (NRZ) est obtenu à partir d'une impulsion rectangulaire $h(t) = A\mathrm{rect}_T(t)$ de durée T et d'amplitude A et de l'alphabet $\{-1,+1\}$. Comme les symboles sont supposés équiprobables et indépendants, $\mathbb{E}\{a_n\} = 0$ et $\mathbb{E}\{a_{n+k}a_n\} = \delta_k$. D'après la formule (2.8), la d.s.p. a pour expression :

$$S_x(f) = A^2 T \frac{\sin^2(\pi f T)}{(\pi f T)^2}$$

et la puissance transmise est $R_x(0) = A^2$.

Le lobe principal est de largeur $2/T$ et contient 91% de la puissance du signal. La décroissance à l'infini est de l'ordre de de $1/f^2$. Remarquons aussi que, plus la vitesse de modulation R est faible, plus l'intervalle de

temps T entre deux symboles successifs est grand et plus la largeur du premier lobe (qui peut être assimilée à l'ordre de grandeur de la bande d'occupation du signal NRZ) diminue. Nous avons représenté figure 2.5 le spectre en dB d'un signal binaire NRZ, de puissance 1, en fonction de fT.

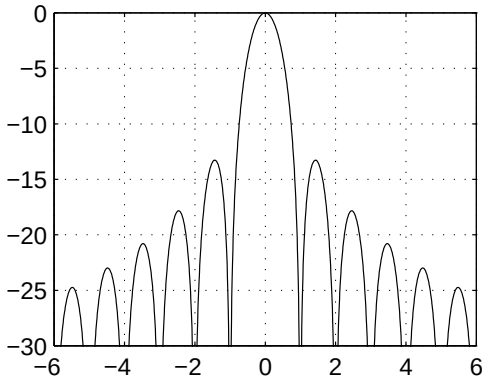

FIG. 2.5 – *Spectre en dB du signal NRZ, de puissance égale à* 1*, en fonction de* fT*.*

Signal binaire RZ

Le signal binaire *Retour à Zéro* (RZ) est obtenu à partir d'une impulsion rectangulaire de durée $\theta < T$ (typiquement $\theta = T/2$) et d'amplitude A et de l'alphabet $\{-1, +1\}$. On en déduit que $R_a(k) = \delta_k$. La d.s.p. a pour expression :

$$S_x(f) = A^2 T \frac{\sin^2(\pi f T/2)}{(\pi f T)^2}$$

et la puissance transmise est $R_x(0) = A^2/2$. La d.s.p. est très semblable à celle du signal NRZ, si ce n'est que les lobes sont deux fois plus larges. Nous l'avons représenté figure 2.6.

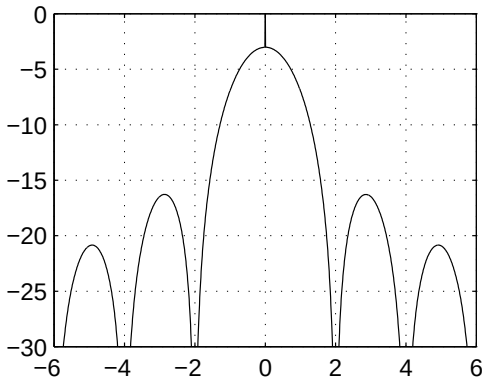

FIG. 2.6 – *Spectre en dB du signal biphase, de puissance égale à* 1*, en fonction de* fT*.*

Signal biphase

Le signal biphase est obtenu à partir de l'impulsion représenté figure 2.7 et de l'alphabet $\{-1, +1\}$. Le bit transmis est caractérisé par la présence d'un front montant ou descendant en $T/2$.

La puissance est $R_x(0) = A^2$ et la d.s.p. est donnée par :

$$S_x(f) = A^2 T \frac{4\sin^4(\pi f T/2)}{(\pi f T)^2}$$

Sa forme est représenté figure 2.8.

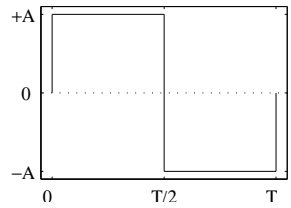

FIG. 2.7 – *Impulsion biphase.*

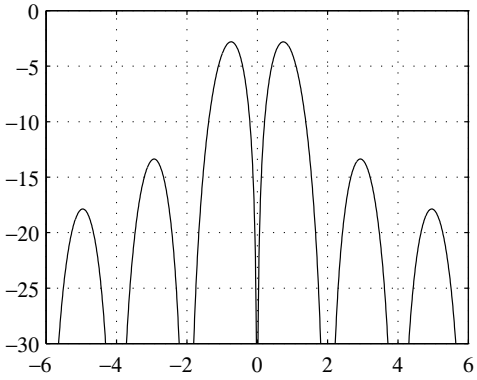

FIG. 2.8 – *Spectre en dB du signal biphase, de puissance* 1*, en fonction de* fT*.*

Signal AMI

Comme on l'a vu pour modifier le spectre on peut agir sur l'impulsion $h(t)$ mais on peut aussi agir sur les corrélations entre symboles. Un exemple fondamental est l'AMI (*Alternate Mark Inversion*). Il s'agit d'une transmission avec un alphabet ternaire $a_k \in \{-1, 0, +1\}$ où l'on code les bits 0 par le symbole 0 et les bits 1 *alternativement* par $+1$ et -1. Montrons que, si les symboles sont équiprobables, alors $m_a = 0$, $R_a(0) = 1/2$, $R_a(\pm 1) = -1/4$ et $R_a(k) = 0$ pour $|k| \geq 2$. En effet le codage AMI peut s'obtenir, de façon itérative, par les deux expressions suivantes :

$$\left\{ \begin{array}{l} a_n = d_n s_{n-1} \\ s_n = (1 - 2d_n) s_{n-1} \end{array} \right. \tag{2.11}$$

où s_n est une variable d'état à valeurs dans $\{-1, +1\}$ et où les variables aléatoires d_n sont supposées à valeurs dans $\{0, 1\}$ indépendantes, et telles que $\mathbb{P}\{d_n = 0\} = \mathbb{P}\{d_n = 1\} = 1/2$. On a alors que :

$$\begin{array}{l} \mathbb{E}\{d_n\} = 1/2 \text{ et donc } \mathbb{E}\{1 - 2d_n\} = 0 \\ \mathbb{E}\{d_n^2\} = 1/2 \text{ et donc } \mathbb{E}\{(1 - 2d_n)^2\} = 1 \\ \mathbb{E}\{d_n d_k\} = 1/4 \text{ si } n \neq k \text{ (indépendance} \Rightarrow \text{non corrélation)} \end{array}$$

Des deux équations (2.11), on déduit que :

$$a_n = d_n(1 - 2d_{n-1}) \dots (1 - 2d_{n-m}) \cdots$$

On a alors, en utilisant l'indépendance des d_k, $\mathbb{E}\{a_n\} = 0$, $\mathbb{E}\{a_n^2\} = 1/2$,

$$\mathbb{E}\{a_n a_{n-1}\} = \underbrace{\mathbb{E}\{d_n\}}_{=1/2} \underbrace{\mathbb{E}\{d_{n-1}(1 - 2d_{n-1})\}}_{=-1/2} \underbrace{\mathbb{E}\{(1 - 2d_{n-2})^2\}}_{=1} \cdots = -1/4$$

et, pour $k > 1$,

$$\begin{array}{rcl} \mathbb{E}\{a_n a_{n-k}\} & = & \mathbb{E}\{d_n\} \underbrace{\mathbb{E}\{(1 - 2d_{n-1})\}}_{=0} \cdots \mathbb{E}\{d_{n-k}(1 - 2d_{n-k})\} \cdots \\ & & \cdots \mathbb{E}\{(1 - 2d_{n-m})^2\} \cdots = 0 \end{array}$$

Comme la fonction d'autocovariance d'une suite réelle est paire, il vient le résultat annoncé. La suite a_n est centrée de fonction d'autocovariance :

$$R_a(k) = \begin{cases} 1/2 & \text{si } k = 0 \\ -1/4 & \text{si } k = \pm 1 \\ 0 & \text{si } |k| \geq 2 \end{cases} \tag{2.12}$$

On en déduit, pour une *impulsion rectangulaire NRZ*, que la puissance moyenne est $R_x(0) = A^2/2$ et que la d.s.p. a pour expression :

$$S_x(f) = A^2 T \frac{\sin^4(\pi f T)}{(\pi f T)^2} \tag{2.13}$$

Nous avons représenté figure 2.9 $S_x(f)$ en dB pour un signal AMI, de puissance égale à 1, en fonction de fT.

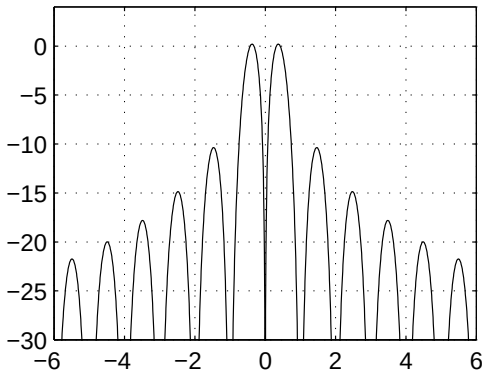

FIG. 2.9 – *Spectre en dB du signal AMI, de puissance égale à* 1*, en fonction de* fT.

On note que, dans le codage AMI, $S_x(0) = 0$. Cette propriété est très souvent requise dans les dispositifs qui transmettent mal l'énergie autour de la fréquence 0 (liaisons capacitives, couplage au réseau par transformateur d'isolation, etc). C'est souvent cette contrainte qui détermine, en tout premier lieu, le choix d'une modulation en bande de base. La condition suivante, dite de la *somme courante bornée* [**?**], fournit un moyen très simple de tester si $S_x(0) = 0$. S'il existe J tel que, quel que soit K :

$$\mathbb{P}\left\{\left|\sum_{k=1}^{K} a_k\right| < J\right\} = 1 \text{ alors } S_x(0) = 0$$

Ainsi dans l'AMI, la somme courante est bornée par 2.

Un autre élément important lors du choix d'une modulation est la possibilité de *récupérer de façon synchrone* le rythme symbole $1/T$. Le codage AMI répond à cette attente. Il permet de récupérer facilement le rythme $1/T$ en synchronisme avec le signal incident. Il suffit, en effet, d'effectuer sur le signal reçu une opération qui transforme les impulsions négatives en impulsions positives (on dit que l'on "redresse" le signal). On obtient alors le signal :

$$y(t) = \sum_k d_k g(t - kT)$$

où d_k représente la suite binaire d'informations et où $g(t) = \text{rect}_{T/2}(t)$. La formule générale (2.7) donne pour le spectre de $y(t)$ l'expression :

$$S_y(f) = S_y^c(f) + \frac{A^2}{4T^2} \sum_k \left|\frac{\sin^2(\pi f T/2)}{\pi^2 f^2}\right|^2 \delta(f - k/T)$$

Le signal $y(t)$ comporte donc une raie d'amplitude $A^2/4\pi^2$ à la fréquence $1/T$. Celle-ci peut être récupérée par simple filtrage linéaire autour de $1/T$ (attention le rythme $1/T$ est connu, l'inconnu ici est la synchronisation de ce rythme).

Signal MDP

Pour les modulations sur fréquence porteuse, nous nous limitons au calcul du spectre de la MDP. Rappelons que l'enveloppe complexe du signal a alors pour expression :

$$x_b(t) = \sum_k a_k h(t - kT)$$

où $h(t)$ est soit l'impulsion rectangulaire NRZ soit, plus généralement, une impulsion filtré. On suppose que les symboles $a_k = e^{j\phi_k}$, avec $\phi_k \in \{0, 2\pi/M, \ldots, 2(M-1)\pi/M\}$, forment une suite de variables aléatoires indépendantes et équiprobables. On en déduit que :

$$m_a = \mathbb{E}\left\{e^{j\phi_k}\right\} = \sum_{k=0}^{M-1} \frac{1}{M} e^{2j\pi k/M} = 0$$

que :

$$R_a(0) = \mathbb{E}\left\{|a_k|^2\right\} = 1$$

et que pour tout $k \neq 0$:

$$R_a(k) = \mathbb{E}\left\{a_{n+k} a_n^*\right\} = \mathbb{E}\left\{a_{n+k}\right\} \mathbb{E}\left\{a_n^*\right\} = 0$$

On rappelle que, pour une suite complexe, la fonction d'autocovariance vérifie $R_a(k) = R_a^*(-k)$ (symétrie hermitienne). Par conséquent $S_{x_b}(f)$ a pour expression :

$$S_{x_b}(f) = A^2 |H(f)|^2$$

D'après (2.10), la puissance de l'enveloppe complexe est donc égale à $A^2 \int |H(f)|^2 df$. D'après les propriétés énoncées paragraphe 1.7.4, la puissance transmise est :

$$R_x(0) = \frac{A^2}{2T} \int_{\mathbb{R}} |h(t)|^2 dt$$

et le spectre du signal numérique (notons que le spectre de $x_b(t)$ est ici une fonction paire) a pour expression :

$$S_x(f) = \frac{1}{4} S_{x_b}(f - f_0) + \frac{1}{4} S_{x_b}(-f - f_0)$$

Dans le cas particulier où $h(t)$ est l'impulsion rectangulaire NRZ, $h(t) = A\text{rect}_T(t)$ et la d.s.p. a pour expression :

$$S_{x_b}(f) = A^2 T \frac{\sin^2(\pi f T)}{(\pi f T)^2}$$

La puissance transmise est $R_x(0) = A^2/2$.

2.3 Performances en présence de bruit pour une transmission en bande de base

2.3.1 Filtre adapté

Schéma général d'une chaîne de transmission

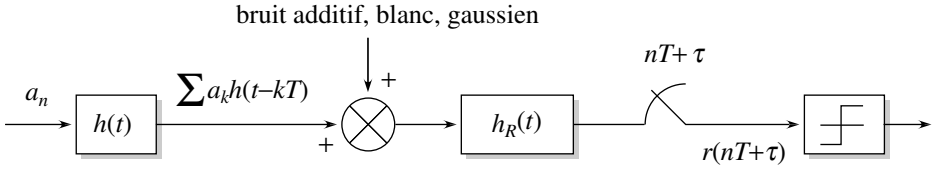

FIG. 2.10 – *Chaîne de transmission en bande de base*

Considérons la modulation numérique décrite par le signal :

$$x(t) = \sum_k a_k h(t - kT)$$

On rappelle que ce signal est réel dans la description des modulations en bande de base, tandis qu'il est complexe dans le cas des modulations sur fréquence porteuse. La transmission se fait à travers un canal non distordant, à bande limitée B, soumis à un bruit $b(t)$ additif, blanc, gaussien de densité spectrale de puissance $N_0/2$. Le signal reçu est :

$$z(t) = x(t) + b(t)$$

Avant d'aborder cette étude et afin d'illustrer ces différents traitements, nous avons représenté figure 2.11 les différents signaux tout au long de la chaîne de transmission : en (i) l'impulsion $h(t) = h_e(t) \star h_c(t)$ combinant le filtre d'émission et le filtre modélisant le canal de transmission. En (ii) le signal émis sans bruit. En (iii) le signal reçu bruité. En (iv) l'impulsion combinant le signal $p(t) = h(t) \star h_R(t)$ combinant le filtre $h(t)$ et le filtre de réception $h_R(t)$. Typiquement $h_R(t) = h(-t)$ (filtre adapté). En (v) la sortie du filtre $h_R(t)$ ainsi que les échantillons prélevés à la cadence T. En opérant une décision symbole par symbole par comparaison des échantillons à 0 (opération que nous justifions plus loin), on observe sur cette réalisation que la présence du bruit a conduit à un certain nombre d'erreurs.

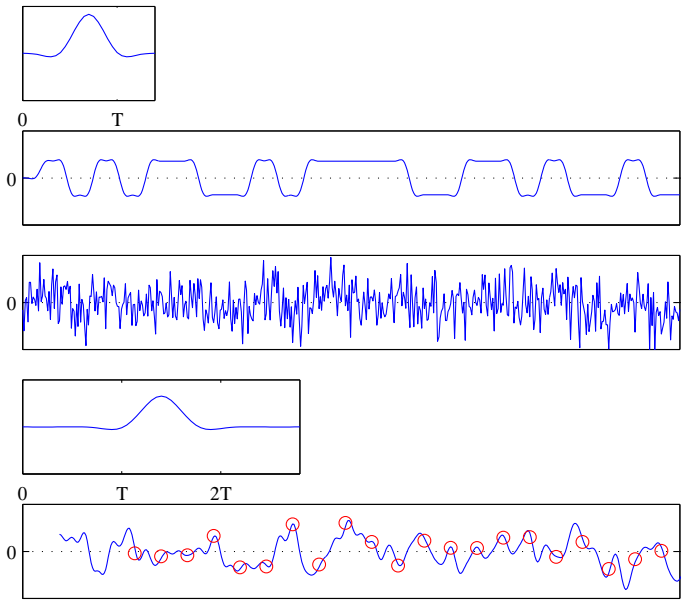

FIG. 2.11 – *Figure du haut : impulsion $h(t)$ combinant le filtre d'émission et le filtre de canal. Figure 2 : signal $x(t)$ émis correspondant à la séquence* 10101100101111001101001 00 *avec un alphabet binaire* $\{-1, +1\}$. *Figure 3 : signal bruité $z(t) = x(t) + b(t)$ reçu. Figure 4 : impulsion $p(t) = h(t) \star h_R(t)$ combinant $h(t)$ et le filtre $h_R(t)$ de réception (typiquement $h_R(t) = h(-t)$). Figure 5 : signal en sortie du filtre de réception. Les points ('o') représentent les échantillons prélevés à la cadence T. La décision se fait en comparant les échantillons à 0 : on obtient* 11000100101101111101001 00. *La présence du bruit conduit, sur cette réalisation, à des erreurs situées en position* 2, 3, 5, 13, 15 *et* 16.

Dans ce paragraphe nous allons précisément considérer le problème du choix du filtre $h(t)$, du filtre $h_R(t)$, de l'instant d'échantillonnage et de la valeur du seuil qui rendent minimale la probabilité d'erreur moyenne. Pour cela nous admettrons le résultat fondamental suivant : on peut atteindre le minimum de probabilité d'erreurs en effectuant la décision uniquement sur les échantillons prélevés à la cadence T en sortie d'un *filtre linéaire* de réponse impulsionnelle :

$$h_R(t) = h^*(\tau - t) \quad \Leftrightarrow \quad H_R(f) = H^*(f) e^{-2j\pi f \tau} \tag{2.14}$$

la valeur de τ étant choisie de façon à ce que l'impulsion $h_R(t)$ soit nulle pour $t < 0$. Ce filtre est appelé *filtre adapté* (sous entendu à la forme de l'impulsion $h(t)$, en anglais *matched filter*). Ce qu'il y a de remarquable

dans ce résultat est que l'on a remplacé le signal $r(t)$, à temps continu, par une suite de valeurs, à temps discret, sans perte de performances. Quant au fait que l'optimalité est obtenue par un dispositif linéaire, cela tient précisément au caractère gaussien du bruit.

Partant de ce résultat il faut ensuite déterminer, à partir de la suite des observations en sortie du filtre adapté, la suite de symboles émis qui minimise la probabilité d'erreur. Dans le cas général l'organe de décision est compliqué et fait appel à un algorithme dit de Viterbi.

Il existe toutefois une situation particulièrement simple et importante en pratique où la décision peut être prise *symbole par symbole*, c'est-à-dire au fur et à mesure que les valeurs en sortie du filtre adapté sont observées. C'est cette situation que nous allons étudier en détail.

2.3.2 Transmission sans IES : canal de Nyquist

Déterminons à présent l'expression des échantillons en sortie du filtre de réception. Pour cela posons :

$$p(t) = h(t) \star h^*(-t) = \int_{-\infty}^{+\infty} h(u)h^*(u-t)du$$

En désignant par $H(f)$ la transformée de Fourier de $h(t)$ et en utilisant les propriétés de la transformation de Fourier ainsi que l'expression (2.14), on a :

$$P(f) = H(f)H^*(f) = |H(f)|^2 = |H_R(f)|^2 \tag{2.15}$$

L'observation $r(t)$ en sortie du filtre adapté a alors pour expression :

$$r(t) = (x(t) + b(t)) \star h_R(t) = \sum_k a_k \int_{-\infty}^{+\infty} h(u - kT)h^*(\tau - t + u)du + b(t) \star h_R(t)$$

En effectuant le changement de variable $v = u - kT$ et en utilisant la définition de $p(t)$, on obtient en sortie du filtre adapté :

$$\begin{aligned} r(t) &= \sum_k a_k \underbrace{\int_{-\infty}^{+\infty} h(v)h^*(v - (t - \tau - kT))dv}_{p(t-\tau-kT)} + b(t) \star h_R(t) \\ &= \sum_k a_k p(t - \tau - kT) + b(t) \star h_R(t) \end{aligned}$$

En sortie de l'échantillonneur, on a alors aux instants $t_n = nT + \tau$:

$$r(nT + \tau) = a_n p(0) + \sum_{k \neq n} a_k p((n-k)T) + w_n \tag{2.16}$$

où w_n désigne une variable aléatoire scalaire, représentant la valeur échantillonnée en sortie du filtre adapté $h^*(\tau - t)$ et correspondant au bruit seul.

Dans l'expression (2.16), donnant la valeur de l'échantillon $r(nT + \tau)$ observé à l'instant $(nT + \tau)$, il apparaît trois termes :

- Le premier est relatif au symbole qui a été émis à l'instant nT ;
- Le second est relatif à tous les autres symboles autres que celui qui a été émis à l'instant nT. Pour cette raison il porte le nom d'*Interférences Entre Symboles* (en abrégé IES, en anglais ISI pour *Inter Symbol Interferences*) ;
- Le troisième est relatif au *bruit* additif sur le canal.

La présence du terme d'IES, qui contient de l'information utile sur plusieurs symboles émis, ne permet pas d'effectuer une décision *symbole par symbole* qui soit en même temps optimale. Pour résoudre le problème, une première approche consiste à faire en sorte que le terme d'IES soit nul ; elle aboutit au canal de Nyquist que nous allons étudier. Toutefois cette approche présente des faiblesses dans la mesure où elle ne permet pas de tirer au mieux profit de la bande disponible. La deuxième approche consiste soit à effectuer un algorithme de déconvolution, pour réduire l'IES à un niveau tel que l'on puisse effectuer une décision symbole par symbole[2], soit utiliser l'algorithme de décision optimale (algorithme de Viterbi). La deuxième approche ne sera pas traitée ici.

[2]Cette solution est sous-optimale

Critère de Nyquist

Supposons que les symboles soient statistiquement indépendants et que la condition :

$$p(mT) = 0 \quad \text{pour} \quad m \neq 0 \tag{2.17}$$

soit vérifiée. Alors le second terme d'IES est nul et l'échantillon prélevé à l'instant $(nT + \tau)$ ne dépend statistiquement que du symbole émis à l'instant nT. On dit qu'il y a suppression de l'interférence entre symboles aux instants d'échantillonnage. La condition s'appelle *condition de Nyquist* et le canal correspondant le canal idéal de Nyquist. Elle permet, comment nous allons le voir, d'effectuer une décision symbole par symbole.

Remarquons que la condition de Nyquist est automatiquement satisfaite si l'impulsion $h(t)$ est de durée inférieure à T. En effet dans ce cas, il ne peut y avoir d'IES, puisque l'impulsion correspondant au n-ième symbole est nulle avant même que l'impulsion correspondant au $(n+1)$-ième symbole ne commence. Le calcul le montre bien, puisque l'impulsion $p(t)$ est alors de durée inférieure à $2T$ et vérifie donc la condition de Nyquist.

La condition de Nyquist donnée par (2.17) porte sur la forme temporelle du signal $p(t)$. Si on note $P(f)$ la transformée de Fourier de $p(t)$, la formule de Poisson (voir annexe) fournit la condition équivalente suivante :

$$\sum_k P(f - k/T) = Tp(0) = \text{constante} \tag{2.18}$$

On en déduit qu'une *condition nécessaire* pour rendre possible une transmission sans IES sur un canal de bande B est que le débit symbole R vérifie :

$$R \leq 2B \tag{2.19}$$

Le cas limite correspond à un spectre $P(f)$ rectangulaire. Une forme largement utilisée est celle des impulsions dites en *cosinus surélevé* (en anglais "raised-cosine") donnée par :

$$C_\alpha(f) = \begin{cases} T & \text{pour} \quad |f| < \dfrac{1-\alpha}{2T} \\ \dfrac{T}{2}\left[1 - \sin\left(\dfrac{\pi T}{\alpha}(f - 1/2T)\right)\right] & \text{pour} \quad \dfrac{1-\alpha}{2T} < |f| < \dfrac{1+\alpha}{2T} \\ 0 & \text{pour} \quad |f| > \dfrac{1+\alpha}{2T} \end{cases} \tag{2.20}$$

avec $\alpha \in (0, 1)$. Notons que $C_\alpha(f)$ est paire. Nous avons représenté figure 2.12 $C_\alpha(f)/T$ pour $\alpha = 0$, 0.3 et 1, pour $f > 0$. L'axe des fréquences en abscisses est gradué en $R = 1/T$.

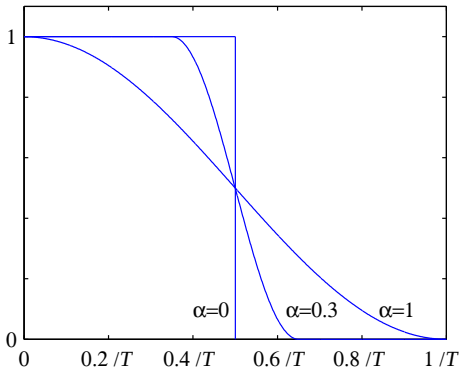

FIG. 2.12 – *Transformée de Fourier de l'impulsion en cosinus-surélevé pour $\alpha = 0$, 0.3 et 1. En abscisses, l'axe des fréquences est gradué en $R = 1/T$. On voit que la bande occupée est $B = (1+\alpha)/2T$.*

Le paramètre α s'appelle le *facteur de débordement* (ou coefficient d'arrondi : en anglais "roll-off"). Il varie entre 0 et 1. Plus il est grand, plus la bande en fréquence donnée par :

$$B = \frac{1+\alpha}{2T} = (1+\alpha)\frac{D_b}{2\log_2(M)} \tag{2.21}$$

nécessaire pour transmettre est grande. En contrepartie l'impulsion $c_\alpha(t)$ (voir exercice 2.6) a les lobes secondaires dont l'amplitude est moindre.

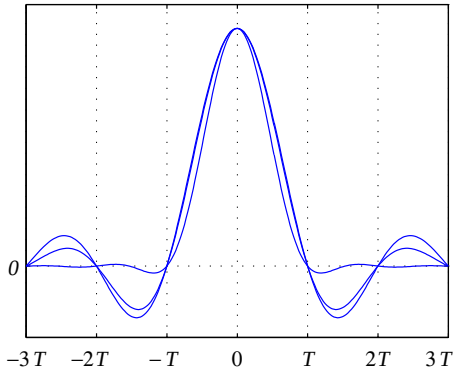

FIG. 2.13 – *Impulsion en cosinus-surélevé pour $\alpha = 0$, 0.3 et 1. En abscisses, l'axe des temps est gradué en T. On voit que plus α est grand plus les lobes secondaires sont de faibles amplitudes.*

Il s'en suit que, lors d'un décalage de l'horloge d'échantillonnage, l'amplitude de l'IES est d'autant plus faible que α est grand.

Répartition du filtrage : racine carrée de Nyquist

En conclusion les résultats précédents montrent que l'on doit répartir à part égale l'impulsion en cosinus surélevé entre émission et réception suivant :

$$\begin{cases} H_R(f) = H^*(f) & \text{(Filtre adapté)} \\ |H(f)| = |H_R(f)| = \sqrt{C_\alpha(f)} & \text{(afin que : } H_e(f)H_R(f) = C_\alpha(f)\text{)} \end{cases}$$

Si $C_\alpha(f)$ est donnée par (2.20), on montre, par un calcul long mais sans difficulté, que la fonction :

$$g_\alpha(t) = \frac{4\alpha}{\pi\sqrt{T}} \frac{\cos((1+\alpha)\frac{\pi t}{T}) + \frac{T}{4\alpha t}\sin((1-\alpha)\frac{\pi t}{T})}{1 - 16\alpha^2 t^2/T^2}$$

vérifie précisément $g(t) \star g(-t) = c_\alpha(t)$ où $c_\alpha(t)$ est la transformée de Fourier inverse de $C_\alpha(f)$. Par continuité on obtient aussi $g(T/4\alpha) = -0.5\cos(\gamma) + (\pi/4)\sin(\gamma)$ avec $\gamma = \pi(1+\alpha)/4\alpha$. En pratique l'expression (2.22) peut être utilisée pour engendrer, sur microprocesseur, les échantillons du signal numérique à une cadence suffisante pour être ensuite mis à l'entrée d'un convertisseur numérique analogique qui délivre le signal numérique modulé.

Remarquons ici que les filtres, donnés par l'expression (2.22), sont définies à une phase $\Phi(f) = -\Phi_R(f)$ près, qui est choisie, en général, de façon à introduire peu de distorsion, en particulier en s'approchant autant que possible d'une phase linéaire (retard pur).

Diagramme de l'œil

Un moyen pratique, très largement utilisé, pour "évaluer" la situation de non interférence entre symboles dans une transmission, est l'observation du *diagramme de l'œil.*

Considérons une transmission binaire d'alphabet $\{-1, +1\}$. Lors d'une suite de symboles successifs, le signal numérique observé en l'absence de bruit est la somme algébrique d'impulsions $p(t)$ émises à la cadence T et multipliées par l'une des deux valeurs -1 ou $+1$; si la durée de l'impulsion est $p(t)$ est supérieure à $2T$, il s'ensuit que le symbole émis à l'instant kT interfère avec les symboles suivants. Si à présent les séquences de bits sont également probables, on observera, dans les intervalles de temps de longueur T, toutes les formes possibles de signaux ; si maintenant on superpose toutes ces formes et que l'on ne conserve que deux intervalles successifs, on obtient le diagramme de l'œil. Pour obtenir cette superposition on observe à l'oscilloscope le signal numérique en se synchronisant sur le temps T.

Nous avons représenté figures 2.14 et 2.15 la forme du diagramme de l'œil pour une impulsion $p(t)$ en cosinus surélevé avec $\alpha = 0{,}3$ pour $M = 2$ et $M = 4$ sans bruit. On notera la forme caractéristique du diagramme de l'œil, lorsque le critère de Nyquist est vérifié : les trajectoires concourent aux instants kT, ce qui a pour conséquence de rendre l'œil très ouvert verticalement à ces instants.

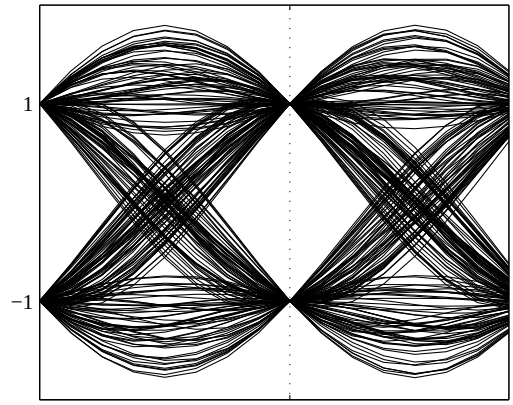

FIG. 2.14 – *Diagrammme de l'œil sur le canal de Nyquist pour $\alpha = 0{,}3$ et $M = 2$, sans bruit.*

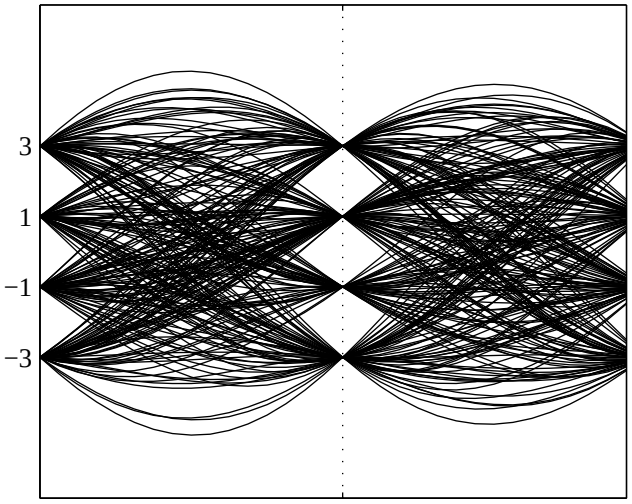

FIG. 2.15 – *Diagrammme de l'œil sur le canal de Nyquist pour $\alpha = 0{,}3$ et $M = 4$, sans bruit.*

L'observation du diagramme de l'œil fournit les indications suivantes :

- L'ouverture verticale mesure les performances contre le bruit. Plus l'œil est ouvert en hauteur, plus il est facile de discriminer les deux symboles en présence de bruit et donc, plus la probabilité d'erreur est faible. Si le diagramme manifeste la présence d'une IES (faible), et que l'on souhaite continuer à utiliser une détection à seuils (solution sous optimale), il faudra venir échantillonner le signal $r(t)$ aux instants où l'œil a une ouverture maximum.
- L'ouverture horizontale indique une résistance à un décalage des instants d'échantillonnage. Ainsi plus l'œil est ouvert en largeur, plus les lobes secondaires de la réponse en temps seront faibles et plus l'accumulation des interférences dues au décalage des instants d'échantillonnage auront une influence moindre en terme de probabilité d'erreur. C'est le cas pour les fonctions en cosinus surélevé lorsque α augmente.

Nous avons représenté figure 2.16 le diagramme de l'œil sur le canal de Nyquist pour $\alpha = 0{,}3$ et $M = 2$, avec un rapport signal sur bruit de 7 dB. On remarque que la décision symbole par symbole sera la meilleure là où l'œil est le plus ouvert verticalement. Malgré le bruit les niveaux significatifs des deux symboles restent relativement bien discriminables.

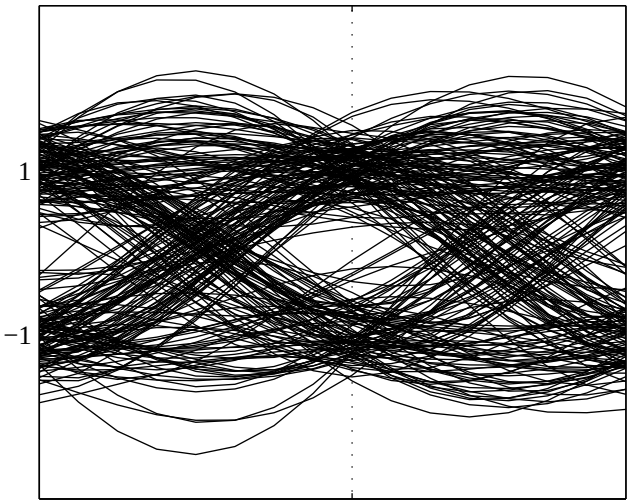

FIG. 2.16 – *Diagrammme de l'œil sur le canal de Nyquist pour* $\alpha = 0{,}3$ *et* $M = 2$*, avec un rapport signal sur bruit de* 7 *dB.*

Distorsion maximale d'IES

La distorsion maximale d'IES est un moyen quantitatif de juger du niveau d'IES. Pour l'introduire, considérons une transmission binaire telle que l'observation en sortie du filtre de réception ait pour expression :

$$r(nT + \tau) = a_n - 0.6a_{n-1} + 0.4a_{n-2} - 0.2a_{n-3} + 0.1a_{n-4}$$

où $a_n \in \{-1, +1\}$. Une détection *symbole par symbole* par comparaison à 0 est clairement désastreuse : si $a_n = +1$, il suffit que la séquence des 4 symboles qui précèdent soit $\{+1, -1, +1, -1\}$ pour que la décision soit erronée. Par conséquent, même sans bruit, la probabilité d'erreur est supérieure à $1/2^4 \sim 6\%$. Ce qui est, en général, inacceptable. Les 4 symboles qui précédent interfèrent de façon destructive avec le symbole à détecter. On en déduit que la décision symbole par symbole sera d'autant plus difficile que le rapport des amplitudes de l'IES par rapport à l'amplitude du symbole à détecter est grand. D'où l'idée de mesurer l'IES par le rapport $(\sum_k |p_k| - |p_{\max}|)/|p_{\max}|$ où $p_{\max} = \max_k p_k$. De façon plus générale pour une modulation à M-aire, on définit la distorsion maximale d'IES par :

$$D_{\max} = (M-1)\frac{\sum_k |p(k)| - |p_{\max}|}{|p_{\max}|}$$

où $p_{\max} = \max_k p(k)$. Si $D_{\max} \ll 1$, une décision symbole par symbole peut être utilisée, cela correspond à un œil très ouvert verticalement. Par contre si $D_{\max}$ est voisin de 1 ou supérieur à 1, une décision optimale symbole par symbole est impossible. Il faut prendre en compte l'ensemble des observations pour détecter la séquence la plus probable. Un algorithme particulièrement efficace existe, il s'agit de l'algorithme de Viterbi.

Performances d'une transmission sur le canal de Nyquist

Revenons au schéma général de la chaîne de transmission figure 2.10 où nous supposons que le critère de Nyquist est vérifié et envisageons tout d'abord une transmission binaire en bande de base. Nous avons alors les hypothèses suivantes :

- a_n est une suite aléatoire i.i.d. à valeurs dans $\{-1, +1\}$ avec $\mathbb{P}\{a_n = 0\} = \mathbb{P}\{a_n = 1\} = 1/2$,
- $h(t)$ est réelle et est tel que $p(t) = h(t) \star h^*(-t)$ vérifie le critère de Nyquist,
- le bruit est *blanc, gaussien* et indépendant de a_n,

D'après le critère de Nyquist, les échantillons en sortie du filtre adapté ont pour expression :

$$r(nT+\tau) = a_n p(0) + w_n$$

D'après les hypothèses a_n et w_n sont indépendantes. On note $w^a(t) = h_R(t) \star b(t)$ le signal à temps continu produit, à la réception, par le bruit en sortie du filtre adapté. Montrons que les variables aléatoires w_n définies par :

$$w_n = w^a(t)|_{t=nT}$$

sont gaussiennes centrées, non corrélées, de variance $\sigma^2 = p(0)N_0/2$. Le caractère gaussien se déduit du caractère linéaire de l'opération entre $b(t)$ et $w^a(t)$ et de l'hypothèse gaussienne sur $b(t)$. Les formules de filtrage donnent tout d'abord pour la moyenne $\mathbb{E}\{w_n\} = \mathbb{E}\{b(t)\} H_R(0) = 0$ puisque $b(t)$ est supposé centré. Ces formules donnent pour la densité spectrale de puissance de $w^a(t)$ l'expression $S_w^a(f) = |H_R(f)|^2 S_b(f)$. En utilisant l'expression (2.15) on obtient alors :

$$S_w^a(f) = \frac{N_0}{2} P(f)$$

Et par conséquent la fonction d'autocovariance, qui est la transformée de Fourier de $S_w^a(f)$, a pour expression :

$$\begin{aligned} \mathbb{E}\{w_n w_k\} &= \mathbb{E}\{w^a(nt) w^a(kT)\} = R_w((n-k)T) \\ &= \int_{\mathbb{R}} S_w^a(f) e^{2j\pi(n-k)Tf} df = \frac{N_0}{2} \int_{\mathbb{R}} P(f) e^{2j\pi(n-k)Tf} df \end{aligned}$$

En revenant à $p(t)$, on obtient :

$$\mathbb{E}\{w_n w_k\} = \frac{N_0}{2} p((n-k)T) = \frac{N_0}{2} p(0) \delta_{k,n} \tag{2.22}$$

la dernière égalité étant justifiée par l'hypothèse que $p(t)$ satisfait le critère de Nyquist. L'expression (2.22) montre que, pour $n \neq k$, les variables aléatoires w_n et w_k ne sont pas corrélées, comme elles sont gaussiennes elles sont en plus indépendantes. Remarquons ici que l'indépendance des variables aléatoires w_k est essentielle pour assurer qu'une décision symbole par symbole est optimale.

En conclusion l'observation :

$$Y_n = r(nT+\tau) = a_n p(0) + w_n \tag{2.23}$$

en sortie du filtre adapté, est une variable aléatoire gaussienne, de variance :

$$\sigma^2 = \frac{p(0)N_0}{2} \tag{2.24}$$

et de moyenne $p(0)$ (respectivement $-p(0)$) si le symbole émis a_n est $+1$ (respectivement -1). Les lois de l'observation conditionnellement à $a_n = +1$ et $a_n = -1$ ont donc pour densités respectives :

$$\begin{aligned} p_{Y_n|a_n=-1}(y) &= \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(y+p(0))^2}{2\sigma^2}\right) \\ p_{Y_n|a_n=+1}(y) &= \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(y-p(0))^2}{2\sigma^2}\right) \end{aligned} \tag{2.25}$$

On choisit un seuil s et on prend comme règle de décision :

$$\hat{a}_n = \begin{cases} +1 & \text{si} \quad Y_n > s \\ -1 & \text{si} \quad Y_n < s \end{cases}$$

La probabilité d'erreur moyenne s'écrit :

$$\begin{aligned} P_e &= \frac{1}{2}\mathbb{P}\{\hat{a}_n = +1 | a_n = -1\} + \frac{1}{2}\mathbb{P}\{\hat{a}_n = -1 | a_n = +1\} \\ &= \frac{1}{2}\mathbb{P}\left\{\hat{Y}_n > s | a_n = -1\right\} + \frac{1}{2}\mathbb{P}\left\{\hat{Y}_n < s | a_n = +1\right\} \end{aligned}$$

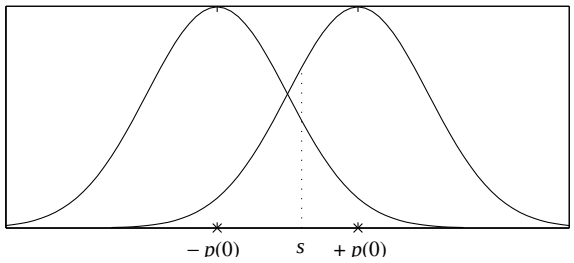

FIG. 2.17 – *Densités des lois d'observation dans le cas binaire.*

où on a utilisé que les deux symboles étaient équiprobables. En utilisant les expressions (2.25), il est immédiat de montrer[3] que le seuil qui minimise P_e est $s = 0$ et que la probabilité d'erreur obtenue est alors donnée par :

$$P_e = \int_0^{+\infty} p_{Y_n|a_n=-1}(y)dy = \int_{p(0)/\sigma}^{+\infty} \frac{1}{\sqrt{2\pi}}\, e^{-t^2/2}dt = Q(p(0)/\sigma)$$

où la fonction $Q(x)$ est définie par :

$$Q(x) = \int_x^{+\infty} \frac{1}{\sqrt{2\pi}}\, e^{-t^2/2}dt$$

Nous allons à présent exprimer P_e en fonction de E_b énergie moyenne par bit. Comme les deux symboles sont supposés équiprobables et transportent un bit, $E_b = \int |h(t)|^2 dt = p(0)$. En utilisant l'expression (2.24), on obtient :

$$P_e = Q\left(\sqrt{\frac{2E_b}{N_0}}\right) \tag{2.26}$$

On vérifie que P_e est une fonction décroissante du rapport E_b/N_0, quantité qui peut être interprétée comme un rapport signal sur bruit, le numérateur étant le nombre de Joules consommés en moyenne par bit transmis.

Transmission M-aire en bande de base sur le canal de Nyquist

Dans le cas d'une transmission M-aire, le récepteur optimal est encore constitué d'un filtre adapté à l'impulsion $h(t)$, suivi d'un échantillonneur. A condition que le critère de Nyquist soit vérifié, les quantités en sortie du filtre adapté ont encore pour expression $r(nT+\tau) = a_n p(0)+w_n$, où $a_n \in \{\pm 1, \pm 3, \cdots, \pm(M-1)\}$ et où les w_n sont des variables aléatoires gaussiennes centrées, non corrélées, de même variance $\sigma^2 = p(0)N_0/2$. Un calcul simple montre que :

$$\mathbb{E}\{a_n\} = 0$$

et que

$$\mathbb{E}\left\{a_n^2\right\} = \frac{2}{M}\left(1^2 + 3^2 + \cdots + (M-1)^2\right) = \frac{M^2-1}{3}$$

Le récepteur est identique à celui de la MIA-2 et l'observation en sortie de l'échantillonneur est donnée par :

$$Y_n = p(0)a_n + w_n$$

Un calcul analogue au calcul fait pour $M = 2$ montre que la loi de Y_n conditionnellement à a_n est une loi gaussienne de moyenne $a_n p(0)$ et de variance $\sigma^2 = N_0 p(0)/2$. Nous avons représenté figure 2.18 les densités de probabilités des lois conditionnelles dans le cas où $M = 8$.

L'organe de décision est un détecteur à seuils, dont les seuils sont situés entre les amplitudes possibles $a_k p(0)$, c'est-à-dire aux points d'amplitude $\{0, \pm 2p(0), \pm 4p(0), \cdots\}$. Si on note :

$$\int_{-p(0)}^{p(0)} \frac{1}{\sigma\sqrt{2\pi}}\, e^{-x^2/2\sigma^2}dx = 1 - 2q \quad \text{où} \quad q = \frac{1}{\sqrt{2\pi}}\int_{p(0)/\sigma}^{+\infty} e^{-t^2/2}dt = Q\left(p(0)/\sigma\right)$$

[3]On peut montrer que le filtre adapté est, parmi tous les filtres linéaires, celui qui maximise le rapport $p(0)/\sigma$.

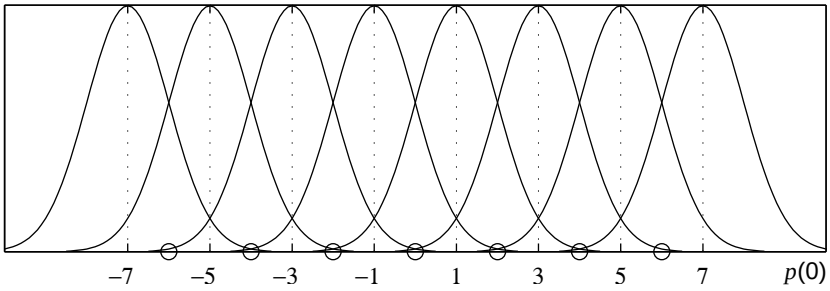

FIG. 2.18 – *Densités des lois de probabilité de l'observation en sortie de l'échantillonneur, pour une transmission 8-aire sur un canal additif, gaussien, blanc. Les points (o) représentent les seuils de décision.*

la probabilité de décision correcte s'écrit :

$$P_c = \frac{1}{M}\,(1-q) + \frac{M-2}{M}\,(1-2q) + \frac{1}{M}\,(1-q) = 1 - \frac{2}{M}\,(M-1)q$$

On en déduit que la probabilité d'erreur est donnée par :

$$P_e = 1 - P_c = \frac{2(M-1)}{M}\,Q\left(p(0)/\sigma\right)$$

En utilisant que l'énergie moyenne par symbole est donnée par :

$$E_s = \mathbb{E}\left\{a_n^2 \int |h(t)|^2 dt\right\} = \frac{M^2-1}{3} p(0)$$

on déduit que l'énergie moyenne par bit est donnée par :

$$E_b = \frac{E_s}{\log_2(M)} = \frac{1}{\log_2(M)} \frac{M^2-1}{3} p(0) \tag{2.27}$$

En portant dans l'expression de P_e et en utilisant que $\sigma^2 = p(0)N_0/2$, on obtient :

$$P_e = 2(1-1/M)Q\left(\sqrt{\frac{2E_b}{N_0}\log_2(M)\frac{3}{M^2-1}}\right) \tag{2.28}$$

TEEB et codage de Gray

En plus de l'expression de la probabilité d'erreurs P_e par symbole, on est aussi intéressé en pratique par le *taux d'erreurs par élément binaire*, en abrégé TEEB (en anglais BER pour *Bit Error Rate*). De façon générale son expression est ;

$$\text{TEEB} = \sum_{a\in\mathcal{A}} \sum_{b\in\mathcal{A}} \tau(a,b)\mathbb{P}\left\{\hat{a}_n = b | a_n = a\right\} \mathbb{P}\left\{a_n = a\right\}$$

où $\tau(a,b)$ représente le taux de bits erronés entre les deux mots-code a et b. a_n et $\hat{a}_n$ désignent respectivement le symbole émis et le symbole décidé à l'instant n. La détermination précise de l'expression du TEEB est toutefois très compliquée et dépend, par l'intermédiaire de $\tau(a,b)$, de l'association entre les éléments binaires et les symboles M-aires.

Une formule approchée pratique peut cependant être obtenue lorsque le rapport signal sur bruit est suffisamment grand et si le code utilisé est un code de Gray. On rappelle que, pour un code de Gray, les mots-code de deux amplitudes voisines ne diffèrent que par *un* bit.

Considérons par exemple le cas $M = 8$. Dans le tableau 2.2 nous avons indiqué, conditionnellement à l'émission du symbole $a_n = +3$, les différentes possibilités de décision pour le symbole $\hat{a}_n$ ainsi que la probabilité d'une telle décision et le taux de bits erronés correspondant. Les approximations sont valides si on suppose que le rapport signal sur bruit est suffisamment grand pour que les seules erreurs correspondent au cas où le symbole choisi est l'un des deux symboles *voisins en amplitude* du symbole émis.

$b \in \mathcal{A}$	e.b.	$\mathbb{P}\{\hat{a}_n = b \mid a_n = +3\}$	$\tau(b, +3)$
-7	111	≈ 0	3/3
-5	110	≈ 0	2/3
-3	100	≈ 0	1/3
-1	101	≈ 0	2/3
$+1$	001	$\approx P_e/2$	1/3
$+3$	000	$= 1 - P_e$	0/3
$+5$	010	$\approx P_e/2$	1/3
$+7$	011	≈ 0	2/3

TAB. 2.2 – *TEEB conditionnellement à $a_n = +3$. Colonne 1 : différentes possibilités de décision. Colonne 2 : suite d'éléments binaires (e.b.) décidée. Colonne 3 : probabilité d'une telle décision. Colonne 4 : taux d'éléments binaires erronés. Le rapport signal sur bruit est supposé grand.*

On en déduit que le taux d'erreurs par éléments binaires, conditionnellement à l'émission de $a_n = +3$, est $\text{TEEB}_3 \approx P_e/3$. Comme les symboles sont supposés équiprobables, le taux d'erreurs par éléments binaires est donc $P_e/3$.

En conclusion, si le rapport signal sur bruit est suffisamment grand et si on utilise un codage de Gray, un seul bit sur les $\log_2(M)$ bits transmis par symbole est faux et l'expression du taux d'erreurs par éléments binaires (TEEB) est alors donnée par :

$$\text{TEEB} = \frac{P_e}{\log_2(M)}$$

Nous avons représenté figure 2.19 le TEEB en fonction du rapport E_b/N_0, pour différentes valeurs de M.

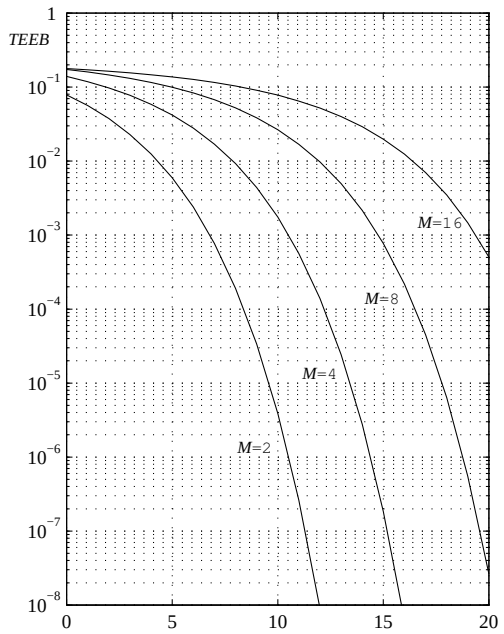

FIG. 2.19 – *TEEB en fonction du rapport E_b/N_0 en dB, pour $M = 2, 4, 8, 16$*

Efficacité spectrale/rapport signal sur bruit

– D'après la formule (2.21), l'efficacité spectrale en MIA-M, sur le canal de Nyquist et avec une impulsion en cosinus sur-élevée, est donnée par :

$$\eta = \frac{D_b}{B} = 2\frac{\log_2(M)}{1+\alpha} \tag{2.29}$$

M	2	4	8	16
$E_b/N_0\,(dB)$	9.6	13.6	18	22.9
$\Delta\,(dB)$	0	4	8.4	13.3

TAB. 2.3 – *Rapport signal sur bruit en MIA-M pour un TEEB* $= 10^{-5}$.

.

Si on considère le cas (le plus favorable) où $\alpha = 0$, on a :

$$\eta = 2\log_2(M) \tag{2.30}$$

soit

M	2	4	8	16
$\eta\,(bits/s/Hz)$	2	4	6	8

– Le tableau 2.3 donne les valeurs du rapport signal sur bruit relevées sur les courbes de la figure 2.19 pour TEEB $= 10^{-5}$. Δ est l'augmentation en dB du rapport signal sur bruit pour compenser en terme de TEEB l'augmentation de M par rapport à $M = 2$. On peut même admettre à la vue des courbes, qui sont quasiment parallèles, que ces écarts sont encore pertinents pour les faibles valeurs du TEEB.

Nous avons reporté figure 2.20 les positions (•) du rapport signal sur bruit en fonction de l'efficacité spectrale. Les valeurs correspondent à un TEEB de 10^{-5}. Nous avons aussi reporté la limite fondamentale de la capacité d'un canal gaussien donnée par l'expression (2.5).

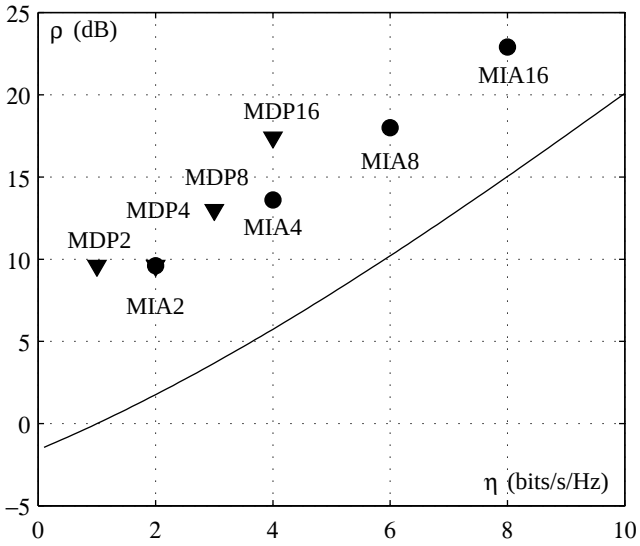

FIG. 2.20 – *Efficacité spectrale en fonction du rapport signal sur bruit pour les modulations MIA-M (•) et pour les modulations MDP-M (▼). Les valeurs correspondent à un TEEB de* 10^{-5}. *La courbe en trait plein représente la limite fondamentale du canal gaussien (équation* (2.5)*).*

2.4 Performances en présence de bruit pour les modulations sur fréquence porteuse

Pour les modulations sur fréquence porteuse, les calculs peuvent être effectués en s'appuyant sur la description faite à partir de l'enveloppe complexe. En effet on rappelle qu'une modulation sur fréquence porteuse est décrite par le signal complexe :

$$x_b(t) = \sum_k a_k h(t - kT)$$

où $h(t)$ est une impulsion en général complexe et où a_k désigne une suite de symboles complexes pris à un alphabet de M symboles et que le signal numérique transmis s'écrit $x(t) = \text{Ré}\{x_b(t)\exp(2j\pi f_0 t)\}$.

Ce signal est soumis à un bruit $n(t)$ additif, gaussien, centré, de d.s.p. constante et égale à $N_0/2$ dans la bande utile du signal $x(t)$. A la réception le signal observé a pour expression $z(t) = x(t) + n(t)$. Son enveloppe complexe (par rapport à f_0) s'écrit $z_b(t) = x_b(t) + n_b(t)$. On sait que :

$$n_b(t) = n_b^r(t) + jn_b^i(t)$$

où $n_b^r(t)$ et $n_b^i(t)$ désignent respectivement les composantes en phase et en quadrature du bruit. On a montré chapitre 1 que ce sont deux processus aléatoires, gaussiens, centrés, indépendants, ayant même densité spectrale de puissance égale à N_0 dans la bande utile du signal $x_b(t)$.

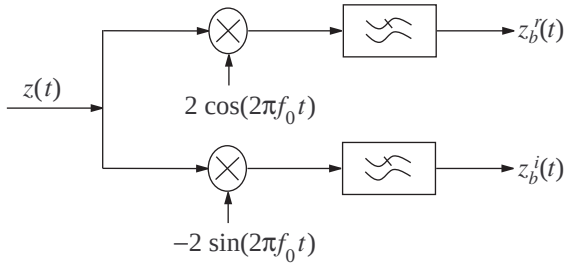

FIG. 2.21 – *Démodulation synchrone.*

On rappelle (voir figure 2.21) que le signal complexe $z_b(t)$ s'obtient à partir du signal reçu $z(t)$ en utilisant un *détecteur synchrone*, c'est-à-dire une opération de multiplication de $z(t)$ par $e^{-2j\pi f_0 t}$ suivie d'un filtrage passe-bas.

Comme pour une transmission en bande de base, on montre que le minimum de probabilité d'erreur s'obtient en effectuant la décision sur les échantillons complexes pris à la cadence T en sortie du filtre de réception, ce dernier étant le filtre *adapté* $h_R(t) = h^*(\tau - t)$. Comme, dans la plupart des cas pratiques, le filtre adapté est un filtre passe-bas plus étroit que le filtre du détecteur synchrone qui élimine les composantes autour de $2f_0$, ce dernier peut être omis dans le schéma de réception.

Dans le cas général, la réponse $h(t)$ est complexe et par conséquent la réponse $h_R(t)$ du filtre adapté est elle-même complexe. Il s'en suit que le filtrage de $z(t)$ par $h_R(t)$ s'écrit :

$$r_b(t) = z_b(t) \star h_R(t) = (z_b^r(t) + jz_b^i(t)) \star (h_R^r(t) + jh_R^i(t))$$

Ce traitement nécessite donc 4 opérations réelles de convolution. La situation se simplifie un peu si $h(t)$ est réel auquel cas le filtrage de $z_b(t)$ ne comporte que deux opérations réelles de filtrage.

Annulation de l'IES : critère de Nyquist

En choisissant pour $p(t) = h(t) \star h^*(-t)$ une fonction qui satisfait le critère de Nyquist (2.17), à savoir $p(mT) = 0$ pour tout $m \neq 0$, il y a annulation de l'interférences entre symboles. Typiquement on prendra pour $p(t)$ une fonction en cosinus sur-élevé dont on déterminera la valeur α du roll-off de façon à satisfaire la contrainte de bande de fréquence du canal. Dans ce cas la sortie échantillonnée a pour expression *complexe* :

$$Y_n = r(nT + \tau) = a_n p(0) + w_n$$

où w_n désigne la partie de l'observation due au bruit $n_b(t)$ filtré par $h_R(t)$. En calculant la d.s.p. du bruit w_n en sortie du filtre $h_R(t)$ et en utilisant le critère de Nyquist, on montre aisément que w_n^r et w_n^i sont deux suites de variables aléatoires gaussiennes, centrées, indépendantes, de même variance $N_0 p(0)$.

L'organe de décision peut donc prendre une décision symbole par symbole : il choisit le symbole d'alphabet le plus proche de la valeur r_n observée. Il est important de noter que l'organe de décision, qui intervient après l'échantillonneur, traite des nombres complexes. Cela signifie que la décision est prise symbole complexe par symbole complexe, mais pas en général partie réelle par partie réelle et partie imaginaire par partie imaginaire (sauf si la constellation est un quadrillage régulier).

Nous allons étudier plus en détails le cas de la modulation de phase.

2.4.1 Cas de la MDP-M

D'après l'équation (2.3), l'enveloppe complexe du signal MDP-M par rapport à la fréquence porteuse f_0 a pour expression :

$$x_b(t) = \sum_k a_k h(t - kT) \tag{2.31}$$

On suppose que $h(t)$ est *réelle*. On rappelle que l'alphabet (ou constellation) est constitué de $M = 2^m$ points régulièrement espacés sur le cercle unité. On peut donc écrire $a_n = \exp(j\phi_n)$ où $\phi_n \in \{0, 2\pi/M, \cdots, 2\pi(M-1)/M\}$. En supposant que ϕ_n est une suite de variables aléatoires indépendantes et réparties uniformément c'est-à-dire $\mathbb{P}\{a_n = \exp(2j\pi k/M)\} = 1/M$, on déduit que $\mathbb{E}\{a_n\} = 0$ et que $\mathbb{E}\{a_n a_k^*\} = \delta(n-k)$.

Annulation de l'IES : critère de Nyquist

Choisissons le filtre $h(t)$ en bande de base de telle façon que l'impulsion $p(t) = h(t) \star h^*(-t)$ vérifie le critère de de Nyquist à savoir $p(mT) = 0$ pour tout $m \neq 0$. Typiquement on prend pour $p(t)$ une fonction en cosinus sur-élevé. La valeur α du roll-off est déterminée en fonction de la bande du canal. Supposons que le canal puisse être considéré comme un *filtre passe-bande idéal de bande B autour de la fréquence f_0*. Pour voir comment ce filtre agit sur le signal il suffit de déterminer le filtre équivalent en bande de base qui lui correspond (voir chapitre 1) et d'appliquer ce filtre au signal $h(t)$. On rappelle que le filtre équivalent en bande de base s'obtient en transaltant de $-f_0$ la partie du gain située dans les fréquences positives. Par conséquent le filtre équivalent en bande de base du filtre de canal est le filtre passe-bas idéal de bande $(-B/2, B/2)$. Pour que l'impulsion en cosinus sur-élevé ne soit pas distordue, il faut que α vérifie :

$$\frac{1}{2T}(1+\alpha) = \frac{B}{2} \tag{2.32}$$

soit $\alpha = \frac{B}{R} - 1$ où R désigne le débit symbole. Il s'en suit que :

$$R < B \tag{2.33}$$

constitue une condition nécessaire à l'absence d'interférence entre symboles. En comparer cette condition à la condition (2.19), établie pour les modulations en bande de base, on voit que la bande nécessaire est deux fois plus grande : ce résultat est à rapprocher du rapport 2 entre la bande occupée en transmission en bande de base et la bande occupée en modulation double bande. On pourrait, comme en modulation analogique BLU, diviser la bande par 2 en filtrant le signal MDP.

Dans la suite nous considérons uniquement le cas d'une MDP double bande sans IES. De l'expression (2.32) on déduit que l'*efficacité spectrale* a pour expression :

$$\eta = \frac{D_b}{B} = \frac{\log_2(M)}{1+\alpha} \tag{2.34}$$

En échantillonnant, en sortie du filtre $h_R(t)$ de réception, aux instants $nT + \tau$, on obtient un échantillon complexe qui ne dépend que d'un *seul* symbole et dont les parties réelle et imaginaire ont pour expressions respectives (on a supposé $p(t)$ réel) :

$$\begin{cases} Y_n^r = p(0)\cos(\phi_n) + w_n^r & \text{composante en phase} \\ Y_n^i = p(0)\sin(\phi_n) + w_n^i & \text{composante en quadrature} \end{cases} \tag{2.35}$$

où ϕ_n est la phase associée au n-ième symbole et où

$$p(0) = \int_{-\infty}^{+\infty} |h(t)|^2 dt$$

Les composantes w_n^r et w_n^i sont deux variables aléatoires, gaussiennes, centrées, indépendantes de même variance :

$$\sigma^2 = N_0 \int_{-\infty}^{+\infty} |h_R(t)|^2 dt = N_0 p(0)$$

Partant de là, conditionnellement à l'émission du symbole ϕ, l'observation (Y_n^r, Y_n^i) est un vecteur aléatoire gaussien de dimension 2, de moyenne $p(0)\cos(\phi), p(0)\sin(\phi)$ et de matrice de covariance $\sigma^2 I_2$. Leur densité s'écrit :

$$p_{Y_n^r Y_n^i|\phi}(y_r, y_i) = \frac{1}{2\pi\sigma^2} \exp\left\{ -\frac{1}{2\sigma^2} \left((y_r - p(0)\cos(\phi))^2 + (y_i - p(0)\sin(\phi))^2 \right) \right\}$$

Une détection symbole par symbole est possible. Elle consiste à tester l'appartenance de $Z_n = (Y_n^r, Y_n^i)$ à l'un des M secteurs angulaires centrés sur les points de la constellation (voir figure 2.22 le cas $M = 8$). En notant Λ_k le secteur associé au symbole $Ae^{2j\pi k/M}$, la probabilité d'erreur par symbole s'écrit :

$$\begin{aligned} P_e &= 1 - \frac{1}{M}\sum_{k=1}^{M} \mathbb{P}\left\{Z_n \in \Lambda_k | \phi_n = 2\pi k/M\right\} \\ &= \frac{1}{M}\sum_{k=1}^{M} \left(1 - \mathbb{P}\left\{Z_n \in \Lambda_k | \phi_n = 2\pi k/M\right\}\right) \\ &= \frac{1}{M}\sum_{k=1}^{M} \mathbb{P}\left\{Z_n \notin \Lambda_k | \phi_n = 2\pi k/M\right\} \\ &\lessapprox \frac{2}{M}\sum_{k=1}^{M} \mathbb{P}\left\{Z_n \notin D_k | \phi_n = 2\pi k/M\right\} \qquad (2.36) \\ &= 2\mathbb{P}\left\{Z_n \notin D_0 | \phi_n = 0\right\} \end{aligned}$$

où D_k désigne le demi-disque dont le diamètre est la médiatrice entre les points M_k et M_{k-1} de la constellation (points associés respectivement aux phases $2\pi k/M$ et $2\pi(k-1)/M$ (voir figure 2.22) et qui contient le point de coordonnées $(p(0)\cos(2\pi k/M), p(0)\sin(2\pi k/M))$. Notons qu'avec ces notations $\bar{D}_0 = D_{M-1}$

L'approximation dans l'équation (2.36) est justifiée dès lors que M est grand et que le rapport signal sur bruit est grand : la région qui est comptée deux fois (voir figure 2.22) étant de probabilité négligeable. On peut à présent calculer $\mathbb{P}\left\{Z_n \in D_0 | \phi_n = 0\right\}$. Il vient :

$$\begin{aligned} \mathbb{P}\left\{Z_n \notin D_0 | \phi_n = 0\right\} &= \iint_{D_1} \frac{1}{2\pi\sigma^2} \exp\left\{-\frac{(y_r - p(0))^2 + y_i^2}{2\sigma^2}\right\} dy_r dy_i \\ &= Q\left(\frac{d}{2\sigma}\right) \end{aligned}$$

où $\sigma^2 = N_0 p(0)$ et où d représente la distance euclidienne entre le point de coordonnées $(p(0), 0)$ et le point de coordonnées $(p(0)\cos(2\pi/M), p(0)\sin(2\pi/M))$, soit $d = 2p(0)\sin(\pi/M)$.

D'après la propriété 5 énoncée paragraphe 1.7.4, l'énergie par symboles associée au signal transmis est la moitié de l'énergie par symboles associée à l'enveloppe complexe et donc :

$$E_s = \frac{1}{2}\int_{\mathbb{R}} |h(t)|^2 dt = \frac{p(0)}{2}$$

On en déduit que l'énergie par bit a pour expression $E_b = p(0)/2\log_2(M)$. En portant dans l'expression de P_e, on obtient :

$$P_e \approx 2Q\left(\sqrt{\frac{p(0)\sin^2(\pi/M)}{N_0}}\right) = 2Q\left(\sqrt{\frac{2E_b}{N_0}\sin^2(\pi/M)\log_2(M)}\right)$$

En utilisant un codage de Gray et si le rapport signal/bruit est suffisamment grand, on obtient pour le taux d'erreur par élément binaire :

$$\text{TEEB} \approx \frac{2}{\log_2(M)}\, Q\left(\sqrt{\frac{2E_b}{N_0}\sin^2(\pi/M)\log_2(M)}\right) \qquad (2.37)$$

Pour $M = 2, 4$ on peut montrer que le TEEB est précisément donné par :

$$\text{TEEB} = Q\left(\sqrt{\frac{2E_b}{N_0}}\right) \qquad (2.38)$$

Nous avons représenté figure 2.22, le résultat d'une simulation : les points ($\Diamond$) indiquent 500 observations obtenues en MDP-8 lors de l'émission du symbole 1 pour un rapport signal sur bruit $E_b/N_0 = 7$ dB. On voit qu'une vingtaine de points sont en dehors de la région Λ_0, points qui conduisent donc à une décision erronée. Cela donne comme estimation de la probabilité d'erreur par symbole $P_e \approx 0.05$ et donc pour le TEEB ≈ 0.015. On peut vérifier que le résultat obtenu est en accord avec la valeur théorique donnée figure 2.23 où nous avons représenté le TEEB d'une MDP-M en fonction du rapport signal sur bruit E_b/N_0 en dB et pour différentes valeurs de M.

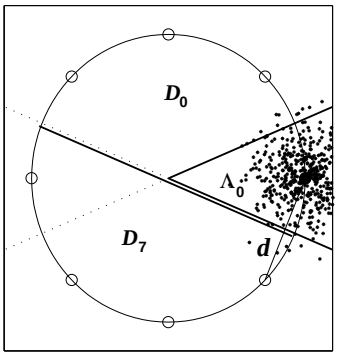

FIG. 2.22 – *Constellation MDP-8. Les deux demi-droites en trait plein représentent les séparatrices de la région Λ_0 de décision de $\phi_n = 0$. Les deux demi-droites en pointillé représentent la partie comptée deux fois dans la borne de* (2.36)*. Le diamètre en trait plein représente la séparatrice entre D_0 et D_7.*

Efficacité spectrale/rapport signal sur bruit

– D'après la formule (2.34) et pour $\alpha = 0$ (cas le plus favorable) on a :

$$\eta = \log_2(M) \tag{2.39}$$

soit

M	2	4	8	16
$\eta\,(bits/s/Hz)$	1	2	3	4

– Le tableau 2.4 donne les valeurs du rapport signal sur bruit relevées sur les courbes de la figure 2.23 pour TEEB $= 10^{-5}$. Δ représente l'écart en dB par rapport à la MDP-2.

M	2	4	8	16
$E_b/N_0\,(dB)$	9.6	9.6	13	17.4
$\Delta\,(dB)$	0	0	3.4	7.8

TAB. 2.4 – *Rapport signal sur bruit en MDP-M pour un TEEB* $= 10^{-5}$.

.

Nous avons reporté figure 2.20 les positions (▼) du rapport signal sur bruit en fonction de l'efficacité spectrale. Les valeurs correspondent TEEB $= 10^{-5}$.

2.5 Exercices

Exercice 2.1 *Qu'appelle-t-on débit binaire ? débit symbole ? Quelle relation y a-t-il entre ces deux quantités ?*

Exercice 2.2 *Décrire la modulation MIA-4 utilisant une impulsion rectangulaire NRZ.*

Exercice 2.3 *On considère le signal de transmission numérique $x(t) = \sum_k a_k g(t - kT)$ où $g(t)$ est une impulsion et a_k une suite de symboles (réels ou complexes) supposée stationnaire au second ordre. On note $m_a = \mathbb{E}\{a_n\}$ et $R_a(k) = \mathbb{E}\{(a_{n+k} - m_a)(a_n - m_a)^*\}$.*

– *Donner une condition nécessaire pour que le spectre de $x(t)$ contienne des composantes harmoniques aux fréquences multiples de $1/T$? Quel intérêt cela a-t-il ?*

– *Donner une condition suffisante pour que le spectre de $x(t)$ soit nul autour de la fréquence 0. Quel intérêt présente cette propriété ?*

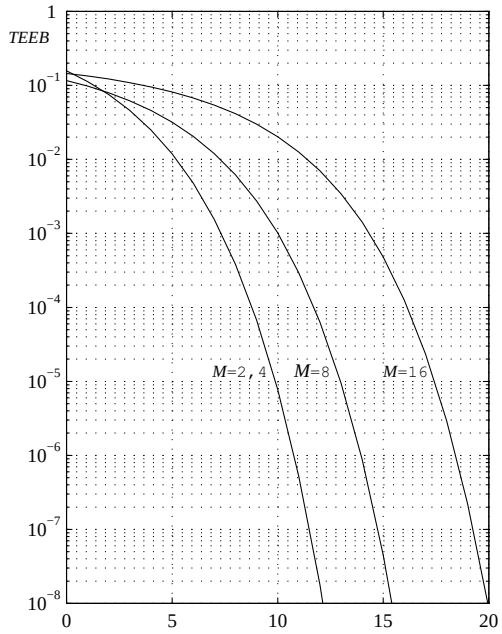

FIG. 2.23 – *TEEB des modulations MDP-M, en fonction du rapport E_b/N_0 en dB, pour $M = 2$, 4, 8 et 16.*

Exercice 2.4 *Qu'appelle-t-on l'interférence entre symboles ? Comment peut-on s'en affranchir ?*

Exercice 2.5 *Donner le critère de Nyquist. Quelle condition sur le débit symbole et la bande impose ce critère ?*

Exercice 2.6 (Fonctions vérifiant le critère de Nyquist) *1. On considère le signal $y(t)$ dont la transformée de Fourier s'écrit $Y(f) = X(f) * R(f)$ où $R(f) = C \times \mathbb{1}_{(-1/2T,1/2T)}(f)$. On rappelle que la transformée de Fourier inverse de $R(f)$ est $r(t) = C\sin(\pi t/T)/\pi t$. Quelle propriété remarquable a $y(t)$?*

2. *On considère $X(f) = 2\cos(\pi f/2b)\mathbb{1}(f \in (-b, b))$. Déterminer l'expression de $x(t)$ (noter que $x(t)$ est la somme de deux exponentielles complexes).*
3. *On pose $b = \alpha/2T$, où $0 < \alpha < 1$. Dans la suite $C = T\pi/8b$. Déterminer, en fonction de α et de T, l'expression de $Y(f) = X(f) * R(f)$.*

Exercice 2.7 *Qu'appelle-t-on le diagramme de l'oeil ? Quelle particularité présente ce diagramme quand l'IES=0 ?*

Exercice 2.8 *Décrire en fréquence les fonctions dites en cosinus-surélevé. Qu'appelle-t-on le "rolloff" ? Comment le "rolloff" intervient-il sur le diagramme de l'oeil ?*

Exercice 2.9 *Qu'appelle-t-on la distorsion maximale ? Montrer que, pour les impulsions en cosinus-surélevé, la distorsion maximale est nulle si on choisit correctement les instants d'échantillonnage en sortie du filtre de réception.*

Si on s'écarte de ces instants, comment varie la distorsion maximale en fonction du "rolloff" ? Que pensez-vous du cas où le "rolloff" est nul ?

Exercice 2.10 *On considère une transmission binaire dont la réponse impulsionnelle globale est représentée à la figure ci-dessous. La transmission est-elle sans IES ? Tracer le diagramme de l'oeil correspondant.*

Exercice 2.11 *Décrire la chaîne d'une transmission binaire, sans IES, sur un canal à bande limitée (en bande de base) et soumis à un bruit additif gaussien et blanc. On précisera les filtres d'émission et de réception.*

Exercice 2.12 *Qu'appelle-t-on taux d'erreur par éléments binaires ? taux d'erreur par symbole ? Quelle relation simple lie ces 2 quantités ?*

Exercice 2.13 *Qu'appelle-t-on l'efficacité spectrale ? Donner en fonction de la taille de l'alphabet de modulation et du "rolloff" l'expression de l'efficacité spectrale.*

Exercice 2.14 *Pour une transmission binaire sans IES, donner l'expression des échantillons y_k en sortie du filtre de réception. Déterminer, en fonction de la d.s.p. $N_0/2$ du bruit et du gain du filtre de réception, l'expression de la variance du terme associé au bruit.*

Exercice 2.15 *On considère un canal passe-bas de bande 2400 Hz sur lequel on désire transmettre, sans IES, un débit de 3600 bits/s à partir d'une modulation de modulation utilisant une impulsion NRZ.*

Déterminer les filtres idéaux (IES nulle et minimum de probabilité d'erreur) d'émission et de réception.

Exercice 2.16 *Pour transmettre un débit binaire de 2 Mbits/s, on utilise une modulation MIA à 4 états. La probabilité d'erreurs par bits est de 10^{-5}. On désire, en conservant une modulation de type MIA, doubler le débit binaire, sur ce même canal (et donc avec la même bande), sans toucher à la probabilité d'erreurs par bits.*

1. *Quelle doit être la nouvelle valeur de la taille de l'alphabet de modulation ?*
2. *De combien de dB faut-il augmenter la puissance nécessaire à l'émission ?*

On donne pour une probabilité d'erreurs par bits $P_e = 10^{-5}$:

M	2	4	8	16
$E_b/N_0(dB)$	9.6	13.6	18	22.9

(2.40)

Exercice 2.17 (Transmission en bande de base) *On considère une source binaire, indépendante, équirépartie, dont le débit est de 10000 bits/s. Le signal émis a pour expression :*

$$s(t) = \sum_k a_k h(t - kT)$$

où les a_k sont une suite de symboles à valeurs dans l'alphabet $\{-3, -1, +1, +3\}$. T désigne l'intervalle de temps entre deux symboles consécutifs. $h(t)$ est un signal réel, d'énergie finie, dont la transformée de Fourier $H(f)$ est nulle à l'extérieur de la bande de fréquence $(-B, B)$. On donne $B = 3000Hz$. Le bruit est une réalisation d'un processus blanc, gaussien, centré, de densité spectrale de puissance $N_0/2$. On note $z(t) = s(t) + b(t)$ le signal reçu.

1. *Calculer la vitesse de modulation $1/T$.*
2. *Indiquer un codage qui minimise le taux d'erreur par bits.*
3. *Donner la fonction de transfert du filtre de réception qui minimise la probabilité d'erreur.*
   On note $P(f)$ la cascade constituée par le filtre d'émission $H(f)$ et le filtre de réception et $p(t)$ sa transformée de Fourier inverse.
4. *A quelle condition sur $p(t)$ a-t-on absence d'interférence entre symboles à la sortie de l'échantillonneur.*
5. *On utilise pour $P(f)$ une fonction en cosinus surélevé. Quelle valeur du taux de débordement faut-il choisir ?*
6. *Calculer, en fonction de $p(t)$, l'énergie moyenne E_s par symbole.*
7. *En déduire l'énergie moyenne E_b par bit.*
8. *Montrer que la variance du bruit à la sortie de l'échantillonneur est égale à $N_0 p(0)/2$.*
9. *Indiquer la règle de l'organe de décision.*
10. *Déterminer l'expression de la probabilité d'erreur par symboles, en fonction de E_b et de N_0.*
11. *En déduire la probabilité d'erreur par bits.*
12. *Y a-t-il des raies spectrales dans le signal émis ? Justifier la réponse.*

Exercice 2.18 (Variance de l'IES) *On considère, dans une transmission numérique, que le signal en sortie du filtre de réception a pour expression* $r(t) = \sum_{k\in\mathbb{Z}} a_k p(t-kT)$ *où* a_k *est une suite de variables aléatoires à valeurs dans un alphabet* M*-aire et telle que* $m_a = \mathbb{E}\{a_k\} = 0$ *et* $R_a(k) = \mathbb{E}\{a_{n+k}a_n^*\}$*. On note* $S_x(f)$ *le spectre de* $x(t)$*. On suppose que la transmission est sans bruit.*

On échantillonne le signal $r(t)$ *aux instants* $nT+\tau$*, ce qui donne, d'après l'expression* (2.16) *:*

$$r(nT+\tau) = a_n p(0) + \underbrace{\sum_{k\neq n} a_k p((n-k)T)}_{IES\ :\ \epsilon_n}$$

où ϵ_n *représente le terme dû à l'interférences entre symboles (IES) dont nous allons calculer la moyenne et la variance.*

1. *Montrer que* ϵ_n *s'exprime sous la forme d'un filtrage linéaire :*

$$\epsilon_n = \sum_{k\in\mathbb{Z}} a_k g(n-k)$$

dont on déterminera la réponse impulsionnelle g_n *et le gain en fréquence en fonction de la transformée de Fourier* $P(f)$ *de* $p(t)$ *et de* $p(0)$*.*

On rappelle que si u_n *désigne un processus aléatoire stationnaire au second ordre de moyenne* m_u *et de d.s.p.* $S_u(f)$ *à l'entrée d'un filtre numérique de réponse impulsionnelle* h_n *et de gain en fréquence* $H(f) = \sum_n h_n e^{-2j\pi nf}$ *alors le processus en sortie* v_n *a pour moyenne* $m_v = m_u H(0)$ *et pour d.s.p. :*

$$S_v(f) = |H(f)|^2 S_u(f)$$

Enfin on a $R_v(n) = \int_{-1/2}^{1/2} S_v(f) e^{2j\pi fn} df$*.*

2. *Déterminer* $\mathbb{E}\{\epsilon_n\}$*.*
3. *Déterminer l'expression de* $\mathrm{var}(\epsilon_n)$ *en fonction de* $S_x(f)$ *(utiliser la formule de Poisson).*
4. *Dans le cas d'une transmission avec une impulsion en cosinus surélevé, comment se comporte* $\mathrm{var}(\epsilon_n)$ *en fonction du facteur de roll-off ?*

Exercice 2.19 (Transmission sur porteuse) *On utilise pour transmettre une source binaire, de débit* $D_b = 30000$ *bits/s, une modulation MDP-8. On suppose que les symboles sont indépendants et équiprobables. Le canal est soumis à un bruit additif, gaussien, centré, blanc de densité spectrale de puissance* $N_0/2$*. On note* $s(t)$ *le signal transmis et* $z(t) = s(t) + b(t)$ *le signal reçu.*

1. *Quel est le débit symbole* $R = 1/T$ *?*
2. *Donner l'expression du signal modulé* $s(t)$ *ainsi que celle de son enveloppe complexe (on précisera l'alphabet de modulation, ainsi que la forme de l'impulsion).*
3. *Indiquer un codage qui minimise le taux d'erreur par éléments binaires.*
4. *Donner le schéma de principe du détecteur qui fournit les composantes en phase et quadrature du signal reçu* $z(t)$*. On note respectivement* $z_p(t)$ *et* $z_q(t)$ *ces deux composantes.*
5. *Donner l'expression du filtre de réception qui minimise la probabilité d'erreur.*
6. *Montrer que l'expression des échantillons, observés à la cadence* $1/T$*, en sortie du filtre de réception s'écrit :*

$$r_k = d_k p(0) + w_k$$

7. *Calculer, en fonction de* $p(t)$*, l'énergie moyenne* E_s *par symbole.*
8. *En déduire l'énergie moyenne* E_b *par bit.*
9. *On pose* $w_k = p_k + jq_k$*. Montrer que* p_k *et* q_k *sont deux variables aléatoires gaussiennes, centrées, de même variance* $N_0 p(0)$*, indépendantes.*
10. *Indiquer la règle de l'organe de décision.*
11. *Donner sous forme d'une intégrale la probabilité d'erreur par symbole.*

Exercice 2.20 (Choix de la modulation MDP-M) *On souhaite transmettre un débit de* 4.8 *kbits/s à travers un canal dont la bande est* 3 *kHz autour de la fréquence* $f_0 = 1.8kHz$*. Pour cela on utilise une MDP-M filtré.*

1. *Décrire la modulation. Donner l'expression du signal en bande de base.*
2. *Comment choisir M pour satisfaire une transmission sans IES ?*
3. *La puissance transmise est 2 W et la d.s.p. du bruit blanc sur le canal est de 0.3 mW/Hz. Quel est le taux d'erreurs par élément binaire ?*

Exercice 2.21 *On désire transmettre en MDP-M, simultanément et avec le même TEEB, deux trains numériques de débits respectifs 4 Mbits/s et 8 Mbits/s.*

1. *Sachant que l'on sait réaliser des filtres de Nyquist de roll-off $\alpha = 0.5$, donner pour chacun des deux débits les bandes nécessaires en MDP-M pour $M = 2$, 4, 8 et 16.*
2. *Le canal a une bande de 9 MHz. En s'aidant du tableau 2.4, déterminer quelle est la façon la plus économique, en terme de puissance moyenne transmise ?*
3. *Même question avec un canal de bande de 7.5 MHz.*

Exercice 2.22 (MDP-8) *On considère un modulation MDP-8 points. On suppose que la transmission se fait sur un canal idéal de bande B autour d'une fréquence f_0 et que le bruit est AGB de densité spectrale $N_0/2$.*

1. *On note $h(t)$ l'impulsion de modulation et $p(t) = h(t) \star h^*(-t)$. Rappeler l'expression de l'enveloppe complexe par rapport à f_0 ainsi que celle du signal modulé.*
2. *Décrire un dispositif qui donne les composantes en phase et en quadrature.*
3. *Décrire la forme du récepteur optimal.*
4. *Donner l'expression des parties réelle et imaginaire des échantillons Y_n en sortie du filtre de réception.*
5. *On suppose que la transmission se fait sans IES. Donner les propriétés de l'observation en sortie du filtre de réception.*
6. *Donner la forme de l'organe de décision.*
7. *Donner l'expression de l'énergie moyenne par bit en fonction de $p(0)$ et de $M = 8$.*
8. *On rappelle que les parties réelle et imaginaire du bruit sont centrées, indépendantes de même variance $\sigma^2 = N_0 p(0)$. Partant de l'expression approchée $P_e = 2Q(d/2\sigma)$ déterminer le TEEB en fonction de E_b/N_0.*

2.6 Annexes

2.6.1 Preuve de (2.7)

Partant $\mathbb{E}\{x(t+\tau)x^*(t)\}$ on a :

$$\begin{aligned}
\mathbb{E}\{x(t+\tau)x^*(t)\} &= \sum_k \sum_m \mathbb{E}\{a_k h(t+\tau-kT+U)a_m^* h^*(t-mT+U)\} \\
&= \sum_k \sum_m \mathbb{E}\{a_k a_m^*\}\,\mathbb{E}\{h(t+\tau-kT+U)h^*(t-mT+U)\} \\
&\qquad\qquad \text{par indépendance de } U \text{ et des } \{a_k\} \\
&= \sum_k \sum_m \mathbb{E}\{a_k a_m^*\} \int_0^T \frac{1}{T} h(t+\tau-kT+u)h^*(t-mT+u)du \\
&\qquad\qquad \text{en utilisant que la loi de } U \text{ est uniforme} \\
&= \frac{1}{T}\sum_k \sum_m (R_a(k-m)+|m_a|^2) \int_{t+\tau-kT}^{t+\tau-kT+T} h(v)h^*(v-(\tau+(m-k)T))dv \\
&= \frac{1}{T}\sum_k \sum_p (R_a(p)+|m_a|^2) \int_{t+\tau-kT}^{t+\tau-kT+T} h(v)h^*(v-(\tau-pT))dv \\
&= \frac{1}{T}\sum_p (R_a(p)+|m_a|^2) \int_{-\infty}^{+\infty} h(v)h^*(v-(\tau-pT))dv \\
&= \frac{1}{T}\sum_p (R_a(p)+|m_a|^2) c_{hh}(\tau-pT) \\
&= \underbrace{\frac{1}{T}\sum_p R_a(p)c_{hh}(\tau-pT)}_{R_x^a(\tau)} + \underbrace{\frac{1}{T}|m_a|^2 \sum_p c_{hh}(\tau-pT)}_{R_x^d(\tau)}
\end{aligned}$$

où nous avons posé $c_{hh}(u) = h(u) \star h^*(-u)$ dont la transformée de Fourier est $C_{hh}(f) = |H(f)|^2$. Pour obtenir le spectre on traite séparément les deux quantités $R_x^a(\tau)$ et $R_x^d(\tau)$.

$R_x^a(\tau)$ possède une transformée de Fourier donnée par :

$$S_x^a(f) = \frac{1}{T}\sum_p R_a(p)C_{hh}(f)e^{-2j\pi pfT} = \frac{|H(f)|^2}{T}\sum_p R_a(p)e^{-2j\pi pfT}$$

$R_x^d(\tau)$ apparaît comme la somme des décalées d'une même fonction $c_{hh}(t)$. C'est donc une fonction périodique de période T, qui, d'après la formule de Poisson, s'écrit :

$$R_x^d(\tau) = \frac{|m_a|^2}{T^2}\sum_k C_{hh}(k/T)e^{-2j\pi k\tau/T} = \frac{|m_a|^2}{T^2}\sum_k |H(k/T)|^2 e^{-2j\pi k\tau/T}$$

$R_x^d(\tau)$ est un mélange harmonique. Son spectre est donc constitué de raies aux fréquences k/T et d'amplitudes respectives $|m_a|^2|H(k/T)|^2/T^2$.

Formule de Poisson

Soit $\phi(t)$ un signal de module intégrable et dont la transformée de Fourier $\Phi(f)$ est elle-même de module intégrable. On a alors pour tout $T > 0$:

$$\sum_{n=-\infty}^{+\infty} \phi(t-nT) = \frac{1}{T}\sum_{k=-\infty}^{+\infty} \Phi(k/T)e^{-2j\pi kt/T} \tag{2.41}$$

Chapitre 3

Introduction aux codes correcteurs d'erreur

3.1 Canal binaire symétrique sans mémoire

Considérons la chaîne de transmission représentée figure 3.1. La modulation est une modulation binaire d'impulsions en amplitude MIA-2. Le canal est supposé idéal, de bande B et soumis à un bruit additif, gaussien, blanc. On suppose que l'interférence entre symboles est nulle aux instants d'échantillonnage (l'ensemble des filtres d'émission et de réception vérifie le critère de Nyquist) et que la probabilité d'erreur est minimale (le filtre de réception est donc le filtre adapté).

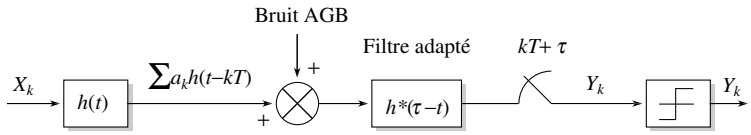

FIG. 3.1 – *chaîne de transmission binaire sur un canal de bande B, soumis à un bruit additif, gaussien, blanc. X_k est une suite de variables aléatoires i.i.d. à valeurs dans $\{0,1\}$ et $a_k = 2X_k - 1$. $h(t)$ est, par exemple, une racine carrée d'un filtre en cosinus-surélevé.*

En considérant le système de transmission de bout en bout, on voit que l'entrée X_k prend ses valeurs dans l'alphabet d'entrée $\mathbb{F}_2 = \{0,1\}$ de taille 2 et que la sortie Y_k du comparateur à seuil prend ses valeurs dans l'alphabet de sortie $\mathbb{F}_2 = \{0,1\}$ de taille 2. On a vu chapitre 2 que, dans les conditions d'une réception idéale sur le canal de Nyquist, la probabilité d'erreurs a pour expression :

$$p = \mathbb{P}\{Y_k = 1 | X_k = 0\} = \mathbb{P}\{Y_k = 0 | X_k = 1\} = Q(\sqrt{2E_b/N_0})$$

où E_b désigne l'énergie par bit et $N_0/2$ la densité spectrale de puissance du bruit blanc, gaussien. Notons que la valeur p ne dépend pas de l'instant k d'utilisation du canal et que, lors d'utilisations successives du canal, les erreurs sont indépendantes : on dit alors que le canal est sans mémoire.

On peut alors représenter la chaîne de transmission entre X_k et Y_k par le schéma de la figure 3.2 que l'on désigne sous le nom de *canal binaire symétrique sans mémoire* (en abrégé CBS). En principe (nous y reviendrons chapitre 4) les alphabets d'entrée et de sortie ne doivent pas être confondus même si, comme c'est le cas ici, les deux alphabets ont la même taille et que, par souci de simplicité, on note les éléments de la même manière.

Tout se passe comme si X_k était soumis à un bruit additif et que l'observation Y_k s'écrivait :

$$Y_k = X_k \oplus B_k$$

où $\oplus$ désigne l'addition modulo 2 (OU exclusif) dans $\mathbb{F}_2 = \{0,1\}$. Le bruit B_k est une suite de variables aléatoires, à valeurs dans $\{0,1\}$, indépendantes et identiquement distribuées telles que $\mathbb{P}\{B_k = 1\} = p$. On peut aussi voir le CBS suivant le schéma de la figure 3.3.

FIG. 3.2 – *Canal binaire symétrique. $q = 1 - p$.*

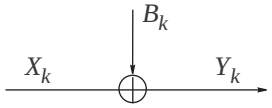

FIG. 3.3 – *Canal CBS vu comme un canal additif (modulo 2) binaire : $Y_k = X_k \oplus B_k$.*

L'utilisation d'une décision dure dans une chaîne de transmission, comme celle construite autour d'une modulation MIA-2, dégrade les performances. Cependant dans beaucoup de systèmes pratiques, par souci de réduction de la complexité, on sépare souvent les opérations de modulation et de codage. En se limitant au cas des entrées binaires, on considère alors que le canal est un canal binaire symétrique sans mémoire dont la probabilité d'erreur est p et on s'intéresse aux performances des codes construits sur ce canal. Ce paragraphe donne une brève introduction aux codes correcteurs sur le canal binaire symétrique sans mémoire.

3.2 Différents types de code

Nous verrons chapitre 4, dans le cadre de la théorie de l'information, la définition générale d'un canal ainsi que celle d'un code (définitions 4.1 et 4.3). Dans ce chapitre nous nous intéressons uniquement aux canaux à entrée binaire, tel que le CBS, et nous adoptons la définition suivante : un codeur sur le canal binaire est un dispositif qui associe à une suite de k bits d'information une suite de n éléments binaires (écrite avec l'alphabet d'entrée du canal). L'ensemble des 2^k mots ainsi construits par le codeur et appartenant à $\mathbb{F}_2^n$ est appelé le *code*. Nous le notons $\mathcal{C}(n, k)$ où k représente le nombre de *bits d'information*, $(n - k)$ le nombre de *bits de redondance* et le rapport :

$$r = \frac{\log_2(M)}{n} = \frac{k}{n}$$

le *taux de codage* (en anglais *code rate*).

Exemple 3.1 (Code à répétitions sur le canal CBS) *On considère le code défini par :*

$$\mathcal{C} : \begin{array}{l} 0 \mapsto 000 \in \mathbb{F}_2^3 \\ 1 \mapsto 111 \in \mathbb{F}_2^3 \end{array}$$

Réponse : Avec les notations précédentes on a $n = 3$ et $k = 1$. Pour des raisons évidentes ce code est dit à répétitions. Son taux de codage est égal à $1/3$.

En pratique on s'intéresse principalement aux codes *linéaires* : il sont tels que, si c_1 et c_2 désignent les deux mots-code respectifs des deux suites de k bits d_1 et d_2, alors à la suite de k bits $d_1 \oplus d_2$ est associé le mot-code $c_1 \oplus c_2$. L'opération $\oplus$ est l'addition bit à bit modulo 2 dans $\mathbb{F}_2$. On vérifie aisément que le code à répétitions est linéaire.

Deux grandes familles de codes linéaires existent : les codes en bloc et les codes convolutifs. En fait, comme nous allons le voir, les codes en bloc ne sont qu'un cas particulier des codes convolutifs.

Code en bloc

Dans un code en bloc, les n éléments binaires des mots-code sont calculés uniquement avec les k bits d'information du *bloc* courant suivant le schéma de codage :

$$\underbrace{[d_1^m \ \cdots \ \cdots \ \cdots \ \cdots \ d_k^m]}_{\text{bloc numéro } m\text{: } k \text{ bits d'information}}$$

$$\downarrow$$

$$\underbrace{[c_1^m \ \cdots \ \cdots \ \cdots \ \cdots \ \cdots \ c_n^m]}_{n \text{ éléments binaires du mot-code}}$$

où m désigne le numéro du bloc courant. Dans la suite, quand il n'y aura pas d'ambiguïté, nous omettrons l'indice m et nous noterons le bloc courant d'information d et le mot-code associé c.

On vérifie aisément que, pour un code en bloc linéaire, le mot-code c s'obtient à partir du mot d'information d par une expression matricielle de la forme :

$$c = dG$$

où d est un vecteur ligne de dimension $1 \times k$ et G une matrice de dimensions $k \times n$ appelée matrice génératrice du code. Dans le cas binaire, qui est le seul que nous traiterons, la matrice G est constituée de 0 et de 1 et les sommes sont calculées modulo 2. Nous y reviendrons plus en détails paragraphe 3.4. En voici deux exemples simples.

Exemple 3.2 (Code à répétitions $(n, 1)$**)** *Nous avons déjà présenté ce code dans l'exemple 3.1. Ce code associe respectivement à* $k = 1$ *bit d'information les deux mots-code suivants :*

$$\begin{aligned} d = 0 \quad &\mapsto \quad c = \underbrace{0 \quad \cdots \quad 0}_{n} \\ d = 1 \quad &\mapsto \quad c = \underbrace{1 \quad \cdots \quad 1}_{n} \end{aligned}$$

Déterminer sa matrice génératrice et son taux de codage.

Réponse : On vérifie aisément que la matrice génératrice de ce code est :

$$G = [\underbrace{1 \quad \cdots \quad 1}_{n}]$$

Son taux de codage est $1/n$.

Exemple 3.3 (Code parité $(n, n-1)$**)** *Ce code ajoute un bit, appelé bit de parité, à une suite de* $k = (n-1)$ *bits d'information de façon à ce que le nombre total de* 1 *du mot-code ainsi formé soit pair. On a :*

$$\begin{aligned} c_j &= d_j, \textit{pour } j = 1, \cdots, n-1 \\ c_n &= d_1 \oplus d_2 \oplus \cdots \oplus d_{n-1} \end{aligned}$$

Déterminer sa matrice génératrice et son taux de codage.

Réponse : Le code de parité a pour matrice génératrice :

$$G = \begin{bmatrix} & & 1 \\ & I_{n-1} & \vdots \\ & & 1 \end{bmatrix}$$

où I_{n-1} désigne la matrice identité de taille $n-1$. Le taux de codage est $(n-1)/n$.

3.3 Décision optimale sur le canal CBS

Une fois codés les bits d'information sont transmis sur le canal. A la réception, partant des mots-code reçus et comportant, à cause du bruit, des erreurs, il faut trouver une règle de décision qui optimise un critère comme par exemple celui de maximiser la probabilité de décision correcte. Ce problème entre dans le cadre de la théorie de la décision statistique que nous avons présenté section 1.8.

3.3.1 Un exemple

Avant de donner la solution générale de la règle de décision optimale sur le CBS, commençons par montrer que le décodage "optimal" du code à répétitions (présenté dans l'exemple 3.1) sur le canal binaire symétrique, se présente comme un problème de décision statistique : à l'entrée du canal on transmet, de façon équiprobable, soit le message 0 soit le message 1. Ici $M = 2$. En sortie du canal, on reçoit un des 8 éléments $y \in \mathbb{F}_2^3$ constitués de 3 composantes (y_1, y_2, y_3) à valeurs dans $\mathbb{F}_2$. En utilisant l'hypothèse que le canal est symétrique et sans mémoire, on déduit que :

$$\mathbb{P}\{Y = y|X = 0\} = (1-p)^{3-d_H(y,000)} p^{d_H(y,000)} \tag{3.1}$$

$$\text{et} \qquad \mathbb{P}\{Y = y|X = 1\} = (1-p)^{3-d_H(y,111)} p^{d_H(y,111)} \tag{3.2}$$

où $d_H(y, c)$ désigne le nombre de bits, en même position, qui diffèrent entre y et c.

Le décodage se fait au moyen d'une fonction de décision caractérisée par la partition de $\mathbb{F}_2^3$ en deux régions Λ_0 et Λ_1 telles que, si l'observation $y \in \Lambda_i$, on choisit i. Du fait que les deux régions Λ_0 et Λ_1 réalisent une partition de l'espace d'observation $\mathbb{F}_2^3$ et que les deux messages sont équiprobables, la probabilité de décision correcte *sur les mots-code* s'écrit :

$$P_c = \frac{1}{2}\sum_{y\in\Lambda_0} \mathbb{P}\{Y = y|X = 0\} + \frac{1}{2}\sum_{y\in\Lambda_1} \mathbb{P}\{Y = y|X = 1\}$$

La maximisation de P_c est simple : il suffit de mettre dans Λ_0 tous les points y de $\mathbb{F}_2^3$ tels que :

$$\mathbb{P}\{Y = y|X = 0\} > \mathbb{P}\{Y = y|X = 1\}$$

Ce qui s'écrit :

$$\Lambda_0 = \{y \in \mathbb{F}_2^3 : \mathbb{P}\{Y = y|X = 0\} > \mathbb{P}\{Y = y|X = 1\}\}$$

et de prendre $\Lambda_1 = \mathbb{F}_2^3 - \Lambda_0$. En utilisant les expressions (3.1) et supposant que $p < 1/2$, on en déduit que :

$$\Lambda_0 = \{y \in \mathbb{F}_2^3 : d_H(y, 000) < d_H(y, 111)\} \quad \text{et} \quad \Lambda_1 = \mathbb{F}_2^3 - \Lambda_0$$

Par conséquent, pour un code à répétitions sur le canal CBS, la fonction de décision optimale (dans le sens du maximum de probabilité de décision correcte sur les mots-code) est :

$$\begin{array}{rll} g: & 000 \to 0 & 100 \to 0 \\ & 001 \to 0 & 101 \to 1 \\ & 010 \to 0 & 110 \to 1 \\ & 011 \to 1 & 111 \to 1 \end{array}$$

Dans ce cas la probabilité d'erreur, $P_e = 1 - P_c$, est minimale et a pour expression :

$$\begin{aligned} P_e &= \frac{1}{2}\mathbb{P}\{Y \in \Lambda_0|X = 1\} + \frac{1}{2}\mathbb{P}\{Y \in \Lambda_1|X = 0\} \\ &= \underbrace{\mathbb{P}\{Y = 011|X = 0\}}_{p^2(1-p)} + \mathbb{P}\{Y = 101|X = 0\} \\ &\qquad + \mathbb{P}\{Y = 110|X = 0\} + \mathbb{P}\{Y = 111|X = 0\} \\ &= 3p^2(1-p) + p^3 \approx 3p^2 \end{aligned}$$

3.3.2 Application au CBS : distance de Hamming

Voyons à présent l'expression de la fonction de décision pour un code $\mathcal{C}(n,k)$ utilisé sur le canal binaire symétrique. On note c_i le mot-code associé au i-ème message. Des hypothèses d'absence de mémoire, on déduit que :

$$\begin{aligned}\mathbb{P}\{Y_1 = y_1, \cdots, Y_n = y_n | X_1 = c_1, \cdots, X_n = c_n\} &= \mathbb{P}\{Y = y | X = c\} &(3.3)\\ &= (1-p)^n \left(\frac{p}{1-p}\right)^{d_H(y,c)} &(3.4)\end{aligned}$$

où $d_H(y,c)$ désigne la distance de Hamming entre le mot reçu $(y_1, \ldots, y_n)$ et le mot-code $c \in \mathcal{C}$.

Définition 3.1 (Distance de Hamming) *Soit $x = (x_1, \ldots, x_n)$ et $y = (y_1, \ldots, y_n)$ deux suites binaires appartenant à $\mathbb{F}_2^n$. On appelle distance de Hamming :*

$$d_H(x,y) = \sum_{i=1}^{n} \mathbb{1}(x_i \neq y_i) \quad \textit{(nombre de bits qui diffèrent entre } x \textit{ et } y\textit{)}$$

On vérifie que $d_H(x,y)$ est une distance sur $\mathbb{F}_2^n$. La distance $d_H(x,0) = \pi_H(x)$ s'appelle le poids de x.

D'après l'expression (**??**) et sous l'hypothèse que les mots-code sont équiprobables, les régions de vraisemblance maximale sont données par :

$$\Lambda_i = \{y \in \mathbb{F}_2^n \text{ t.q. } \mathbb{P}\{Y = y | X = i\} \geq \mathbb{P}\{Y = y | X = j\} \ \ \forall i \neq j \in \mathcal{C}\}$$

En supposant que $p < 1/2$ (et donc que $p/(1-p) < 1$), on peut encore simplifier sous la forme :

$$\Lambda_i = \{y \in \mathbb{F}_2^n \text{ t.q. } d_H(y|i) \leq d_H(y|j) \ \ \forall i \neq j \in \mathcal{C}\} \tag{3.5}$$

Règle de décision optimale

D'après (3.5), la règle de décision optimale sur le canal binaire symétrique consiste à associer au mot reçu y le mot-code le plus proche au sens de la distance de Hamming à savoir :

$$\hat{c} = \arg \min_{c \in \mathcal{C}(n,k)} d_H(y,c) \tag{3.6}$$

où y désigne le mot reçu[1].

Insistons tout particulièrement sur le fait que le minimum de distance de Hamming est la règle de décision optimale *sur le canal binaire symétrique*. Sur un autre canal, la règle de décision est a priori différente ; celle-ci doit être établie en utilisant l'expression générale (**??**) après avoir déterminé les expressions de $\mathbb{P}\{Y = y | X = i\}$.

Probabilité d'erreur

En revenant au canal CBS et en utilisant l'expression (3.3), on déduit l'expression de la probabilité d'erreur par mot-code :

$$\begin{aligned}\bar{P}_e(M,n) &= \frac{1}{M} \sum_{i=1}^{M} \sum_{y \in \bar{\Lambda}_i} \mathbb{P}\{Y = y | X = c_i\}\\ &= \frac{1}{M} \sum_{i=1}^{M} \sum_{y \in \bar{\Lambda}_i} (1-p)^{n - d_H(y,c_i)} p^{d_H(y,c_i)} &(3.7)\end{aligned}$$

où $\bar{\Lambda}_i$ désigne le complémentaire de Λ_i par rapport à $\mathbb{F}_2^n$, ainsi que l'expression de la probabilité de décision correcte :

$$\begin{aligned}\bar{P}_c(M,n) &= \frac{1}{M} \sum_{i=1}^{M} \sum_{y \in \Lambda_i} \mathbb{P}\{Y = y | X = c_i\}\\ &= \frac{1}{M} \sum_{i=1}^{M} \sum_{y \in \Lambda_i} (1-p)^{n - d_H(y,c_i)} p^{d_H(y,c_i)} &(3.8)\end{aligned}$$

[1] $\arg f$ désigne l'argument de la fonction f.

Notons que ces expressions se prêtent mal au calcul car les régions Λ_i, qui dépendent du choix du code, ont en général des formes compliquées. C'est pourquoi on s'intéresse, en pratique, à l'obtention de bornes inférieures ou supérieures plus facilement calculables. En particulier on peut montrer (voir annexe 3.6.1) que :

$$\bar{P}_e(M,n) \leq (M-1)[4p\ (1-p)]^{d_{\min}/2} \tag{3.9}$$

où $d_{\min}$ est la distance minimale entre deux mots-code définie par :

Définition 3.2 (Distance minimale d'un code) *Soit un code $\mathcal{C}(n,k)$ sur le canal binaire. On appelle distance minimale du code la quantité :*

$$d_{\min} = \min_{x \neq y \in \mathcal{C}} d_H(x,y)$$

Dans la suite nous noterons un code sous la forme $\mathcal{C}(n,k,d_{\min})$. Le résultat suivant est fondamental :

Propriétés 3.1 (Capacité de correction d'un code) *On considère un code $\mathcal{C}(n,k,d_{\min})$ sur le canal binaire symétrique. On note*[2] *:*

$$t = \left\lfloor \frac{d_{\min}-1}{2} \right\rfloor$$

Alors ce code corrige t erreurs.

En effet, choisissons un entier u qui vérifie $2u+1 \leq d_{\min}$ et supposons que le mot reçu $\mathbf{y}$ comporte $\mathbf{u}$ erreurs avec le mot émis $\mathbf{c}$. Cela s'écrit $d_H(\mathbf{y},\mathbf{c}) = u$. En utilisant l'inégalité triangulaire, on a alors quel que soit le mot-code $\mathbf{c}' \neq \mathbf{c}$, $d_H(\mathbf{y},\mathbf{c}') \geq d_H(\mathbf{c},\mathbf{c}') - d_H(\mathbf{y},\mathbf{c}) \geq d_{\min} - u \geq u+1 > d_H(\mathbf{y},\mathbf{c})$. En conclusion, $\forall \mathbf{c}' \neq \mathbf{c}$, on a $d_H(\mathbf{y},\mathbf{c}') > d_H(\mathbf{y},\mathbf{c})$, ce qui signifie que $\mathbf{y}$ est bien le mot-code le plus proche de $\mathbf{c}$. Par conséquent la règle du maximum de vraisemblance conduit à une décision correcte, à condition que le nombre t d'erreurs soit inférieur au plus grand entier u qui vérifie $2u+1 \leq d_{\min}$, ce qui donne $t = \lfloor (d_{\min}-1)/2 \rfloor$.

Dans le cas où le mot reçu comporte entre $(t+1)$ et $(d_{\min}-1)$ erreurs et donc que $t+1 \leq d_H(\mathbf{y},\mathbf{c}) \leq d_{\min}-1$, alors quel que soit le mot-code $\mathbf{c}' \neq \mathbf{c}$, $d_H(\mathbf{y},\mathbf{c}') \geq 1$ qui signifie que l'observation reçue n'est pas un mot-code. Mais la règle du minimum de distance peut conduire à une décision erronée, puisqu'il peut exister un mot-code $\mathbf{c}'$ qui vérifie $d_H(\mathbf{y},\mathbf{c}') < d_H(\mathbf{y},\mathbf{c})$. De façon plus générale, on a :

Propriétés 3.2 (Correction, détection) *Un code peut corriger α erreurs et détecter β erreurs, avec $\beta \geq \alpha$, si $d_{\min} \geq \alpha + \beta + 1$.*

Remarques sur les effacements

Souvent lorsque le symbole reçu est "ambigu" lors une décision dure (par exemple, niveau proche de 0 pour une MIA-2) ou quand le récepteur s'aperçoit d'une anormalité dans la démodulation, un ou plusieurs symboles peuvent être effacés au lieu d'être vu comme une valeur de l'alphabet. Alors tout mot comportant g effacements peut être corrigé si $d_{\min} \geq g+1$.

Tout mot contenant α erreurs et g effacements peut être corrigé si $d_{\min} \geq 2\alpha + g + 1$.

Une borne supérieure sur t

Une condition nécessaire pour qu'un code corrige t erreurs est que les ensembles des mots à la distance 1, 2, ..., t (voir figure 3.4) des $M = 2^k$ mots-code soient disjoints. Il faut donc que :

$$2^n \geq 2^k(1 + C_n^1 + C_n^2 + \cdots + C_n^t)$$

Quand l'égalité est vérifiée on dit que le code est parfait.

Exemple 3.4 (Code à répétitions $(n,1,n)$) *La distance minimale du code à répétitions est $d_{\min} = n$. La règle de décision optimale consiste à effectuer un "vote" majoritaire. Ce code corrige $\lfloor (n-1)/2 \rfloor$ erreurs.*

Exemple 3.5 (Code parité $(n,n-1,2)$) *Ce code est défini dans l'exemple 3.3. Sa distance minimale est $d_{\min} = 2$. Ce code ne corrige pas d'erreur par contre il permet d'en détecter une.*

[2] $\lfloor x \rfloor$ désigne la partie entière de x définie par $\lfloor x \rfloor \in \mathbb{N}$ et tel que $\lfloor x \rfloor \leq x < \lfloor x \rfloor + 1$.

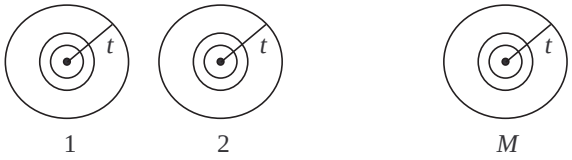

FIG. 3.4 – *Les $M = 2^k$ mots-code et leurs sphères de décision associées aux distances 1,..., t. Si on veut que le code corrige t erreurs, il faut que les sphères de rayon t soient disjointes, et par conséquent que $2^n > M(1 + n + \cdots + C_n^t)$.*

Taux d'erreurs par élément binaire (TEEB)

Considérons un code $\mathcal{C}(n, k, d_{\min})$, sur le canal CBS de probabilité d'erreur p. Ce code corrige t erreurs. Par conséquent les régions de vraisemblance maximale Λ_i contiennent au moins les C_n^j mots distants de $j \leq t$ du mot-code associé à cette région. En se reportant à la formule (3.8), on peut écrire :

$$\sum_{y\in\Lambda_i} (1-p)^{n-d_H(y,C_i)} p^{d_H(y,C_i)} \geq (1-p)^n + C_n^1(1-p)^{n-1}p + \cdots + C_n^t(1-p)^{n-t}p^t$$

On en déduit une borne inférieure de la probabilité de décision correcte :

$$\bar{P}_c \geq (1-p)^n + C_n^1(1-p)^{n-1}p + \cdots + C_n^t(1-p)^{n-t}p^t$$

et donc, pour la probabilité d'erreur, la borne supérieure :

$$\bar{P}_e \leq 1 - (1-p)^n - C_n^1(1-p)^{n-1}p = \sum_{j=t+1}^{n} C_n^j p^j (1-p)^{n-j} \tag{3.10}$$

Si on s'intéresse, à présent, au *taux d'erreur par éléments binaires* (TEEB), dont l'expression est donnée par :

$$\text{TEEB} = \frac{1}{M}\sum_{i=1}^{M}\sum_{y\in\bar{\Lambda}_i} \frac{d_H(y,c_i)}{n}(1-p)^{n-d_H(y,c_i)}p^{d_H(y,c_i)} = \frac{1}{M}\sum_{i=1}^{M}\text{TEEB}_i \tag{3.11}$$

le calcul est, là encore, en principe très compliqué. Il faut en effet se donner un code et en connaître l'ensemble des distances entre mots-code[3] On peut toutefois obtenir l'expression approchée suivante :

$$\text{TEEB} \approx \sum_{j=t+1}^{n} \frac{j+t}{n}\, C_n^j p^j (1-p)^{n-j} \tag{3.12}$$

En effet considérons un code qui corrige t erreurs et supposons que le canal introduise j erreurs, avec $j \geq t+1$ au mot-code émis c_i. Au décodage on va, dans le pire des cas, en modifier t de plus (pour peu qu'il existe un mot-code à la distance t du mot reçu !). Le taux d'erreur est alors de $(j+t)/n$ et, donc conditionnellement au mot-code émis c_i, on a :

$$\text{TEEB}_i \approx \sum_{j=t+1}^{n} \frac{j+t}{n}\, C_n^j p^j (1-p)^{n-j}$$

En portant dans l'expression (3.11) on obtient l'expression (3.12). Dans le cas où $p \ll 1$, cette expression se simplifie si on considère uniquement le terme d'ordre $t+1$. On aboutit alors à une formule pratique du taux d'erreurs par éléments binaires dont l'expression est :

$$\text{TEEB} \approx \frac{2t+1}{n}\, C_n^{t+1} p^{t+1} \tag{3.13}$$

[3]Il pourrait paraître plus raisonnable, de prime abord, de choisir le code et de construire la règle de décision de façon à minimiser directement le TEEB. La raison qui conduit à préférer minimiser, comme nous l'avons fait, la probabilité d'erreur par mot-code est que la résolution de ce problème est bien plus simple.

Gain de codage en modulation MIA-2

Dans le cas où le canal CBS est obtenu à partir d'une transmission binaire sur le canal de Nyquist, on a vu que sans codage la probabilité d'erreur est donnée par :

$$p = Q\left(\sqrt{\frac{2E_0}{N_0}}\right)$$

où $N_0/2$ désigne la d.s.p. du bruit AGB et où E_0 désigne l'énergie mise en jeu à chaque utilisation du canal. Si on dispose, pour transmettre la source d'une énergie moyenne par bit égale à E_b et que l'on utilise un code (n, k, d), l'énergie moyenne disponible à chaque utilisation du canal est donc :

$$E_0 = \frac{k}{n}\, E_b$$

Par conséquent la probabilité d'erreur *avec codage*, en effectuant une décision ferme, a pour expression :

$$\mathrm{TEEB}^{(ac)} = \frac{2t+1}{n}\ C_n^{t+1}\ Q^{t+1}\left(\sqrt{\frac{k}{n}\frac{2E_b}{N_0}}\right) \tag{3.14}$$

Cette expression est à comparer à celle de la probabilité d'erreur *sans codage* qui s'écrit :

$$\mathrm{TEEB}^{(nc)} = Q\left(\sqrt{\frac{2E_b}{N_0}}\right) \tag{3.15}$$

Pour une même probabilité d'erreur, le système sans codage nécessite en principe un rapport signal sur bruit ρ fois plus grand. La quantité $20\log_{10}(\rho)$, exprimée en dB, s'appelle le *gain de codage*. On verra, page 75, une application numérique pour un code de Hamming.

Supposons que l'on veuille obtenir un code $\mathcal{C}(70, 50)$ pour le canal CBS, c'est-à-dire un code transmettant 50 bits d'information avec 20 bits de redondance. On montre alors qu'en tirant au hasard les 2^{50} mots-code parmi le 2^{70} mots possibles, il y a de grandes chances de tomber sur un très bon code. Toutefois, la fonction de décision, donnée par (3.6) bien que simple, est difficile à mettre en œuvre avec un algorithme pratique. Il faut en effet comparer le mot reçu aux 2^{50} mots-code ; même à raison de une comparaison par nanoseconde cela fait tout de même 13 jours de calcul ! C'est pourquoi on a été conduit à rechercher des codes dont la structure facilite l'algorithme de détermination du mot-code le plus proche. Les codes en bloc et, plus généralement, les codes convolutifs ont une telle propriété.

3.4 Codes linéaires en bloc

3.4.1 Propriétés générales

Définition 3.3 *Le code (n, k) est dit linéaire en bloc si les 2^k mots de codes ont pour expression :*

$$c = \sum_{i=1}^{k} d_i g_i \tag{3.16}$$

où les coefficients $d_i \in \mathbb{F}_2$ et où les g_i sont k vecteurs-ligne, constitués de $\mathbb{F}_2^n$, supposés linéairement indépendants. Les opérations d'addition et de multiplication sont effectuées modulo 2.

Notons que les vecteurs-ligne g_i sont eux même des mots-code. Ils sont supposés linéairement indépendants car il est, en effet, raisonnable d'exiger que, si $d_1 \neq d_2$, alors $c_1 \neq c_2$. Ils forment donc une base du sous-espace de $\mathbb{F}_2^n$ engendré par G. Ce sous-espace est de dimension k. En écrivant l'expression (3.16) sous forme matricielle on a :

$$c = d\ G \quad \text{où} \quad G = \begin{bmatrix} g_1^1 & \cdots & g_1^n \\ \vdots & & \vdots \\ g_k^1 & \cdots & g_k^n \end{bmatrix}$$

et où d est l'un des $M = 2^k$ vecteurs-ligne de longueur k qui représentent les messages d'information à transmettre. La matrice G est appelée *matrice génératrice* du code. Elle est, d'après la remarque ci-dessus, de rang plein k. Et on a :

$$d_1 \neq d_2 \ \Leftrightarrow d_1 G \neq d_2 G$$

Code mis sous forme systématique

Rappelons qu'un code est principalement le choix de 2^k points dans l'espace $\mathbb{F}_2^n$ et que les performances sont essentiellement liées aux inter-distances entre ces points. Les deux remarques qui suivent conduisent à la notion de forme systématique d'un code en bloc linéaire :

- combinaisons linéaires des lignes de G : d'après la définition d'un code linéaire en bloc, toute opération de changement de base, qui s'obtient par combinaisons linéaires des lignes de G, conduit aux mêmes mots-code. Ce qui change est l'étiquetage des 2^k mots-code par les suites de k bits d'information. Par conséquent, la probabilité d'erreur par mot-code sur le CBS est inchangée.
- permutation des colonnes : soit $\mathcal{C}$ un code linéaire en bloc associé à la matrice G. La permutation des colonnes de G conduit à un code $\mathcal{C}'$ dont la matrice génératrice s'écrit $G' = G\Pi$ où Π désigne une matrice de permutation. On sait que cette opération laisse la matrice génératrice G' de rang plein. Par contre elle change le code. Les mots-code de $\mathcal{C}'$ s'obtiennent par une même permutation des mots-code de $\mathcal{C}$. Toutefois cette opération ne modifie les inter-distances des mots-code de ces deux codes et, par conséquent, la probabilité d'erreur par mot-code sur le CBS reste inchangée.

En conclusion tout changement de base (sur les lignes) et toute permutation des colonnes de la matrice génératrice d'un code laissent inchangée la probabilité d'erreur par mot-code sur le CBS. On peut alors, par un choix judicieux de la base et des permutations, amener la matrice génératrice[4] sous la forme :

$$G = \left[\begin{array}{c|c} I_k & P_{k\times(n-k)} \end{array} \right] \tag{3.17}$$

où I_k est la matrice identité de dimension k. On dit alors que le code est mis sous *forme systématique*. La matrice P de dimensions $k \times (n-k)$ est appelée la *matrice de parité* du code. Dans ce cas, les 2^k mots-code "commencent" par les k bits d'information et sont complétés par $(n-k)$ bits de redondance, sous la forme (ligne) :

$$c = d\,G = \left[\begin{array}{cccccc} d_1 & \cdots & d_k & p_1 & \cdots & p_{n-k} \end{array} \right]$$

On verra (propriétés 3.5) que, lorsque G est mise sous forme systématique, la détermination de la matrice H telle que $HG^T = 0$ est simple.

Exemple 3.6 *On considère le code linéaire* $(7, 4)$ *défini par la matrice génératrice :*

$$G = \left[\begin{array}{ccccccc} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{array} \right]$$

Vérifier que la matrice est de rang plein puis la mettre sous forme systématique.

Réponse : Il faut choisir k colonnes de G tels que la sous-matrice de dimensions $k \times k$ soit inversible, ce qui est toujours possible puisque la matrice G est supposée de rang plein. Sur notre exemple on note que la sous-matrice :

$$\tilde{G}_4 = \left[\begin{array}{cccc} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{array} \right]$$

construite sur les 4 premières colonnes est inversible puisque son déterminant est égal à 1. Il n'est donc pas utile d'effectuer de permutation.

Il faut ensuite de trouver 4 combinaisons linéaires des vecteurs-ligne de G, caractérisées par une matrice Λ de dimension 4×4, telles que :

$$\Lambda G = \left[\begin{array}{c|c} I_4 & P \end{array} \right]$$

où I_4 est la matrice identité de taille 4. On note F la matrice obtenue. On en déduit $\Lambda\tilde{G}_4 = I_4$. Un calcul sans difficulté donne :

$$\Lambda = \tilde{G}_4^{-1} = \left[\begin{array}{cccc} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{array} \right]$$

[4] Il est clair que les k vecteurs-ligne de G, définie par (3.17), sont indépendants et donc que G est de rang plein.

En conséquence la matrice sous forme systématique de ce code a pour expression :

$$F = \tilde{G}_4^{-1} G = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}$$

Distance minimale

Nous avons vu le rôle déterminant de la distance minimale dans le calcul des performances. En principe, la détermination de l'expression de la distance minimale d'un code est souvent compliquée : il faut en principe calculer C_M^2 distances puis déterminer la plus petite. Pour les codes linéaires on a le résultat suivant :

Propriétés 3.3 (Distance minimale) *La distance minimale d'un code linéaire en bloc est le poids du mot-code de poids minimale autre que* 0 *:*

$$d_{\min} = \min_{c \neq 0} \pi_H(c)$$

En effet soit $c_1 \neq c_2$, $d(c_1, c_2) = \pi_H(c_1 \oplus c_2) \neq 0$. Comme $c_1 \oplus c_2$ appartient au code (puisque le code est linéaire), la recherche du minimum de $d(c_1, c_2) \neq 0$ est équivalent à la recherche du minimum de $\pi_H(c)$ pour tout mot-code $c \neq 0$. La détermination de $d_{\min}$ ne nécessite alors que le calcul de $2^k - 1$ poids.

Codes de Hamming

Un code de Hamming est défini, pour une valeur de m donnée, par $(2^m - 1, 2^m - 1 - m, 3)$. On note que, pour un code de Hamming, on a $2^n = 2^k(1+n)$. Cela signifie que chacun des 2^k mots utiles est entouré des n mots à la distance 1. Par conséquent un code de Hamming ne corrige qu'une erreur et sa distance minimale est 3. Son taux de codage est donné par :

$$r = (2^m - 1 - m)/(2^m - 1)$$

qui est une fonction croissante de m.

Exemple 3.7 (Code de Hamming $(4, 7, 3)$**)** *le code de Hamming* $(7, 4, 3)$ *a pour matrice génératrice :*

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix} = \left[\begin{array}{c|c} I_4 & P_{4,3} \end{array} \right] \quad \text{où} \quad P_{4,3} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

Définition 3.4 (Matrice de contrôle de parité) *Soit un code linéaire* $\mathcal{C}(n, k)$ *de matrice génératrice* $\mathbf{G}$*. On appelle matrice de contrôle de parité une matrice* $\mathbf{H}$ *de dimensions* $(n - k) \times n$ *et de rang plein* $(n - k)$ *qui vérifie :*

$$\mathbf{H}\mathbf{G}^T = 0 \Longleftrightarrow \mathbf{G}\mathbf{H}^T = 0 \tag{3.18}$$

On en déduit que $\mathbf{G}$ et $\mathbf{H}$ décomposent l'espace $\mathbb{F}^n$ en deux sous-espaces complémentaires $\mathcal{C}$ et $\mathcal{C}'$ de dimensions respectives k et $(n - k)$ et tels que, pour tout $\mathbf{x} \in \mathbb{F}^n$, il existe un unique couple $(\mathbf{x}_1, \mathbf{x}_2) \in \mathcal{C} \times \mathcal{C}'$, tel que $\mathbf{x} = \mathbf{x}_1 \oplus \mathbf{x}_2$. Comme $\mathbf{G}$ et $\mathbf{H}$ sont de rang plein, on en déduit alors qu'il existe un unique couple $(\mathbf{d}_1, \mathbf{d}_2) \in \mathbb{F}^k \times \mathbb{F}^{n-k}$, tel que :

$$\mathbf{x} = \mathbf{d}_1\mathbf{G} \oplus \mathbf{d}_2\mathbf{H} \tag{3.19}$$

On a le résultat fondamental suivant :

Propriétés 3.4 (Caractérisation d'un mot-code) $\mathbf{x} \in$ *code* $\Longleftrightarrow \mathbf{x}\mathbf{H}^T = \mathbf{0}$

Condition nécessaire : si $\mathbf{x} \in$ code $\Rightarrow \mathbf{x} = \mathbf{a}\,\mathbf{G} \Rightarrow \mathbf{x}\mathbf{H}^T = \mathbf{a}\,\mathbf{G}\mathbf{H}^T = 0$ d'après la définition (3.18) de $\mathbf{H}$. Condition suffisante : supposons que $\mathbf{x}\mathbf{H}^T = 0$. D'après (3.19), $\mathbf{x}$ peut s'écrire $\mathbf{x} = \mathbf{d}_1\mathbf{G} \oplus \mathbf{d}_2\mathbf{H}$. En multipliant à droite par $\mathbf{H}^T$, il vient $\mathbf{0} = \mathbf{0} \oplus \mathbf{d}_2\mathbf{H}\mathbf{H}^T = 0$ qui implique que $\mathbf{d}_2 = \mathbf{0}$ puisque $\mathbf{H}$ est de rang plein. Et donc $\mathbf{x}$ s'écrit $\mathbf{x} = \mathbf{d}_1\mathbf{G}$ qui signifie que $\mathbf{x}$ appartient au code.

Définition 3.5 (Syndrome) *On appelle syndrome associé à* $\mathbf{y}$ *le vecteur de* $\mathbb{F}^{n-k}$ *défini par* $\mathbf{s} = \mathbf{H}\mathbf{y}^T$.

Une conséquence directe de la propriété (3.4) est que le syndrome de $\mathbf{y}$ est nul si et seulement si $\mathbf{y}$ appartient au code.

Propriétés 3.5 *Soit un code linéaire dont la matrice génératrice* $\mathbf{G}$ *est mise sous forme systématique. Alors la matrice de contrôle de parité* $\mathbf{H}$ *a pour expression :*

$$\mathbf{H} = \left[\; (\mathbf{P}^T)_{(n-k)\times k} \quad | \quad \mathbf{I}_{(n-k)} \; \right] \tag{3.20}$$

En effet, d'une part $\mathbf{H}\mathbf{G}^T = \mathbf{P}^T \oplus \mathbf{P}^T = 0$, et d'autre part $\mathbf{H}$ est de rang $(n-k)$, puisque les $(n-k)$ lignes de $\mathbf{H}$ sont constituées des vecteurs de $\mathbf{I}_{n-k}$ complétés par ceux de $\mathbf{P}$.

Une conséquence directe est que :

Propriétés 3.6 *Pour un code linéaire en bloc la distance minimale* $d_{\min}$ *est égale au plus petit nombre de colonnes dépendantes de la matrice* $\mathbf{H}$.

En effet, d'après la propriété 3.3, pour un code linéaire la distance minimale est égale au poids du mot-code de poids minimal. Or d'après la propriété 3.4, $\mathbf{H}\mathbf{x}^T = 0$. Ce qui démontre la propriété.

Il s'ensuit une borne supérieure très simple de la distance minimale.

Propriétés 3.7 (Borne supérieure simple de $d_{\min}$**)** *Pour un code linéaire en bloc :*

$$d_{\min} \leq n - k + 1 \tag{3.21}$$

En effet, considérons un code linéaire dont la distance minimale est égale à $d_{\min}$ et supposons que le rang r de $\mathbf{H}$ vérifie $r < d_{\min} - 1$. On peut alors trouver $(r+1)$ colonnes de $\mathbf{H}$ qui soient liées et donc construire un vecteur $\mathbf{u}$ non nul de poids $r+1$ tel que $\mathbf{H}\mathbf{u}^T = 0$. Mais d'après la propriété 3.4, $\mathbf{u}$ appartient au code et son poids est strictement inférieur à $d_{\min}$. Ce qui est impossible puisque le code est supposé avoir comme distance minimale $d_{\min}$. On en déduit que le rang de $\mathbf{H}$, qui est égal à $(n-k)$, vérifie $r = (n-k) \geq d_{\min} - 1$ qui est le résultat annoncé.

Propriétés 3.8 *Pour qu'un code linéaire en bloc corrige* t *erreurs, il doit avoir au moins* $2t$ *bits de redondance.*

En effet, d'après la propriété 3.1, pour corriger t erreurs, il faut que $2t \leq d_{\min} - 1$. En appliquant (3.21) il vient le résultat annoncé.

Notons que la borne supérieure donnée par (3.21) est très grande. Elle indique toutefois que si on veut augmenter $d_{\min}$, dans le but de diminuer la probabilité d'erreur, alors il faut augmenter $n - k = n(1 - k/n)$ et donc augmenter n, si on souhaite garder k/n constant.

3.4.2 Décodage par le syndrome

La propriété 3.4 permet de tester l'appartenance d'un mot quelconque de longueur n au code engendré par la matrice $\mathbf{G}$. Supposons qu'on transmette sur le CBS le mot-code $\mathbf{c}$. Alors le mot reçu s'écrit $\mathbf{y} = \mathbf{c} \oplus \mathbf{b}$ où le vecteur de "bruit" $\mathbf{b}$ est un vecteur aléatoire à n composantes $\in \mathbb{F}_2$ dont les 1 indiquent les positions des erreurs introduites par le canal. On a alors, d'après la propriété 3.4, $\mathbf{H}\mathbf{y}^T = \mathbf{H}\mathbf{b}^T$. Par conséquent on a :

Propriétés 3.9 *Soit un code linéaire* $\mathcal{C}$ *de matrice de contrôle de parité* $\mathbf{H}$. *On note* $\mathbf{y} = \mathbf{c} \oplus \mathbf{b}$ *le mot reçu sur le CBS. Alors le vecteur de bruit* $\mathbf{b}$ *vérifie :*

$$\mathbf{s} = \mathbf{H}\mathbf{b}^T$$

où $\mathbf{s} = \mathbf{H}\mathbf{y}^T$ *désigne le syndrome associé à* $\mathbf{y}$.

Ce résultat est une conséquence directe de la propriété 3.4 et du fait que $\mathbf{y} = \mathbf{c} \oplus \mathbf{b}$.

La propriété est remarquable car elle indique que le syndrome, qui se calcule à partir de l'observation et de la matrice de contrôle de parité, ne dépend que du bruit mais pas du mot-code émis. On peut donc espérer pouvoir déduire $\mathbf{b}$ de la connaissance de $\mathbf{H}$ et de l'observation $\mathbf{y}$. Une fois $\mathbf{b}$ connu, on peut déduire $\mathbf{c}$ simplement en écrivant que $\mathbf{y} = \mathbf{c} \oplus \mathbf{b}$ est équivalent à :

$$\mathbf{c} = \mathbf{y} \oplus \mathbf{b}$$

Syndrome nul

La règle du maximum de vraisemblance sur le CBS consiste à prendre le mot-code le plus proche du mot reçu. Par conséquent, si le syndrome nul, il faut prendre le mot reçu comme le mot-code le plus vraisemblable. Si le code est systématique, il suffit alors d'extraire les k premiers bits du mot reçu. Cela ne signifie pas que le mot émis soit le mot reçu. Il peut avoir eu un nombre d'erreurs tel que, partant d'un mot-code, on aboutisse à un autre mot-code. Ce que nous enseigne la théorie de la détection statistique est que la décision optimale, dans le sens du minimum de la probabilité d'erreur moyenne, est de faire confiance à la règle du minimum de distance.

Syndrome non nul

Si le syndrome $\mathbf{s}$, calculé à partir de l'observation, n'est pas nul et qu'il existe un unique vecteur $\mathbf{b}$ *de poids minimum* qui vérifie $\mathbf{s} = \mathbf{H}\mathbf{b}^T$, alors le mot-code *le plus proche du mot observé* $\mathbf{y}$ est donné par $\hat{\mathbf{c}} = \mathbf{y} \oplus \mathbf{b}$. Dans le cas où l'équation $\mathbf{s} = \mathbf{H}\mathbf{b}^T$ a plusieurs solutions de poids minimum, on peut choisir l'une quelconque pour déterminer $\hat{\mathbf{c}} = \mathbf{y} \oplus \mathbf{b}$. Dans le cas d'un code systématique, on en déduit ensuite $\hat{\mathbf{d}}$ en ne conservant que les k premiers bits.

Exemple 3.8 (Code de Hamming $(7,4,3)$**)** *Reprenons le code de Hamming* $(7,4,3)$ *dont la matrice génératrice est :*

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}$$

On en déduit la matrice de contrôle de parité :

$$\mathbf{H} = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

1. *On reçoit* $\mathbf{y} = [1\,0\,1\,0\,0\,1\,1]$. *Déterminer le message émis le plus vraisemblable.*
2. *Même question si on reçoit* $\mathbf{y} = [0\,0\,1\,1\,1\,0\,0]$.
3. *On émet le message* $\mathbf{d} = [0\,1\,1\,0]$. *Déterminer le mot-code émis. Déterminer les différents syndromes associés à zéro erreur, à une erreur en position* 3*, à deux erreurs en position* 4 *et* 5*, à deux erreurs en position* 2 *et* 4. *Déduire dans chaque cas le mot-code décidé.*

Exemple corrigé 3

1. *Si* $\mathbf{y} = [1\,0\,1\,0\,0\,1\,1]$, *alors* $\mathbf{s} = \mathbf{H}\mathbf{y}^T = [0\ 0\ 0]^T$. *On en déduit que* $\mathbf{y}$ *est un mot-code. On décide que le mot-code émis est* $[1\,0\,1\,0\,0\,1\,1]$. *Comme le code est systématique, le message est donc* $\hat{\mathbf{d}} = [1\,0\,1\,0]$.
2. *Si* $\mathbf{y} = [0\,0\,1\,1\,1\,0\,0]$, *alors* $\mathbf{s} = \mathbf{H}y^T = [0\ 0\ 1]^T$. *On note que* $\mathbf{s}$ *est la* 7*-ème colonne de* $\mathbf{H}$. *Par conséquent* $\mathbf{H}[0\,0\,0\,0\,0\,0\,1]^T = [0\ 0\ 1]^T$ *et* $\mathbf{b} = [0\,0\,0\,0\,0\,0\,1]$ *est le vecteur contenant un seul* 1 *et qui, ajouté à* $\mathbf{y}$, *donne un syndrome nul. On en déduit que le mot-code le plus proche de* $\mathbf{y}$ *est* $[0\,0\,1\,1\,1\,0\,1]$. *On décide donc* $\hat{\mathbf{d}} = [0\,0\,1\,1]$.
3. $\mathbf{d} = [0\,1\,1\,0] \rightarrow \mathbf{c} = [0\,1\,1\,0\,0\,0\,1]$ *et on a :*
   - *s'il n'y a pas d'erreur, le syndrome est nul. Le mot reçu est précisément le mot-code émis et la décision est* $\hat{\mathbf{d}} = [0\,1\,1\,0]$. *On n'a aucun bit erroné.*
   - *s'il y a une seule erreur en position* 3*, le mot reçu est* $\mathbf{y} = [0\,1\,0\,0\,0\,0\,1]$ *et le syndrome est* $[1\,1\,0]$ *qui est la* 3*-ème colonne de* $\mathbf{H}$. *Dans ce cas, le mot-code le plus proche du mot reçu est* $\hat{\mathbf{c}} = [0\,1\,1\,0\,0\,0\,1]$ *et* $\hat{\mathbf{d}} = [0\,1\,1\,0]$. *Le code corrige le bit erroné.*
   - *s'il y a deux erreurs, respectivement en position* 4 *et* 5*, le mot reçu est* $\mathbf{y} = [0\,1\,1\,1\,1\,0\,1]$ *et le syndrome est alors* $[1\,1\,1]$ *qui est aussi le syndrome d'une seule erreur en position* 2. *Dans ce cas le mot-code le plus proche du mot reçu est* $\hat{\mathbf{c}} = [0\,0\,1\,1\,1\,0\,1]$. *Par conséquent la règle de décision optimale nous oblige à dire que* $\hat{\mathbf{d}} = [0\,0\,1\,1]$. *On a donc* 2 *bits erronés.*
   - *s'il y a deux erreurs, respectivement en position* 2 *et* 4*, le mot reçu est* $\mathbf{y} = [0\,0\,1\,1\,0\,0\,1]$ *et le syndrome est alors* $[1\,0\,0]$ *qui est aussi le syndrome d'une seule erreur en position* 5. *Dans ce cas le mot-code le plus proche du mot reçu est* $\hat{\mathbf{c}} = [0\,0\,1\,1\,1\,0\,1]$. *Par conséquent la règle de décision optimale nous oblige à dire que* $\hat{\mathbf{d}} = [0\,0\,1\,1]$. *On a encore* 2 *bits erronés.*

En résumé, sur le canal binaire symétrique :

- l'absence d'erreur donne un syndrome nul,
- la présence de t erreurs donne un syndrome qui est la somme de t colonnes de $\mathbf{H}$. Par conséquent, si le syndrome observé correspond à plusieurs sommes possibles de colonnes de $\mathbf{H}$, la règle du minimum de distance de Hamming nous conduit à prendre la somme comportant le moins de termes pour modifier les éléments du mot reçu. Si plusieurs sommes comportant le moins de termes sont égales, cela signifie qu'il y a plusieurs mots-code équidistants du mot reçu et on choisit l'une quelconque de ces combinaisons pour modifier le mot reçu.
  On peut aussi, si on dispose d'une voie de retour, adopter un protocole de demande de réémission lorsque la distance de Hamming entre le mot reçu et le mot-code *trouvé* dépasse un certain seuil.

Insistons sur le fait que ces conclusions sont étroitement liées à l'utilisation du code sur le CBS. En effet elles se déduisent, en partie, de l'équation $\mathbf{s} = \mathbf{H}\mathbf{b}^T$. Le même code considéré sur un autre canal donne des capacités de détection et de correction différentes. Ainsi sur le canal à effacements, un code $\mathcal{C}(n, k, d_{\min})$ corrige $d_{\min} - 1$ effacements (voir exercice 3.2).

Signalons enfin que la méthode de décodage utilisant le calcul du syndrome est rapidement impraticable dès lors que n et k deviennent grands. C'est pourquoi des classes particulières de codes ont été étudiées, en particulier les codes BCH (voir section 3.4.3) pour lesquels des algorithmes de décodage simples ont été trouvés.

Gain de codage et limite fondamentale

En se reportant à l'expression (3.14) donnant le taux d'erreur par éléments binaires sur le canal CBS obtenu à partir d'une transmission MIA-2 sur le canal de Nyquist avec décision ferme, on a pour un code de Hamming qui corrige 1 erreur :

$$\mathrm{TEEB}^{(ac)} = \frac{3(n-1)}{2}\ Q^2\left(\sqrt{\frac{k}{n}\frac{2E_b}{N_0}}\right) \quad \text{où} \quad n = 2^m - 1 \quad \text{et} \quad k = 2^m - 1 - m \tag{3.22}$$

Nous avons représenté, figure 3.5, le taux d'erreur par éléments binaires en fonction du rapport signal sur bruit E_b/N_0, exprimé en dB, pour le système sans codage (équation 3.15) et pour le système utilisant un code de Hamming (équation 3.22) pour $m = 3$,4 et 5. Pour une probabilité d'erreur de 10^{-5}, on relève respectivement pour $m = 3$,4 et 5, comme *gains de codage* environ 0.5, 1 et 1.6 dB.

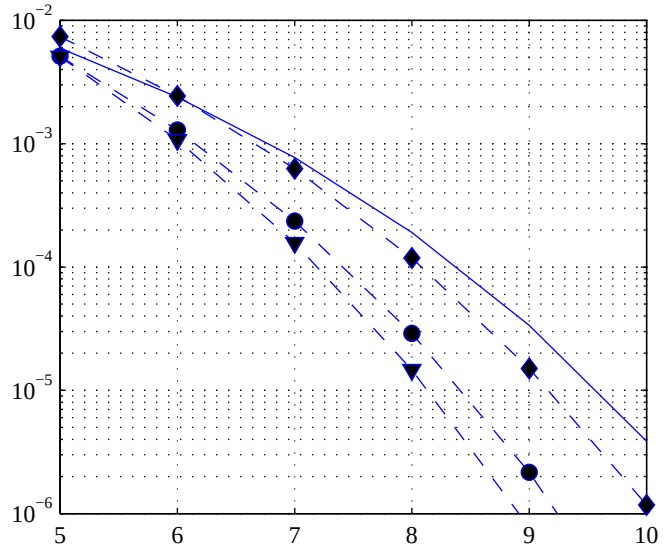

FIG. 3.5 – *Probabilité d'erreur en fonction du rapport signal sur bruit E_b/N_0 exprimé en dB. Courbe en trait plein : probabilité d'erreur sans codage. Courbes en trait pointillé : probabilité d'erreur avec codage de Hamming : $m = 3$ ('◇'),$m = 4$ ('∘'),$m = 5$ ('∇'). A TEEB $= 10^{-5}$, on relève comme gains de codage* 0.5, 1 *et* 1.6 *dB.*

La figure 3.5 indique au concepteur d'un système de transmission numérique le gain qu'il obtient en terme de rapport signal sur bruit avec un codage. Considérons, par exemple, un système de transmission qui utilise

une modulation numérique MIA-2 sur un canal idéal avec un filtrage de Nyquist. La courbe de la figure 3.5 indique qu'il faut un rapport signal sur bruit de 9.6 dB pour avoir, sans codage, un TEEB égal à 10^{-5} tandis qu'il ne faut que 8 dB pour un système utilisant un code de Hamming $(31, 26, 3)$, soit un gain d'environ 1.6 dB.

L'aspect exceptionnel de la découverte de C. Shannon est qu'il existe, dans un certain sens, une limite fondamentale au gain de codage. En effet pour une efficacité spectrale $\eta = 2$ bit/s/Hz, qui est celle de la MIA-2, la formule 2.5 de la capacité d'un canal gaussien, que nous rappelons ici $E_b/N_0 = \frac{2^\eta - 1}{\eta}$, nous indique qu'il suffit que $E_b/N_0 = 3/2$ soit 1.76 dB pour obtenir les mêmes performances (en fait "infiniment" mieux, puisqu'on peut atteindre, en principe, une probabilité d'erreur aussi faible que l'on veut !). Comparé au 9.6 dB d'un système sans codage, cela signifie que l'on peut, en théorie, avoir un gain de codage d'environ 8 dB. Des codes plus compliqués que les codes de Hamming précédents, utilisés en transmission spatiale dans les années 80, donnaient des gains de l'ordre de 6 à 7 dB. En tout il manquait encore 1 à 2 dB. La découverte, en 1993, des turbo-codes par C. Berrou et A. Glavieux a repoussé cette limite à quelques dixièmes de dB de la limite fondamentale prévue par C. Shannon.

3.4.3 Codes cycliques

Les principales propriétés qui ont orienté la recherche des codes en bloc sont la facilité d'implantation et de décodage, la capacité à corriger un grand nombre d'erreurs, la capacité à corriger des erreurs intervenant par paquets, la capacité à retrouver une perte de synchronisation, etc. Tous ces objectifs ont fait que la recherche s'est concentrée plus particulièrement sur une petite sous-classe de codes en bloc : les *codes dits cycliques.*

Définition 3.6 (Code cyclique) *On dit qu'un code linéaire en bloc est cyclique si tout décalage circulaire d'un mot-code est un autre mot-code.*

Cette définition suggère de traiter les mots-code d'un code cyclique (n, k) comme les coefficients d'un polynôme de degré inférieur ou égal à $(n-1)$, et d'associer au mot-code $\mathbf{c} = \{c_{n-1}, c_{n-2}, \cdots, c_1, c_0\}$ le polynôme-code :

$$c(D) = c_{n-1}D^{n-1} \oplus \cdots \oplus c_1 D \oplus c_0$$

où les coefficients de ce polynôme sont les composantes d'un mot-code à valeurs dans $\mathbb{F}_2 = \{0, 1\}$. Notons qu'il y a au total 2^k polynômes codes.

De la même manière, partant du mot-code $\mathbf{c}_m = \{c_{n-m-1}, c_{n-m-2}, \cdots, c_1, c_0, c_{n-1}, \cdots, c_{n-m}\}$, obtenu en décalant circulairement le mot-code $\mathbf{c}$ de m positions vers la gauche avec $0 \leq m \leq n-1$, on peut introduire le polynôme :

$$c_m(D) = c_{n-m-1}D^{n-1} \oplus \cdots \oplus c_{n-m+1}D \oplus c_{n-m}$$

Exemple 3.9 *On considère un code cyclique $\mathcal{C}(n, k)$ avec $n = 4$. On note $c(D)$ et $c_m(D)$ les polynômes-code associés à un mot-code et à son décalé de m vers la gauche. Déterminer l'expression du polynôme $p(D) = D^m c(D) \oplus c_m(D)$ pour $m = 3$. Montrer que $p(D)$ est divisible par $(D^n \oplus 1)$.*

Exemple corrigé 4 *Il vient :*

$$\begin{aligned} p(D) &= D^3(c_3D^3 \oplus c_2D^2 \oplus c_1D^1 \oplus c_0) \oplus c_0D^3 \oplus c_3D^2 \oplus c_2D \oplus c_1 \\ &= c_0(D^3 \oplus D^3) \oplus c_1(D^4 \oplus 1) \oplus c_2(D^5 \oplus D^1) \oplus c_3(D^6 \oplus D^2) \\ &= (D^4 \oplus 1)(c_1 \oplus c_2D \oplus c_3D^2) \end{aligned}$$

qui montre que $p(D)$ est divisible par $D^n \oplus 1$. On en déduit que $D^m c(D) = (D^4 \oplus 1)(c_1 \oplus c_2D \oplus c_3D^2) \oplus c_m(D)$.

L'exemple 3.9 se généralise sans difficulté et on peut énoncer :

Propriétés 3.10 *Le polynôme-code $c(D)$ associé à un mot-code et le polynôme-code $c_m(D)$ associé à son décalé circulaire de m pas vers la gauche, avec $0 \leq m \leq n-1$, vérifient :*

$$c_m(D) = D^m c(D) \bmod (D^n \oplus 1) \tag{3.23}$$

Voyons à présent comment engendrer un code cyclique. Pour cela choisissons un polynôme $g(D)$ de degré $(n-k)$ qui soit un diviseur de $(D^n \oplus 1)$ et considérons les 2^k polynômes, de degré inférieur ou égal à $(k-1)$, de la forme :

$$d(D) = d_{k-1}D^{k-1} \oplus \cdots \oplus d_1 D \oplus d_0 \tag{3.24}$$

Rappelons que 2^k est le nombre de mots-code. Posons :

$$c(D) = g(D)d(D) \tag{3.25}$$

On notera que, par construction, il y a 2^k polynômes qui sont à la fois de degré inférieur ou égal à $(n-1)$ et qui sont multiples de $g(D)$. Il y a donc une correspondance 1-pour-1 entre les polynômes-code et les 2^k polynômes multiples de $g(D)$. Vérifions que le polynôme-code $c(D)$ représente le polynôme-code d'un code cyclique. En effet un décalage circulaire de $c(D)$ donne le polynôme :

$$c_m(D) = D^m c(D) \oplus q(D)(D^n \oplus 1)$$

En utilisant $c(D) = g(D)d(D)$ il vient :

$$c_m(D) = D^m d(D)g(D) \oplus q(D)(D^n \oplus 1)$$

Comme, par hypothèse, $g(D)$ divise $(D^n \oplus 1)$, $g(D)$ divise aussi $c_m(D)$ ce qui s'écrit $c_m(D) = g(D)d_m(D)$ où $d_m(D)$ est un polynôme de degré inférieur ou égal à $(k-1)$. $d_m(D)$ est donc l'un des 2^k polynômes dont la forme est donnée par (3.24). Par conséquent, un décalage cyclique d'un mot-code, défini par (3.25), fournit un autre mot-code.

Réciproque

Commençons par une propriété sur la multiplication de deux polynômes : considérons deux polynômes $a(D)$ et $b(D)$ de degré $(n-1)$. On note a_k et b_k leurs coefficients respectifs. Alors le polynôme $c(D) = a(D)b(D) \bmod (D^n \oplus 1)$ a pour coefficients pour k allant de 0 à $n-1$

$$c_k = \sum_{i=0}^{k} a_i b_{k-i} \mod 2$$

où la convolution est *circulaire* dans le sens où les indices de b_j se calcule modulo n. On pourra montrer ce résultat à titre d'exercice. On utilisera que pour $2n-1 \geq j \geq n$ on a

$$D^j \bmod (D^n \oplus 1) = D^{n-j}$$

A présent notons H la matrice de contrôle de parité d'un code cyclique et G sa matrice génératrice. On rappelle les deux propriétés suivantes :

- les lignes de G, en particulier la dernière, appartiennent au code.
- $c \in \mathcal{C} \Leftrightarrow Hc^T = 0$. Cela signifie en particulier que toutes les lignes de H sont orthogonales à tous les mots-code.

En particulier le vecteur obtenu par décalage successif de la dernière ligne de G est orthogonale à la première ligne de H. Rappelons que, dans sa forme systématique, la dernière ligne de G s'écrit :

$$\left[\underbrace{0 \cdots\cdots\cdots 0}_{(k-1) \text{ zéros}} \quad 1 \quad b_1 \quad b_2 \quad \cdots \quad b_{n-k}\right]$$

Notons $g(D)$ le polynôme construit sur la dernière ligne de G mise sous forme systématique et notons $h(D)$ le polynôme construit sur la première ligne de H *inversée*. Il s'ensuit que

$$g(D)h(D) = 0, \bmod (D^n \oplus 1)$$

Par conséquent un code cyclique est tel que le polynôme construit sur la dernière ligne de G (mise sous forme systématique) est de degré $(n-k)$ et est un diviseur de $(D^n \oplus 1)$. Il s'ensuit que les 2^k polynômes de degré $(n-1)$ multiples de $g(D)$ sont en bijection avec les 2^k mots-code de longueur n.

Notons que les 2^k multiples de $g(D)$ de degré $(n-1)$ forment un groupe pour l'addition des polynômes modulo 2, ce qui est équivalent à dire que le code est linéaire.

En conclusion on a le résultat suivant :

Théorème 3.1 *Tout code cyclique $\mathcal{C}(n,k)$ est engendré par un diviseur $g(D)$ de degré $(n-k)$ de $(D^n \oplus 1)$ et tout diviseur de degré $(n-k)$ de $(D^n \oplus 1)$ engendre un code cyclique $\mathcal{C}(n,k)$. $g(d)$ s'appelle le polynôme générateur du code.*

Les mots-code d'un code cyclique (n,k), associé au polynôme générateur $g(D)$ de degré $(n-k)$ diviseur de $(D^n \oplus 1)$, sont donnés par :

$$c(D) = d(D)g(D)$$

où le polynôme $d(D)$ de degré $(k-1)$ peut prendre 2^k valeurs. Dans la suite on note :

$$g(D) = g_{n-k}D^{n-k} \oplus \cdots \oplus g_0, \text{ où } g_{n-k} = g_0 = 1$$

Remarque : si $g(D)$ est diviseur de $D^n \oplus 1$ et de $D^m \oplus 1$ avec $m < n$, implique que $D^m \oplus 1$ est un multiple de $g(D)$ de degré inférieure à n. C'est par conséquent un mot-code et son poids est égal à 2. Comme le code est linéaire cela entraîne que la distance minimale est inférieure ou égale à 2. C'est pourquoi on choisit en général un diviseur de $D^n \oplus 1$ qui ne soit pas diviseur de $D^m \oplus 1$ avec $m < n$.

Calcul des bits de redondance du code mis sous forme systématique

Partons du polynôme $D^{n-k}d(D)$ et effectuons la division euclidienne de ce polynôme par $g(D)$. Il vient :

$$D^{n-k}d(D) = q(D)g(D) \oplus r(D) \Rightarrow D^{n-k}d(D) \oplus r(D) = q(D)g(D)$$

avec $\partial^\circ r(D) < n-k$ et où $\partial^\circ q(D) < k$. Par conséquent $q(D)g(D)$ est un polynôme-code et donc $D^{n-k}d(D) \oplus r(D)$ est un polynôme-code. Ce polynôme comporte deux parties : les k coefficients de poids les plus grands correspondent aux coefficients de $D^{n-k}d(D)$ qui sont précisément les bits d'information et les coefficients de $r(D)$ qui représentent les bits de redondance.

Il en découle d'une part que les coefficients de $D^{n-k}d(D) \oplus r(D)$ correspondent à l'écriture d'un mot-code écrit sous *forme systématique* et d'autre part un moyen simple de calculer les bits de redondance : il suffit d'effectuer la division euclidienne de $D^{n-k}d(D)$ par le polynôme générateur et de prendre le reste.

Exemple 3.10 *On considère $g(D) = D^3 \oplus D^2 \oplus 1$ diviseur de $D^7 \oplus 1$. On vérifie en effet que $(D^7 \oplus 1) = (D^3 \oplus D^2 \oplus 1)(D^4 \oplus D^3 \oplus D^2 \oplus 1)$. On veut coder* 0101*. Déterminer les bits de redondance.*

Exemple corrigé 5 *Le code est de paramètres $n = 7$ et $k = 4$. On effectue la division euclidienne de $D^3(0 \oplus D^2 \oplus 0 \oplus D)$. Il vient*

$$D^5 \oplus D^3 = g(D)(D^2 \oplus D) + D^2 \oplus D$$

Par conséquent $r(D) = D^2 \oplus D$ et donc les bits de redondance sont 110.

Code cyclique mis sous forme systématique

On se propose de construire la matrice génératrice, mise sous forme systématique, du code cyclique associé au polynôme $g(D)$ diviseur de $D^n \oplus 1$. Rappelons tout d'abord que, dans sa forme systématique, la dernière ligne de G s'écrit :

$$\left[\underbrace{0 \cdots\cdots 0}_{(k-1) \text{ zéros}} \quad 1 \quad b_1 \quad b_2 \quad \cdots \quad b_{n-k}\right]$$

et que les coefficients sont précisément ceux de $g(D)$. La ligne précédente est pris comme la décaler sur la gauche de cette ligne. Si un '1' apparaît en colonnes k, alors on ajoute $g(D)$ ce qui fait disparaître le terme en D^{n-k}.

Exemple 3.11 *On vérifie que le polynôme $g(D) = D^4 \oplus D \oplus 1$, de degré $n-k = 4$, divise $D^{15} \oplus 1$. Déterminer à partir de $g(D)$ la matrice du code cyclique $(15, 11)$ écrite sous forme systématique.*

Exemple corrigé 6 *A chaque étape le polynôme obtenu est multiplié par D. S'il apparaît alors un terme de degré 4, on ajoute $g(D)$. En passant, par récurrence, de la ligne 11 à la ligne 1, on obtient :*

polynôme	*ligne*
$\ell_{11}(D) = g(D) = D^4 \oplus D \oplus 1$	11
$\ell_{10}(D) = D\ell_{11}(D) = D^5 \oplus D^2 \oplus D$	10
$\ell_9(D) = D\ell_{10}(D) = D^6 \oplus D^3 \oplus D^2$	9
$\ell_8(D) = D\ell_9(D) \oplus g(D) = D^7 \oplus D^3 \oplus D \oplus 1$	8
$\ell_7(D) = D\ell_8(D) \oplus g(D) = D^8 \oplus D^2 \oplus 1$	7
$\ell_6(D) = D\ell_7(D) = D^9 \oplus D^3 \oplus D$	6
$\ell_5(D) = D\ell_6(D) \oplus g(D) = D^{10} \oplus D^2 \oplus D \oplus 1$	5
$\ell_4(D) = D\ell_5(D) = D^{11} \oplus D^3 \oplus D^2 \oplus D$	4
$\ell_3(D) = D\ell_4(D) \oplus g(D) = D^{12} \oplus D^3 \oplus D^2 \oplus D \oplus 1$	3
$\ell_2(D) = D\ell_3(D) \oplus g(D) = D^{13} \oplus D^3 \oplus D^2 \oplus 1$	2
$\ell_1(D) = D\ell_2(D) \oplus g(D) = D^{14} \oplus D^3 \oplus 1$	1

qui donne la matrice génératrice :

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \end{bmatrix}$$

Transmission sur le canal binaire symétrique

Supposons que la transmission sur un canal bruyant s'accompagne d'un vecteur d'erreurs $\mathbf{b}$ de longueur n. Le mot reçu s'écrit alors $\mathbf{y} = \mathbf{c} \oplus \mathbf{b}$. Il s'ensuit que $y(D) = c(D) \oplus b(D)$ où :

$$y(D) = y_{n-1}D^{n-1} \oplus \cdots \oplus y_1 D \oplus y_0 \quad \text{et} \quad b(D) = b_{n-1}D^{n-1} \oplus \cdots \oplus b_1 D \oplus y_0$$

Comme pour les codes linéaires généraux, le syndrome joue un rôle important dans la détection et la correction des erreurs sur le canal CBS.

Définition 3.7 (Syndrome) *On appelle syndrome associé au polynôme $y(D)$ de degré $(n-1)$ le reste $s(D)$ de la division de $y(D)$ par $g(D)$.*

Propriétés 3.11 *On a :*
- *$s(D) = 0$ si et seulement si $\mathbf{y}$ appartient au code.*
- *$s(D)$ est le reste de la division de $b(D)$ par $g(D)$.*

En effet en utilisant la relation $y(D) = c(D) + b(D)$ et $c(D) = d(D)g(D)$, on obtient :

$$b(D) = (\alpha(D) \oplus d(D))g(D) \oplus s(D)$$

La deuxième propriété est importante pour le décodage. Ayant observé $\mathbf{y}$, on en déduit tout d'abord $s(D)$. Puis :
- si $s(D) = 0$, alors le mot reçu est un mot-code,
- si $s(D) \neq 0$, alors :

- si l'un des n polynômes $b(B)$ où un seul b_i est non nul donne comme reste $s(D)$, il y a un mot-code à la distance 1 du mot reçu. Il s'obtient en faisant $\mathbf{y} \oplus \mathbf{b}$.
- sinon si l'un des C_n^2 polynômes $b(D)$ où seuls deux coefficients sont non nuls donne comme reste $s(D)$, il y a un mot-code à la distance 2 du mot reçu. Il s'obtient en faisant $\mathbf{y} \oplus \mathbf{b}$.
- etc.

Propriétés 3.12 *Soit un code cyclique $\mathcal{C}(n,k)$ de polynôme générateur $g(D)$. Si $s(D)$ est le syndrome associé au bruit $b(D)$, alors le syndrome associé au bruit obtenu par décalage circulaire de $b(D)$ de ℓ positions vers la gauche est le reste de la division de $D^\ell b(D)$ par $g(D)$.*

En effet $D^\ell b(D)$ caractérise le vecteur-bruit décalé de ℓ positions sur la gauche. On a $D^\ell b(D) = D^\ell q(D)g(D) \oplus D^\ell s(D)$ qui montre que le reste de la division de $D^\ell b(D)$ par $g(D)$ est aussi le reste de la division de $D^\ell s(D)$ par $g(D)$.

Exemple 3.12 *On considère $g(D) = D^3 \oplus D^2 \oplus 1$. On reçoit $y(D) = [0101010]$. On calcule le reste de la division de $y(D) = D^5 \oplus D^3 \oplus D$ par $g(D)$. Il vient $s(D) = D^2$. Par conséquent le reste de la division de $b(D)$ par $g(D)$ est D^2. Il s'ensuit que $b(D) \neq 0$. Reste à chercher le polynôme ayant un seul coefficient dont le reste de la division par $g(D)$ égal D^2. Il vient $b = [0000100]$ et donc le mot émis le plus probable est $c(D) = [0101110]$.*

Paquet d'erreurs

Supposons que le canal introduise p erreurs successives. On dit que les erreurs arrivent par *paquet* (en anglais *burst*). Cette situation se rencontre, par exemple, lorsque le récepteur se trouve à proximité d'un système qui produit un intense et brusque rayonnement électromagnétique comme la mise en route d'un moteur.

Indiquons ici qu'une technique pratique pour traiter les paquets d'erreurs est d'utiliser un entrelaceur (en anglais *interleaver*). Un entrelaceur est un dispositif qui ré-ordonne les suites de symboles. Typiquement il s'agit d'une matrice que l'on remplit par lignes et que l'on relit par colonnes. Ainsi les symboles d'un même mot-code ne seront plus transmis consécutivement.

En général les codes bons pour le canal binaire symétrique ne sont pas nécessairement bons pour détecter/corriger des paquets d'erreurs. Cependant les codes cycliques ont quelques bonnes propriétés.

Théorème 3.2 *Tout code cyclique (n,k) peut détecter tout paquet d'erreurs dont la longueur est inférieure ou égale à $(n-k)$.*

Pour démontrer ce résultat considérons tout d'abord le syndrome $s(D)$ associé au vecteur de bruit b et le syndrome $s_\ell(D)$ associé au vecteur de bruit b décalé de ℓ positions sur la gauche. Alors on montre simplement que $s_\ell(D)$ est le reste de la division de $D^\ell s(D)$ par $g(D)$. Supposons à présent que b comporte p "1" (ce que produit p erreurs successives) à partir de la position ℓ. Il s'ensuit que le syndrome $s(D)$ est le reste de la division de $D^\ell(1 \oplus D \oplus \cdots \oplus D^{p-1})$ par $g(D)$. Notons que D^ℓ est premier avec $g(D)$. Par conséquent si $p-1 < \partial^\circ g(D) = n-k$ alors $(1 \oplus D \oplus \cdots \oplus D^{p-1})$ n'est pas divisible par $g(D)$ est le syndrome est différent de 0 : on détecte donc la présence d'erreur ! En conclusion si le paquet d'erreurs successives est de longueur inférieure ou égale à $n-k$, le code cyclique détecte ce paquet d'erreurs.

Exemple 3.13 (Code de Golay) *On considère le polynôme générateur $g(D) = D^{11} \oplus D^9 \oplus D^7 \oplus D^6 \oplus D^5 \oplus D \oplus 1$. On vérifie que $g(D)$ est un diviseur de $D^{23} \oplus 1$. Par conséquent $g(D)$ engendre un code cyclique $(23,12,7)$ appelé code de Golay. On vérifie que $2^n = 2^k(1 + C_n^1 + C_n^2 + C_n^3)$. Le code de Golay est un code parfait. Sa distance minimale est $d_{\min} = 7$. On en déduit qu'il corrige $t = 3$ erreurs. Déterminer la probabilité d'erreur sur les mots-code sur le CBS de probabilité d'erreur p.*

Exemple corrigé 7 *puisque le code de Golay est parfait la probabilité d'erreur par mot-code est donnée par l'expression :*

$$P_e = \sum_{j=4}^{n} C_{23}^j p^j (1-p)^{23-j} \approx C_{23}^4 p^4 \tag{3.26}$$

où l'approximation suppose que $p \ll 1$. Dans le cas, où le CBS considéré correspond à une transmission binaire sur le canal AGB avec décision ferme, $p = Q(\sqrt{\rho})$ avec $\rho = (k/n)2E_b/N_0$.

Figure 3.6, nous avons reporté la probabilité d'erreur, en fonction du rapport E_b/N_0, pour le code de Golay sur le canal à entrée binaire issu d'une transmission sur le canal AGB. Nous avons considéré le cas du CBS issu d'une décision ferme. Nous avons représenté l'expression exacte (3.26) *, la borne supérieure expression* (3.9)*. Nous avons aussi indiqué la borne supérieure sur la canal à décision ferme expression* (**??**)*.*

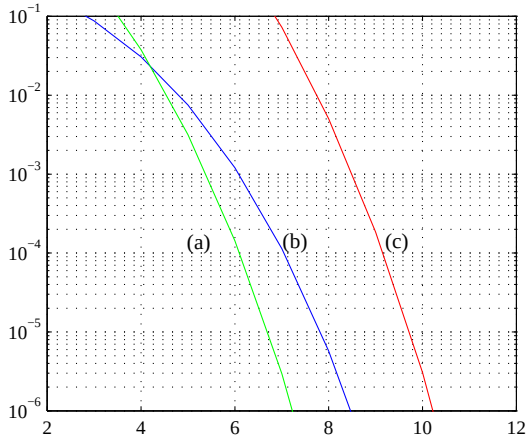

FIG. 3.6 – *Comparaison des expressions de la probabilité d'erreur par mot-code, en fonction du rapport signal sur bruit E_b/N_0, pour le code de Golay* $(23, 12)$ *sur le canal à entrée binaire, issu d'une transmission idéale sur le canal AGB. Figure (b) : expression exacte* (3.26) *sur le canal à décision ferme (CBS). Figure (c) : expression* (3.9) *de la borne supérieure pour une décision ferme. Figure (a) : expression* (**??**) *de la borne supérieure pour une décision douce.*

Codes BCH

On montre que, pour tout couple d'entiers m et t, il existe un code cyclique de paramètres :

$$n = 2^m - 1 \quad \text{et} \quad k \geq n - mt$$

Ces codes sont dits *BCH* des noms de leurs découvreurs : *Bose, Chaudhuri et Hocquenghem*. On montre que la distance minimale est donnée par :

$$d_{\min} = 2t + 1$$

En faisant $t = 1$, on trouve que les codes de Hamming sont un cas particulier de codes BCH. Comme nous l'avons montré, les polynômes générateurs sont les diviseurs de $D^n \oplus 1$. Ceux-ci ont été tabulés pour des degrés très élevés allant jusqu'à $m = 34$. Les codes BCH forment une sous-classe très importante des codes cycliques. L'une des principales raisons est que Peterson et Weldon [**?**] ont trouvé un algorithme de décodage efficace qui a rendu possible, en pratique, l'implantation de codes BCH de très grande longueur dans les systèmes de communications numériques.

3.4.4 Stratégie FEC/ARQ

Comme on l'a vu les codes permettent de corriger et/ou de détecter des erreurs introduites par le canal de transmission. La stratégie qui consiste à corriger s'appelle FEC pour *Forward Error Correction*. La stratégie qui consiste à demander une ré-émission dans le cas où on détecte une erreur est dite ARQ pour *Automatic Repeat Request*.

FEC

Les avantages sont les suivants :
- ne nécessite pas une voie de retour,
- n'introduit pas de retard variable

Les inconvénients sont les suivants :
- la correction d'erreur nécessite des algorithmes plus complexes qu'une simple détection d'erreurs.
- la correction d'erreur nécessite en règle générale un nombre de bits de redondance élevé.

Typiquement cette stratégie s'applique à des systèmes ayant des retards importants ne permettant pas l'utilisation d'une voie de retour, par exemple les liaisons spatiales.

ARQ

Les inconvénients sont les suivants :
- l'ARQ nécessite une voie de retour qui peut être surchargée s'il y a de nombreuses demandes de ré-émission.
- les demandes de ré-émission introduisent des retards de durées variables qui posent un problème dans les applications temps-réel.

3.5 Exercices

Exercice 3.1 *Un code $\mathcal{C}(100, 50, 10)$ est-il préférable à un code $\mathcal{C}(50, 25, 5)$? Existe-t-il un code linéaire $\mathcal{C}(24, 16, 10)$?*

Exercice 3.2 (Canal à effacement) *On considère le canal à effacement représenté figure 3.7*

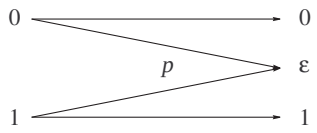

FIG. 3.7 – *Canal à effacements*

On utilise sur ce canal un code $\mathcal{C}(n, k, d_{\min})$. Combien ce code corrige-t-il d'effacements ? (remarque : ce canal ne fait pas d'erreur contrairement au canal CBS). Réponse : $t = d_{\min} - 1$.

Exercice 3.3 *Soit le code linéaire de matrice génératrice :*

$$G = \left[\begin{array}{cccc} & 1 & 0 & 1 & 1 \\ I_3 & 1 & 1 & 0 & 1 \\ & 1 & 1 & 1 & 0 \end{array}\right]$$

1. *Déterminer la matrice de contrôle de parité.*
2. *On suppose qu'après transmission il y a une erreur en position i. Déterminer le syndrome correspondant. En déduire que l'on peut corriger une erreur.*
3. *On suppose qu'il y a deux erreurs en position i et j. Déterminer le syndrome correspondant. Peut-on corriger 2 erreurs ? Peut-on détecter deux erreurs ?*

3.6 Annexes

3.6.1 Preuve de (3.9)

Partant de l'expression (3.8), la probabilité d'erreur est donnée par :

$$P_e(M, n) = \frac{1}{M} \sum_{i=1}^{M} \mathbb{P}\left\{\mathbf{Y} \in \bar{\Lambda}_i | \mathbf{X} = \mathbf{c}_i\right\}$$

où la suite $\mathcal{C} = \{\mathbf{c}_i\}$ désigne l'ensemble des mots-code et M le nombre de mots-code supposés de même probabilité $1/M$. Les régions de décision sont définies par :

$$\Lambda_i = \{\mathbf{y} \in \mathbb{F}_2^n \ t.q. \ \mathbb{P}\{\mathbf{Y} = \mathbf{y} | \mathbf{X} = \mathbf{c}_i\} > \mathbb{P}\{\mathbf{Y} = \mathbf{y} | \mathbf{X} = \mathbf{c}_j\}, \ \forall j \neq i\}$$

On note :

$$L_{j|i} = \{\mathbf{y} \in \mathbb{F}_2^n \ t.q. \ \mathbb{P}\{\mathbf{Y} = \mathbf{y}|\mathbf{X} = \mathbf{c}_j\} > \mathbb{P}\{\mathbf{Y} = \mathbf{y}|\mathbf{X} = \mathbf{c}_i\}\}$$

$L_{j|i}$ représente la région associée à $\mathbf{c}_j$ dans une détection binaire de $\mathbf{c}_i$ contre $\mathbf{c}_j$. Il est clair que :

$$\bar{\Lambda}_i = \cup_{j \neq i} L_{j|i}$$

En utilisant la borne de l'union, qui dit que $\mathbb{P}\{\cup_i A_i\} \leq \sum_i \mathbb{P}\{A_i\}$, on a :

$$P_e(M,n) \leq \frac{1}{M} \sum_{i=1}^{M} \sum_{j \neq i} \mathbb{P}\left\{\mathbf{Y} \in L_{j|i} | \mathbf{X} = \mathbf{c}_i\right\} \tag{3.27}$$

Nous allons à présent déterminer, pour le CBS, un majorant de $\mathbb{P}\left\{\mathbf{Y} \in L_{j|i} | \mathbf{X} = \mathbf{c}_i\right\}$ en fonction de la distance de Hamming $d_H(i,j)$ entre $\mathbf{c}_i$ et $\mathbf{c}_j$. Rappelons qu'il s'agit à présent d'une décision binaire entre les deux mots-code $\mathbf{c}_i$ et $\mathbf{c}_j$ qui diffèrent par $d_H(i,j)$ bits. On pose $u_{ij} = \lfloor d_H(i,j)/2 \rfloor$ qui implique que $d_H(i,j)/2 < u_{ij} + 1$. La règle de décision utilisant le minimum de distance de Hamming permet d'écrire que :

$$\begin{aligned} \mathbb{P}\left\{\mathbf{Y} \in L_{j|i} | \mathbf{X} = \mathbf{c}_i\right\} &\leq \sum_{m=u_{ij}+1}^{d_H(i,j)} C_{d_H(i,j)}^m p^m q^{d_H(i,j)-m} \\ &= \sum_{m=0}^{d_H(i,j)} C_{d_H(i,j)}^m p^m q^{d_H(i,j)-m} \mathbb{1}pm \geq u_{ij} + 1 \end{aligned} \tag{3.28}$$

où on a posé $q = 1 - p$. On suppose que $p < 1/2$ et donc $\alpha = q/p > 1$. Montrons que, pour tout m, on a :

$$\mathbb{1}pm \geq u_{ij} + 1 \leq \alpha^{m - d_H(i,j)/2} \tag{3.29}$$

En effet, si $m < u_{ij} + 1$, alors le membre gauche est nul. Et si $m \geq u_{ij} + 1$, alors $1 \leq \alpha^{m-d_H(i,j)/2}$ puisque $\alpha > 1$ et que $m - d_H(i,j)/2 > m - (u_{ij} + 1) \geq 0$ qui est une conséquence du choix de $d_H(i,j)/2 < u_{ij} + 1$. En portant (3.29) dans (3.28), il vient :

$$\begin{aligned} \mathbb{P}\left\{\mathbf{Y} \in L_{j|i} | \mathbf{X} = \mathbf{c}_i\right\} &\leq \sum_{m=0}^{d_H(i,j)} C_{d_H(i,j)}^m p^m q^{d_H(i,j)-m} (q/p)^{m - d_H(i,j)/2} \\ &= [4p(1-p)]^{d_H(i,j)/2} \end{aligned}$$

En portant cette expression dans (3.27), on obtient :

$$P_e \leq \frac{1}{M} \sum_i \sum_{j \neq i} [4p(1-p)]^{d_H(i,j)/2}$$

Comme $4p(1-p) \leq 1$ et que $\delta_{i,j} \geq d_{\min}$, un autre majorant est :

$$P_e \leq (M-1)[4p(1-p)]^{d_{\min}/2}$$

qui est l'expression (3.9).

Chapitre 4

Eléments de théorie de l'information

4.1 Capacité d'un canal de transmission

4.1.1 Notion de canal de transmission

Que doit-on entendre par a envoie un message à b ? Nous entendons par là que a agit physiquement de façon à induire chez b un état désiré. Le processus physique qui sert à l'accomplissement de cette action est toujours soumis à des perturbations incontrôlées qui rend la réception incertaine. On dira que la communication est réussie si a et b sont d'accord ce qui a été envoyé. Le nombre de messages que l'on peut envoyer de façon réussie lors de n utilisations du canal est une fonction exponentielle de n. L'exposant s'appelle la capacité du canal.

Plus précisément la source est un mécanisme qui émet un message parmi M messages possibles. A chaque message de source est associée une suite de n actions sur les entrées du canal. Le nombre d'entrée peut être fini, dénombrable ou non dénombrable. Cette suite produit n réalisations parmi les états de sortie du canal. Le nombre d'états de sortie peut être fini, dénombrable ou non dénombrable. L'objectif du transmetteur est de retrouver les suites d'entrées émises et donc, par conséquent, les symboles de source envoyés. Malheureusement les suites observées côté réception sont, à cause des perturbations sur le canal, aléatoires. Cela provoque une certaine confusion sur les symboles émis par la source. Le théorème de codage de canal dit que, malgré la confusion apparente, on peut retrouver le message d'entrée avec une probabilité d'erreur aussi faible que l'on veut, à condition que $\log_2(M)/n$ soit inférieur à une quantité appelée la capacité du canal.

Le cas le plus simple est un canal comportant deux entrées et deux sorties : l'action sur l'une des deux entrées provoque chez le destinataire la réalisation de l'une des deux sorties. Si la valeur de cette sortie Y ne dépend que de la dernière action sur l'entrée X, on dira que le canal est sans mémoire. Il peut alors être représenté par le schéma de la figure 4.1 qui est caractérisé par les 4 probabilités conditionnelles $\mathbb{P}\{Y_k = i | X_k = j\}$, où $j \in \mathcal{X} = \{0, 1\}$ et $i \in \mathcal{Y} = \{0, 1\}$, avec $\mathbb{P}\{Y_k = 0 | X_k = j\} + \mathbb{P}\{Y_k = 1 | X_k = j\} = 1$.

FIG. 4.1 – *Canal binaire sans mémoire. Les probabilités conditionnelles* $\mathbb{P}\{Y_k = 0 | X_k = 0\} = p$, $\mathbb{P}\{Y_k = 1 | X_k = 0\} = 1 - p$, $\mathbb{P}\{Y_k = 0 | X_k = 1\} = 1 - q$, $\mathbb{P}\{Y_k = 1 | X_k = 1\} = q$ *ne dépendent pas des entrées passées.*

Il en ressort qu'un canal est caractérisé par un alphabet d'entrée, un alphabet de sortie et une loi de probabilité de la sortie conditionnellement à l'entrée.

4.1.2 Exemples

Dans ce paragraphe nous allons voir que, partant de l'exemple physique d'un système de communication, il est possible de mettre en évidence plusieurs types de canaux de transmission suivant les endroits où l'on place, le long de la chaîne, l'entrée et la sortie du canal. Ces canaux diffèrent par leur alphabet d'entrée, leur alphabet de sortie et par leur loi de probabilité conditionnelle.

Considérons une chaîne de transmission binaire utilisant une modulation d'impulsions à deux niveaux (en abrégé MIA-2), sur un canal de bande B, soumis à un bruit additif, gaussien, blanc. On suppose que l'interférence entre symboles est nulle aux instants d'échantillonnage (l'ensemble des filtres d'émission et de réception vérifie le critère de Nyquist) et que la probabilité d'erreur est minimale (le filtre de réception est le filtre adapté).

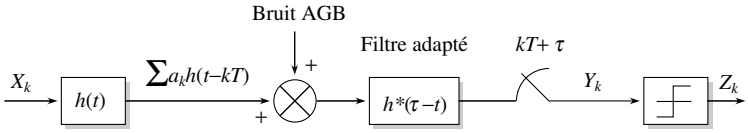

FIG. 4.2 – *chaîne de transmission binaire sur un canal de bande B, soumis à un bruit additif, gaussien, blanc. X_k est une suite de variables aléatoires i.i.d. à valeurs dans $\{0,1\}$ et $a_k = 2X_k - 1$. $h(t)$ est, par exemple, une racine carrée d'un filtre en cosinus-surélevé.*

Canal binaire symétrique

Considérons tout d'abord le système de transmission de bout en bout, dont l'entrée X_k est l'entrée du codeur, qui prend ses valeurs dans l'alphabet d'entrée $\{0,1\}$ de taille 2 et dont la sortie Z_k est la sortie du comparateur à seuil, qui prend ses valeurs dans l'alphabet de sortie $\{0,1\}$ de taille 2. On a vu dans le chapitre communications numériques que, dans les conditions d'une réception idéale sur le canal de Nyquist, on a les probabilités suivantes :

$$p = \mathbb{P}\{Z_k = 1 | X_k = 0\} = \mathbb{P}\{Z_k = 0 | X_k = 1\} = Q(\sqrt{2E_b/N_0})$$

où E_b désigne l'énergie par bit et $N_0/2$ la densité spectrale de puissance du bruit blanc.

On peut alors représenter l'ensemble par le schéma de la figure 4.3 que l'on désigne sous le nom ce *canal binaire symétrique* (en abrégé CBS).

FIG. 4.3 – *Canal binaire symétrique. $q = 1 - p$.*

Notons que d'une part p est indépendant de l'instant d'utilisation du canal et que d'autre part, lors d'utilisations successives du canal, les erreurs sont indépendantes : on peut dire que le canal est alors sans mémoire. Tout se passe comme si la sortie était soumise à un bruit additif et s'écrivait :

$$Z_k = X_k \oplus B_k$$

où $\oplus$ désigne l'opération OU EXCLUSIF et où B_k est un bruit binaire, c'est-à-dire une suite de variables aléatoires, à valeurs dans $\{0,1\}$, indépendantes et identiquement distribuées telles que $\mathbb{P}\{B_k = 1\} = p$.

Canal binaire à décision douce

Considérons à présent que, sur la chaîne de transmission précédente, nous plaçons la sortie Y_k après l'échantillonneur qui suit le filtre de réception (qui est le filtre adapté dans le cas d'une réception optimale). Depuis l'entrée X_k jusqu'à la sortie Y_k de l'échantillonneur, on dispose d'un canal dont l'alphabet d'entrée est l'ensemble $\{0,1\}$ et dont l'alphabet de sortie est $\mathbb{R}$. Nous avons vu chapitre 2 que l'observation s'écrit $Y_k = a_k + W_k$ où $a_k = 2X_k - 1$ et où W_k est une suite de variables gaussiennes, centrées, indépendantes entre elles, de même variance $\sigma^2 = \mathbb{E}\left\{W_k^2\right\} = N_0/2E_b$. Par suite Y_k suit une loi de probabilité gaussienne dont la densité de probabilité conditionnellement aux deux symboles d'entrée a pour expression :

$$p_{Y_k|X_k=0}(y) = \frac{1}{\sigma_b\sqrt{2\pi}}\exp(-(y+1)^2/2\sigma^2)$$

et

$$p_{Y_k|X_k=1}(y) = \frac{1}{\sigma\sqrt{2\pi}}\exp(-(y-1)^2/2\sigma^2)$$

où $\sigma^2 = N_0/2E_b$.

Le canal obtenu est représenté figure 4.4. Il est dit à décision douce[1]. Nous verrons plus loin la raison du choix de ce terme.

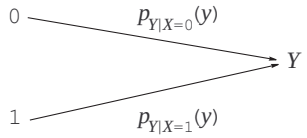

FIG. 4.4 – *Canal binaire à décision douce*

Canal à effacement

Partons toujours d'une transmission binaire sur la canal de Nyquist en présence d'un bruit additif, gaussien, blanc. Mais considérons à présent que la sortie du canal est la sortie d'un organe de décision, mis après l'échantillonneur, et qui fonctionne de la façon suivante :

- si $Y_k > e > 0$ il affiche la valeur $V_k = 1$,
- si $-e \leq Y_k \leq e$ il affiche la valeur $V_k = \epsilon$,
- enfin si $Y_k < -e$ il affiche la valeur $V_k = 0$.

On a alors un canal dont l'alphabet d'entrée est binaire $\{0,1\}$ et dont l'alphabet de sortie est ternaire $\{0,\epsilon,1\}$. D'après l'expression des lois de probabilités de Y_k conditionnellement à X_k, on obtient comme loi conditionnelle de V_k par rapport à X_k les expressions suivantes :

$$\begin{aligned}
&\mathbb{P}\{V_k = 0|X_k = 0\} = \mathbb{P}\{V_k = 1|X_k = 1\} = P_c\\
&\mathbb{P}\{V_k = 0|X_k = 1\} = \mathbb{P}\{V_k = 1|X_k = 0\} = q\\
&\mathbb{P}\{V_k = \epsilon|X_k = 1\} = \mathbb{P}\{V_k = \epsilon|X_k = 0\} = p
\end{aligned}$$

où

$$\begin{aligned}
&q = \mathbb{P}\{V_k \geq e|0\} = Q(\sqrt{2(E_b+e)/N_0})\\
&P_c = \mathbb{P}\{V_k \geq e|1\} = 1 - Q(\sqrt{2(E_b-e)/N_0})\\
&p = 1 - P_c - q
\end{aligned}$$

En pratique $q \approx 0$ et le canal peut être représenté par le schéma de la figure 4.5 avec $p \approx Q(\sqrt{2(E_b-e)/N_0})$. Le canal obtenu porte le nom de *canal à effacements*.

[1]En anglais on emploie le terme *soft decision* par opposition au terme *hard decision* utilisé pour désigner la décision ferme

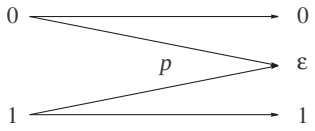

FIG. 4.5 – *Canal à effacements*

Canal gaussien

Reprenons encore le cas d'une transmission sur le canal de Nyquist en présence d'un bruit additif, gaussien, blanc. Mais considérons à présent que l'entrée du canal est la sortie du filtre d'émission *échantillonnée*. L'échantillonnage en sortie du filtre d'émission est justifié puisque le signal est à bande limitée et peut donc être remplacé par la suite de ses échantillons, si la cadence d'échantillonnage est supérieure à deux fois la bande. On note U_k la suite obtenue. La sortie du canal est la sortie Y_k de l'échantillonneur de réception.

On a alors un canal dont l'entrée U_k et à la sortie Y_k sont à valeurs dans $\mathbb{R}$. Tout se passe comme si on avait :

$$Y_k = U_k + B_k$$

où B_k est un bruit additif gaussien, blanc, c'est-à-dire une suite de variables aléatoires gaussiennes, indépendantes, centrées, de même variance σ_B^2. En supposant que U_k et B_k sont indépendantes, l'observation Y_k, conditionnellement à U_k, suit une loi de probabilité gaussienne dont la densité est :

$$p_{Y_k|U_k}(y) = \frac{1}{\sigma_B\sqrt{2\pi}} \exp\left(-\frac{(y-u_k)^2}{2\sigma_B^2}\right)$$

4.1.3 Définitions

Définition 4.1 (Canal discret) *Un canal discret est constitué d'un alphabet d'entrée $\mathcal{X}$ de taille finie, d'un alphabet de sortie $\mathcal{Y}$ de taille finie et d'une loi de probabilité de transition notée $p(y|x)$. On le note $\{\mathcal{X}, p(y|x), \mathcal{Y}\}$.*

La loi $p(y|x)$ est caractérisée par les probabilités :

$$\mathbb{P}\{Y = y|X = x\}$$

où $x \in \mathcal{X}$ et $y \in \mathcal{Y}$ et qui vérifie pour tout x :

$$\sum_{y\in\mathcal{Y}} \mathbb{P}\{Y = y|X = x\} = 1$$

Définition 4.2 (Extension d'ordre n sans mémoire et sans voie de retour) *L'extension d'ordre n sans mémoire et sans voie de retour d'un canal discret $\{\mathcal{X}, p(y|x), \mathcal{Y}\}$ est le canal $\{\mathcal{X}^n, p(y^n|x^n), \mathcal{Y}^n\}$ où :*

$$\mathbb{P}\{Y_n = y_n, \ldots, Y_1 = y_1|X_n = x_n, \ldots, X_1 = x_1\} = \prod_{k=1}^{n} \mathbb{P}\{Y_k = y_k|X_k = x_k\} \tag{4.1}$$

Remarquons que l'expression (4.1) peut se déduire de :

- absence de mémoire

$$\mathbb{P}\{Y_k = y_k|X_k = x_k, \ldots, X_1 = x_1, Y_{k-1} = y_{k-1}, \ldots, Y_1 = y_1\} = \mathbb{P}\{Y_k = y_k|X_k = x_k\}$$

- absence de voie de retour

$$\mathbb{P}\{X_k = x_k|X_{k-1} = x_{k-1}, \ldots, X_1 = x_1, Y_{k-1} = y_{k-1}, \ldots, Y_1 = y_1\} = \mathbb{P}\{X_k = x_k|X_{k-1} = x_{k-1}, \ldots, X_1 = x_1\}$$

Définition 4.3 ((n, M)-code) *Un (n, M)-code pour le canal $\{\mathcal{X}^n, p(y^n|x^n), \mathcal{Y}^n\}$ est défini par la donnée :*
- *d'un ensemble $\mathcal{M} = \{1, \ldots, M\}$ de M indices (messages),*
- *d'une application $c : \mathcal{M} \mapsto \mathcal{X}^n$,*
- *et d'une application $g : \mathcal{Y}^n \mapsto \mathcal{M}$, appelée règle de décision.*

$c(i) = (c_1(i), \cdots, c_n(i))$ est appelé un mot-code.

Définition 4.4 *On appelle probabilité d'erreur maximale d'un (n, M)-code sur le canal $\{\mathcal{X}^n, p(y^n|x^n), \mathcal{Y}^n\}$:*

$$\begin{aligned} P_e(M, n) &= \max_{i \in \mathcal{M}} \mathbb{P}\{g(Y_1, \ldots, Y_n) \neq i | c(i)\} \\ &= \max_{i \in \mathcal{M}} \sum_{y \in \mathcal{Y}^n} \mathbb{P}\{(Y_1, \ldots, Y_n) = y | c(i)\} \mathbb{1}(g(y) \neq i) \end{aligned}$$

Définition 4.5 (Capacité) *Soit un (n, M)-code sur le canal $\{\mathcal{X}^n, p(y^n|x^n), \mathcal{Y}^n\}$. On note $r = \log_2(M)/n$ le taux de codage (qui se mesure en bits par utilisation du canal). On dit que r est acceptable si $P_{e,\max}(M, n)$ tend vers 0 quand n tend vers l'infini. La plus grande valeur de r est appelée la capacité.*

Théorème 4.1 (Codage de canal) *Soit $\{\mathcal{X}, p(y|x), \mathcal{Y}\}$ un canal sans mémoire. Sa capacité est donnée par :*

$$C = \max_{\mathcal{P}_\mathcal{X}} I(X, Y) \tag{4.2}$$

où $\mathcal{P}_\mathcal{X}$ désigne l'ensemble de toutes les lois de probabilité sur $\mathcal{X}$ et où

$$I(X, Y) = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} \mathbb{P}\{X = x, Y = y\} \log_2 \frac{\mathbb{P}\{X = x, Y = y\}}{\mathbb{P}\{X = x\} \mathbb{P}\{Y = y\}} \tag{4.3}$$

$I(X, Y)$ s'appelle l'information mutuelle entre X et Y.

Théorème 4.2 (Réciproque du théorème de capacité) *Soit un (n, M)-code défini sur le canal $\{\mathcal{X}^n, p(y^n|x^n), \mathcal{Y}^n\}$ (voir définition 4.3) avec $M = 2^{nr}$, soit $r = \log_2(M)/n$. On note $\bar{P}_e(M, n)$ la probabilité d'erreur moyenne définie par :*

$$\begin{aligned} \bar{P}_e(M, n) &= \sum_{i=1}^{M} \frac{1}{M} \mathbb{P}\{g(Y_1, \ldots, Y_n) \neq i | X_n = c_n^n(i), \ldots, X_1 = c_1^n(i)\} \\ &= \mathbb{P}\{i \neq g(Y_1, \ldots, Y_n)\} \end{aligned}$$

Alors on a :

$$\bar{P}_e(M, n) \geq 1 - \frac{C}{r} - \frac{1}{nr}$$

où C désigne la capacité du canal. Si $r > C$, $\bar{P}_e(M, n)$ est bornée inférieurement par une quantité strictement positive.

Il existe une forme plus forte de ces résultats : on montre que, si $r < C$, la probabilité d'erreur tend exponentiellement vers 0 quand n tend vers l'infini et si $r > C$ la probabilité d'erreur tend exponentiellement vers 1. La capacité représente, pour le débit, une valeur critique qui sépare le cas des communications sûres de celles qui ne le sont pas.

Lien avec le débit binaire en bits/s : si D_b désigne le débit binaire, en nombre de bits par unité de temps, de la source à l'entrée du canal et W_c le nombre moyen d'utilisations du canal par unité de temps, il s'en suit que l'on a :

$$r = \frac{\log_2(M)}{n} = \frac{D_b}{W_c}$$

et donc $D_b < CW_c$, formule qui établit la limite, en bits/s, du débit pour avoir une communication sûre. En pratique C est souvent déterminé par le rapport signal sur bruit et W_c par la bande passante en fréquence du canal.

Propriétés 4.1 *La capacité vérifie :*
- *$C \geq 0$,*

– $C \leq \min\{\log_2(\text{card}\mathcal{X}), \log_2(\text{card}\mathcal{Y})\}$

Nous verrons paragraphe 4.2.2 comment montrer ces propriétés. Rappelons simplement ici les résultats suivants qui sont souvent utiles dans le calcul de la capacité :
– $\mathbb{P}\{X = x, Y = y\} = \mathbb{P}\{Y = y|X = x\}\,\mathbb{P}\{X = x\}$
– $\mathbb{P}\{X = x\} = \sum_{y\in\mathcal{Y}} \mathbb{P}\{X = x, Y = y\}$
– $\mathbb{P}\{Y = y\} = \sum_{x\in\mathcal{X}} \mathbb{P}\{X = x, Y = y\}$.

Par conséquent l'expression (4.3) peut encore s'écrire :

$$I(X,Y) = \sum_{x\in\mathcal{X}} \sum_{y\in\mathcal{Y}} \mathbb{P}\{X = x, Y = y\} \log_2 \frac{\mathbb{P}\{Y = y|X = x\}}{\mathbb{P}\{Y = y\}}$$

La formule de la capacité donnée par l'expression (4.2) est encore vraie :
– dans le cas des alphabets "continus" c'est-à-dire lorsque les lois de probabilité possèdent une densité (par rapport à la mesure de Lebesgue). Il suffit, dans ce cas, de remplacer, dans $I(X,Y)$, le signe *somme* par le signe *intégrale* et les *probabilités* par des *densités* :

$$\begin{aligned} I(X,Y) &= \int_{\mathcal{X}}\int_{\mathcal{Y}} p_{XY}(x,y) \log_2 \frac{p_{XY}(x,y)}{p_X(x)p_Y(y)}\, dxdy \\ &= \int_{\mathcal{X}}\int_{\mathcal{Y}} p_{XY}(x,y) \log_2 \frac{p_{Y|X}(x,y)}{p_Y(y)}\, dxdy \end{aligned}$$

– dans le cas où l'on impose des contraintes sur l'entrée c'est-à-dire des contraintes sur l'ensemble $\mathcal{P}_\mathcal{X}$ des lois de probabilités sur $\mathcal{X}$. Ainsi dans l'exemple 4.6 on calcule la capacité du canal gaussien sous la contrainte que l'entrée est de moyenne nulle et de variance égale à σ_X^2.

4.1.4 Calculs de capacité

Canal CBS

Partant de la loi de transition de paramètre $p = \mathbb{P}\{Z = 0|X = 1\} = \mathbb{P}\{Z = 1|X = 0\}$ et d'une loi de probabilité quelconque sur l'entrée, loi qui est caractérisée par $\mathbb{P}\{X = 0\} = \alpha$ et $\mathbb{P}\{X = 1\} = 1 - \alpha$ avec $\alpha \in (0,1)$, on a :

$$\begin{aligned} I(X,Y) = \alpha(1-p)\log_2\left(\frac{1-p}{\alpha(1-p)+(1-\alpha)p}\right) \\ +\alpha p \log_2\left(\frac{p}{\alpha p+(1-\alpha)(1-p)}\right) \\ +(1-\alpha)p\log_2\left(\frac{p}{\alpha(1-p)+(1-\alpha)p}\right) \\ +(1-\alpha)(1-p)\log_2\left(\frac{1-p}{\alpha p+(1-\alpha)(1-p)}\right) \end{aligned}$$

FIG. 4.6 – *Canal binaire symétrique.* $q = 1 - p$.

Par raison de symétrie on trouve que $I(X,Y)$ est maximum pour $\alpha = 1/2$ et on en déduit que :

$$C(p) = 1 + p\log_2(p) + (1-p)\log_2(1-p) \tag{4.4}$$

Nous avons représenté, figure 4.7, C en fonction de p. On voit que, si $p = 1/2$, la capacité est nulle. En effet si la probabilité de recevoir 0 ou 1, conditionnellement à l'émission d'un 0 ou d'un 1, est égale à $1/2$, on

comprend bien qu'un tel canal ne puisse rien transmettre de façon sûre. D'un autre côté un canal tel que $p = 1$, qui systématiquement conduit à décider 0 à la place de 1 et inversement, est très fiable et il est normal que sa capacité soit maximale (pour s'en convaincre il suffit d'inverser les décisions).

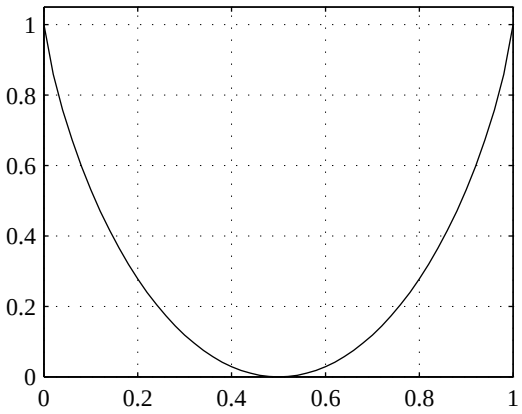

FIG. 4.7 – *Capacité du canal binaire symétrique en fonction de la probabilité de transition p.*

Explication intuitive : on peut, en s'appuyant sur le canal CBS et en utilisant un argument lié à la loi des grands nombres, fournir une explication intuitive du théorème de codage de canal : supposons que l'on envoie, sur le canal CBS de probabilité d'erreur p, le mot-code c_0 d'un code (n, M). La loi des grands nombres nous dit que, quand n est grand, le mot reçu "comportera" asymptotiquement np symboles erronés et $n(1-p)$ symboles justes : par conséquent "tous" les mots reçus tombent donc sur la "sphère" de centre c_0 et de rayon np (distance de Hamming). On dit qu'il y a durcissement de la sphère (en anglais sphere hardening). Comptons combien chaque sphère contient de mots reçus. Tous les mots reçus, d'après l'hypothèse d'indépendance (le canal est supposé sans mémoire) sont asymptotiquement équiprobables de probabilité $\pi_t = p^{np}(1-p)^{n(1-p)}$. Leur nombre est donc égal à :

$$N_t = \frac{1}{\pi_t} = p^{-np}(1-p)^{-n(1-p)} = 2^{-n(p\log_2(p)+(1-p)\log_2(1-p))} = 2^{n(1-C(p))}$$

D'un autre côté, on voit que, si les M sphères associées aux M mots-code sont disjointes, la règle de décision optimale donnera une probabilité d'erreur nulle. Par conséquent, pour que la probabilité d'erreur soit asymptotiquement nulle, il faut que :

$$MN_t < 2^n$$

En posant $r = \log_2(M)/n$ (soit $M = 2^{nr}$), il vient $2^{nr} \times 2^{n(1-C(p))} < 2^n$ soit :

$$r < C(p)$$

La recherche des bons codes : le code à répétitions est de la forme $(n, 2^{rn})$ avec un taux de codage $r = 1/n$. On a vu que, pour que la probabilité d'erreur tende vers 0, il fallait que n tende vers l'infini et donc que r tende vers 0. L'aspect fabuleux du théorème de Shannon est qu'il assure l'existence d'un code $(n, 2^{rn})$ sur le canal CBS dont la probabilité d'erreur tend vers 0, quand n tend vers l'infini, à condition que r reste strictement inférieur à la quantité $C(p) = 1 + p\log_2(p) + (1-p)\log_2(1-p)$. r n'a donc pas besoin de tendre vers 0 quand n tend vers l'infini. Pour la petite histoire, il a fallu attendre 1972 avant de trouver un code de longueur n qui corrige λn erreurs, avec $\lambda > 0$, et dont le taux de codage ne tende pas vers 0 quand n tend vers l'infini.

Canal à effacement

La capacité en fonction de p a pour expression :

$$C_\epsilon(p) = 1 - p \tag{4.5}$$

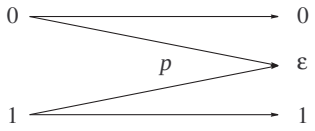

FIG. 4.8 – *Canal à effacements*

En effet notons $\alpha = \mathbb{P}\{X = 0\}$. On en déduit la loi conjointe de (X, Y) pour X appartenant à $\{0, 1\}$ et Y appartenant à $\{0, \epsilon, 1\}$:

$X \setminus Y$	0	ϵ	1
0	$\alpha(1-p)$	αp	0
1	0	$(1-\alpha)p$	$(1-\alpha)(1-p)$

Par conséquent l'information mutuelle a pour expression :

$$\begin{aligned} I(X,Y) &= -\alpha(1-p)\log_2(\alpha) - (1-\alpha)(1-p)\log_2(1-\alpha) \\ &= -(1-p)(\alpha\log_2(\alpha) + (1-\alpha)\log_2(1-\alpha)) \end{aligned}$$

Le maximum est obtenu pour $\alpha = 1/2$ et vaut $C = 1 - p$.

Imaginons que l'on dispose d'une voie de retour *sans bruit* et que l'on adopte la procédure suivante :
- si $Y_i = 0$ on décide $\hat{X} = 0$ et on arrête,
- si $Y_i = 1$ on décide $\hat{X} = 1$ et on arrête,
- si $Y_i = \epsilon$ on demande, au moyen de la voie de retour, la $(i+1)$-ème ré-émission du symbole.

On note N l'instant d'arrêt aléatoire de la procédure c'est-à-dire la plus petite valeur de k telle que $Y_k \neq \epsilon$. Plus mathématiquement $N = \inf\{k \in \mathbb{N} : Y_k \neq \epsilon\}$. Sous l'hypothèse que le canal est sans mémoire, on a pour n fixé :

$$\begin{aligned} \mathbb{P}\{N > n | X = 1\} = \mathbb{P}\{N > n | X = 0\} = \mathbb{P}\{Y_1 = \epsilon, \cdots, Y_n = \epsilon\} = p^n \\ \iff \mathbb{P}\{N \leq n | X = 1\} = \mathbb{P}\{N \leq n | X = 0\} = 1 - p^n \end{aligned}$$

On suppose que $\mathbb{P}\{X = 0\} = \mathbb{P}\{X = 1\} = 1/2$. Par conséquent la probabilité de décider sans erreur après n émissions est égale à $P_c = 1 - p^n$. Quand on fait tendre n vers l'infini, P_c tend vers 1. Par conséquent le codage avec voie de retour proposé a une probabilité d'erreur aussi petite que l'on veut. Examinons à présent le débit. Pour cela calculons le nombre moyen d'émissions. Il vient conditionnellement à l'émission de $X \in \{0, 1\}$:

$$\bar{n}_i = 1 \times (1-p) + 2 \times p(1-p) + \cdots + k \times p^{k-1}(1-p) + \cdots = \frac{1}{1-p}$$

Il faut donc en moyenne $1/(1-p)$ utilisations du canal pour transmettre, asymptotiquement sans erreur, un bit ou ce qui est équivalent : "on transmet $(1-p)$ bit par utilisation du canal". Par conséquent la voie de retour a permis d'atteindre un débit asymptotiquement sans erreur et égal à la capacité. On pourrait rétorquer que la capacité propre de la voie de retour sans bruit, doit être ajouter à celle du canal à effacement. Il n'en est rien car on démontre le résultat surprenant suivant :

Théorème 4.3 (Canal avec voie de retour) *On considère un canal discret sans mémoire $(\mathcal{X}, p(y|x), \mathcal{Y})$ disposant d'une voie de retour sans bruit. On peut alors considérer que le code $\mathcal{C}(M, n)$, défini comme une application de $\mathcal{M}$ dans $\mathcal{X}^n$, soit tel que la composante $c_i(k)$ soit à la fois une fonction de $i \in \mathcal{M}$ et des valeurs $Y_1, \cdots, Y_{k-1}$ obtenues en sortie du canal. Alors la capacité du canal est égale à celle du canal sans voie de retour :*

$$C = C_{\textit{sans voie de retour}}$$

Comme le montre l'exemple ci-dessus, la voie de retour facilite la transmission sur le canal mais n'augmente pas sa capacité.

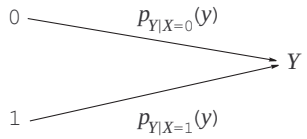

FIG. 4.9 – *Canal binaire à décision douce*

Canal binaire à décision douce

L'entrée X prend ses valeurs dans l'alphabet binaire et la sortie Y dans $\mathbb{R}$. On note $p_0(y)$ et $p_1(y)$ les densités des lois de transition conditionnellement à $X = 0$ et $X = 1$.

Un calcul sans difficulté donne pour sa capacité l'expression :

$$\begin{aligned} C_S &= 1 + \frac{1}{2}\int_{\mathbb{R}} p_0(y)\log_2\left(\frac{p_0(y)}{p_0(y)+p_1(y)}\right)dy \\ &\quad + \frac{1}{2}\int_{\mathbb{R}} p_1(y)\log_2\left(\frac{p_1(y)}{p_0(y)+p_1(y)}\right)dy \end{aligned} \tag{4.6}$$

Canal AGB

L'entrée X et la sortie $Y = X + B$ sont à valeurs dans $\mathbb{R}$. On suppose que B est une variable aléatoire gaussienne, centrée, de variance σ_B^2 et indépendante de X. On impose que X soit centrée et de variance σ_X^2. Le calcul de la capacité est un peu long mais sans difficulté (voir exemple 4.6). On vérifie tout d'abord que Y est centrée et de variance $\sigma_X^2 + \sigma_B^2$. On montre ensuite que :

$$I(X,Y) = -\int_{\mathbb{R}} p_Y(y)\log_2(p_Y(y))dy + \int_{\mathbb{R}} p_B(b)\log_2(p_B(b))db$$

On note que la seconde intégrale ne dépend pas du choix de $p_X(x)$. On montre ensuite (voir exemple 4.5) que la première intégrale est maximale si Y suit une loi gaussienne (ce qui est possible. Il suffit que (X, B) soient conjointement gaussiennes). On en déduit que :

$$C = \frac{1}{2}\log_2\left(1 + \frac{\sigma_X^2}{\sigma_B^2}\right) \tag{4.7}$$

4.1.5 Canal CBS/Canal binaire à décision douce/Canal AGB

Reprenons le système de transmission MIA-2 représenté figure 4.10. On peut alors appliquer les résultats précédents.

Canal CBS (décision ferme) l'alphabet d'entrée est $\mathcal{X} = \{0, 1\}$, où l'alphabet de sortie $\mathcal{Z} = \{0, 1\}$ et où les probabilités de transition sont données par :

$$\mathbb{P}\{Z = 0|X = 1\} = \mathbb{P}\{Z = 1|X = 0\} = Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$$

D'après (4.4) sa capacité est donnée par $C_H = 1 + p\log_2(p) + (1-p)\log_2(1-p)$.

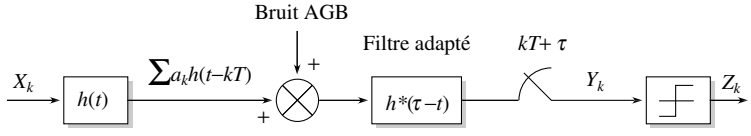

FIG. 4.10 – *Canal de transmission MIA-2.*

Dans le cas où on impose comme contrainte sur la loi d'entrée que l'énergie E_b soit proportionnelle à $\log_2(M)/n$, on a :

$$E_b = E_0\log_2(M)/n \leq E_0 C_H$$

Par conséquence l'expression de C_H en fonction de E_0 vérifie :

$$C_H = 1 + p\log_2(p) + (1-p)\log_2(1-p) \text{ où } p = Q\left(\sqrt{\frac{2E_0C_H}{N_0}}\right) \tag{4.8}$$

Canal binaire à décision douce Considérons à présent le canal dont l'alphabet d'entrée est $\mathcal{X} = \{0,1\}$, l'alphabet de sortie $\mathcal{Y} = \mathbb{R}$ et dont la probabilité de transition est donnée par :

$$p_{Y|X=0}(y) = \frac{1}{\sigma\sqrt{2\pi}}\exp(-(y+1)^2/2\sigma^2)$$

et

$$p_{Y|X=1}(y) = \frac{1}{\sigma\sqrt{2\pi}}\exp(-(y-1)^2/2\sigma^2)$$

où $\sigma^2 = N_0/2E_b$. En utilisant l'expression (4.6), on obtient :

$$C_S = 1 - \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}}\ e^{-u^2/2}\ \log_2(1 + e^{-2/\sigma^2}e^{2u/\sigma})du$$

En prenant la même contrainte que dans le cas à décision dure, l'expression de la capacité du canal binaire à décision douce en fonction de E_0 vérifie :

$$C_S = 1 - \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}}\ e^{-u^2/2}\ \log_2(1 + e^{-2/\sigma^2}e^{2u/\sigma})du \tag{4.9}$$

où $1/\sigma^2 = 2E_0C_S/N_0$.

Numériquement on vérifie que $C_S \geq C_H$. Ce résultat n'est pas surprenant. En effet notons :

$$X \in \{0,1\} \to Y \in \mathbb{R}$$

le canal binaire à décision douce. Et faisons le suivre d'un détecteur à seuil 0. On obtient :

$$X \in \{0,1\} \to Y \in \mathbb{R} \to Z \in \{0,1\}$$

qui est précisément le canal CBS. Or dans une telle cascade on ne peut que réduire la capacité. Une autre façon est de remarquer que Y permet de déduire Z alors que la réciproque est fausse : il y a donc nécessairement perte de l'information mutuelle par passage de Y à Z. Le théorème 4.5, que nous énonçons plus loin, donne une explication rigoureuse de ce résultat.

Le canal $X \to Z$ est dit à décision dure ou ferme (en anglais *hard decision*) tandis que le canal $X \to Y$ est dit à décision douce (en anglais *soft decision*). Dans le calcul du gain de codage nous verrons qu'en terme de probabilité d'erreur le canal à décision douce est meilleur que le canal à décision ferme.

Canal additif, Gaussien, blanc, de bande B l'alphabet d'entrée $\mathcal{X} = \mathbb{R}$ et l'alphabet de sortie $\mathcal{Y} = \mathbb{R}$. On impose que X soit centrée, de variance $P_X = \mathbb{E}\left\{X^2\right\} = E_0D_b$. Le bruit est supposé blanc dans la bande B, par conséquent sa puissance est donnée par $P_B = N_0B$. En utilisant l'expression (4.7), la capacité vérifie, en fonction de E_0 et N_0 :

$$\frac{2^{C_G} - 1}{C_G} = \frac{E_0}{N_0} \tag{4.10}$$

Nous avons représenté figure 4.11 les courbes de capacité correspondant aux expressions (4.8), (4.9), (4.10). On observe que le canal binaire symétrique à décision ferme a une capacité C_H plus faible que celle du canal à décision douce et que leur capacité respective est bornée par 1 ($C \leq \log_2(2) = 1$). Par contre celle du canal additif gaussien, blanc, dont l'entrée et la sortie sont infinies, peut prendre des valeurs supérieures à 1. On peut montrer que, quand E_0/N_0 tend vers 0, C_G et C_S tendent vers $10\log_{10}(\log(2)) \approx -1.59$.

Exemple 4.1 *On considère des modulations MIA-M sur le canal additif, Gaussien, blanc, de bande B. On suppose la condition de Nyquist vérifiée. Le rapport E_b/N_0 est fixé.*

1. *En se reportant aux courbes de probabilité d'erreur en fonction de E_b/N_0 pour différentes valeurs de M, comment doit-on faire pour diminuer la probabilité d'erreur ?*

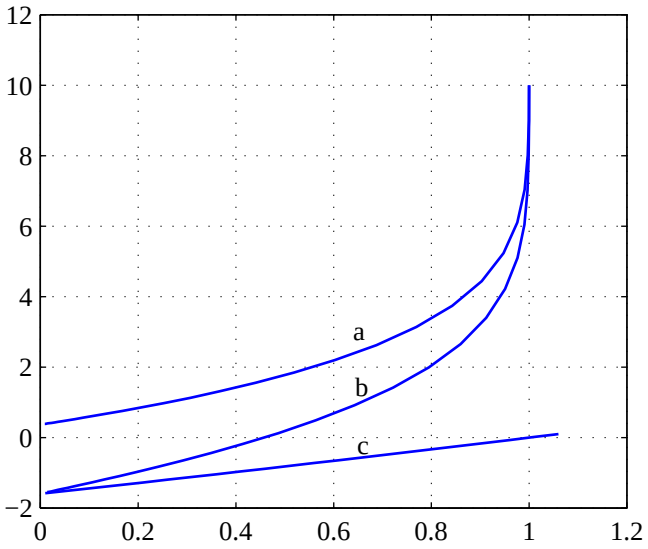

FIG. 4.11 – *Capacités en fonction de E_0/N_0 (en dB) pour une modulation MIA-2 sur le canal de Nyquist. Courbe (a) : alphabet d'entrée binaire et alphabet de sortie binaire. Courbe (b) : alphabet d'entrée binaire et alphabet de sortie $\mathbb{R}$. Courbe (c) : alphabet d'entrée $\mathbb{R}$ et alphabet de sortie $\mathbb{R}$.*

2. *Quelle est la conséquence sur le débit ?*
3. *Le résultat précédent est exact. En quoi est-il insatisfaisant ?*

Réponse :

1. Si on fixe E_b/N_0 pour diminuer la probabilité d'erreur, il faut diminuer M.
2. Comme $D_b = \log_2(M)/T$ et que $1/T = B/2$ (limite de la condition de Nyquist), $D_b = B\log_2(M)/2$. Il faut donc diminuer le débit.
3. Shannon nous apprend que, pour E_b/N_0 fixé, il existe une valeur η_M de l'efficacité spectrale maximale telle la probabilité d'erreur soit aussi près de 0 que l'on veut. Pour cette valeur de η_M le débit n'a pas besoin d'être diminué : il est égal à $D_b = B\eta_m$. Malheureusement ce n'est pas les MIA-M qu'il faut utiliser.

Exemple 4.2 (Code à répétitions sur le canal à décision douce) *Les conditions de codage sont les mêmes que dans l'exemple 3.1 :*

$$f : \quad \begin{array}{l} A \mapsto 000 \in \{0,1\}^3 \\ B \mapsto 111 \in \{0,1\}^3 \end{array}$$

Mais à présent à la sortie du canal, l'observation y est celle obtenue en sortie du filtre adapté pour une modulation MIA-2 sur le canal de Nyquist. Par conséquent $Y \in \mathbb{R}^3$ et :

$$p_{Y|X=000}(y_1,y_2,y_3) = \frac{1}{\sigma^3(2\pi)^{3/2}} \exp\left(-\frac{(y_1+1)^2+(y_2+1)^2+(y_3+1)^2}{2\sigma^2}\right)$$

et

$$p_{Y|X=111}(y_1,y_2,y_3) = \frac{1}{\sigma^3(2\pi)^{3/2}} \exp\left(-\frac{(y_1-1)^2+(y_2-1)^2+(y_3-1)^2}{2\sigma^2}\right)$$

*où $\sigma^2 = N_0/2E_b$. En utilisant (**??**), on montre aisément que la fonction de décision qui minimise la probabilité d'erreur moyenne est :*

$$g : \quad \begin{array}{l} d_E(y,y_A) < d_E(y,y_B) \mapsto A \\ d_E(y,y_B) < d_E(y,y_A) \mapsto B \end{array}$$

où $d_E(x,y)$ désigne la distance euclidienne dans $\mathbb{R}^3$ et $y_A = (-1,-1,-1)$ et $y_B = (1,1,1)$. On en déduit la probabilité d'erreur moyenne :

$$P_e = Q\left(\frac{d_E(y_A, y_B)}{2\sigma_B}\right) = Q(\sqrt{6E_b/N_0})$$

Le gain est plus important dans le cas où on effectue une décision "douce" entre A et B à partir de $y \in \mathbb{R}^3$ plutôt que de décider de façon ferme après chaque utilisation du canal et ensuite de décider entre A et B en minimisant la distance de Hamming c'est-à-dire le nombre d'éléments qui diffèrent.

Nous avons représenté figure 4.12, les régions de décision pour une décision ferme et une décision douce. Dans le cas de la décision douce, la séparatrice est le plan médiateur des points représentatifs des deux messages. Dans le cas de la décision ferme, où on compare tout d'abord les 3 observations à 0 puis où on applique un minimum de distance de Hamming, la séparatrice est constituée de 6 demi-plans. Dans ce cas, la probabilité d'erreur moyenne est plus grande que celle obtenue avec le plan médiateur.

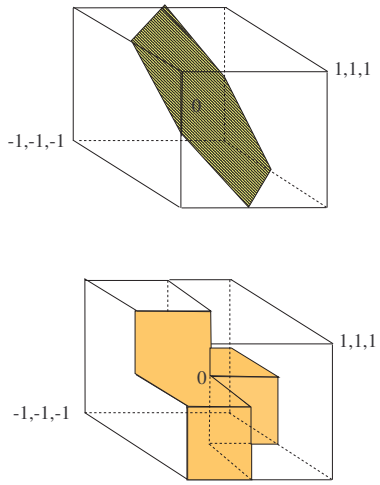

FIG. 4.12 – *Régions de décision en décisions ferme et douce sur le canal gaussien.*

4.2 Outils de la théorie de l'information

4.2.1 Quantité d'information

De façon à pouvoir calculer précisément le coût d'une transmission, C. Shannon a été conduit à se poser la question de savoir comment définir la quantité d'information. Sans être un grand cruciverbiste, il est clairement plus facile, en français, de trouver un mot commençant par les deux lettres w et h qu'un mot commençant par les deux lettres d et e. On peut dire que "wh" contient plus d'information que "de". Et donc, indépendamment de l'usage que l'on fait d'une information ou du contenu sémantique d'un message, on peut dire que ce qui est rare contient plus d'information que ce qui est fréquent. Définir une quantité d'information revient donc à définir une mesure de l'incertain. En communication c'est la même idée qui est contenue dans la remarque suivante apparemment très banale : "si le message était parfaitement connu du destinataire on n'aurait pas besoin de le lui transmettre". Partant de là, C. Shannon a adopté la définition suivante :

Définition 4.6 (Quantité d'information) *Soit un espace de probabilité $\{\Omega, \mathcal{F}, \mathbb{P}\}$ et un événement $E \in \mathcal{F}$. On appelle quantité d'information contenue dans E la valeur positive donnée par $h(E) = -\log_2(\mathbb{P}\{E\})$*[2].

Notons que, par suite de l'utilisation de la fonction logarithme, la quantité d'information de l'union d'événements indépendants est la somme des quantités d'information respectives de chaque événement.

[2]Lorsque la base du logarithme est 2 l'unité est le *bit*. Lorsque la base du logarithme est e l'unité est le *nat*. Dans la suite nous utiliserons uniquement le *bit*.

Définition 4.7 (Entropie) *Soit X une variable aléatoire à valeurs dans un ensemble $\mathcal{X} = \{1, \cdots, M\}$ de cardinalité finie M. On note $p_X(x) = \mathbb{P}\{X = x\}$ où $x \in \mathcal{X}$ et $h(X) = -\log_2(p_X(X))$. On appelle entropie la quantité :*

$$H(X) = \mathbb{E}\{h(X)\} = -\sum_{x=1}^{M} p_X(x) \log_2(p_X(x))$$

Par convention on prend $0 \log_2(0) = 0$.

Le lemme suivant est à la base de la démonstration de plusieurs inégalités portant sur l'entropie.

Lemme 4.1 (Fondamental) *Soit $\mathbb{P}$ et $\mathbb{Q}$ deux mesures de probabilité définies sur le même espace dénombrable $\mathcal{X}$ muni de l'ensemble de ses parties. $\mathbb{P}$ et $\mathbb{Q}$ sont caractérisées par la mesure des singletons de $\mathcal{X}$. On note, pour $x \in \mathcal{X}$, $p(x) = \mathbb{P}\{x\}$ et $q(x) = \mathbb{Q}\{x\}$. Alors*

$$K(\mathbb{P}, \mathbb{Q}) = \sum_{x \in \mathcal{X}} p(x) \log_2 \left(\frac{p(x)}{q(x)} \right) \geq 0$$

L'égalité a lieu si et seulement si $\mathbb{P}$ et $\mathbb{Q}$ coïncident.

Il suffit d'utiliser l'inégalité de Jensen qui dit que :

Théorème 4.4 (Jensen) *Si f est une fonction convexe et U est une variable aléatoire, alors $\mathbb{E}\{f(U)\} \geq f(\mathbb{E}\{U\})$. De plus si la fonction f est strictement convexe, alors $\mathbb{E}\{f(U)\} = f(\mathbb{E}\{U\})$ implique que $U = \mathbb{E}\{U\}$ en probabilité (c'est-à-dire que X est constant en probabilité).*

Pour démontrer le lemme 4.1, il suffit d'appliquer le théorème 4.4 à la fonction strictement convexe $f(U) = -\log_2(U)$ pour la variable aléatoire $U(X) = q(X)/p(X)$. Il vient :

$$\begin{aligned} K(\mathbb{P}, \mathbb{Q}) &= \mathbb{E}\left\{ -\log_2 \left(\frac{q(X)}{p(X)} \right) \right\} \geq -\log_2 \left(\mathbb{E}\left\{ \frac{q(X)}{p(X)} \right\} \right) \\ &= -\log_2 \left(\sum_{x \in \mathcal{X}} p(x) \frac{q(x)}{p(x)} \right) \\ &= -\log_2 (\sum_{x \in \mathcal{X}} q(x)) = 0 \end{aligned}$$

L'égalité a lieu si et seulement si $q(X)/p(X)$ est égale à la constante. Comme $p(x)$ et $q(x)$ sont normés, cette constante vaut 1 et on $p(x) = q(x)$ (l'égalité a lieu en probabilité).

$K(\mathbb{P}, \mathbb{Q})$ s'appelle la *divergence de Kullback* entre les deux mesures. Noter que ce n'est pas une distance, en particulier $K(\mathbb{P}, \mathbb{Q}) \neq K(\mathbb{Q}, \mathbb{P})$.

Propriétés 4.2

1. $H(X) \geq 0$
2. *Si le cardinal de $\mathcal{X}$ est fini et égal à M, alors $H(X) \leq \log_2(M)$, l'égalité ayant lieu pour l'équidistribution,*
3. $\mathcal{H}(P) = -\sum_{i=1}^{M} p_i \log_2(p_i)$ *est une fonction continue et concave sur l'ensemble convexe $\{P = (p_1, \cdots, p_M); p_i > 0, \sum_i p_i = 1\}$.*

Pour démontrer le point 2, il suffit d'appliquer le lemme 4.4 en prenant $q(x) = 1/M$.

Exemple 4.3 *On considère une variable aléatoire discrète X à valeurs dans un ensemble de cardinalité M avec les probabilités $(p_1, p_2, \ldots, p_M)$. Montrer que :*

$$H(X) \leq h(p_1) + (1 - p_1) \log(M - 1)$$

où $h(x) = -x \log(x) - (1 - x) \log(1 - x)$.

Réponse : on considère la loi de probabilité sur $\mathcal{X}$, définie par $q_1 = p_1$, $q_i = (1-p_1)/(M-1)$ pour $2 \leq i \leq M$. Puis on applique le lemme fondamental. Il vient :

$$\begin{aligned} H(X) &\leq -\sum_{i=1}^{M} p_i \log(q_i) \\ &= -p_1 \log(p_1) - (1-p_1)\log(1-p_1) + (p_2 + \cdots + p_M)\log(M-1) \end{aligned}$$

qui est le résultat demandé.

4.2.2 Information Mutuelle

Définition 4.8 (Entropie conjointe) *Soit X, Y deux variables aléatoires discrètes à valeurs dans $\mathcal{X} \times \mathcal{Y}$ de loi conjointe $p_{XY}(x,y) = \mathbb{P}\{X = x, Y = y\}$ où $(x,y) \in \mathcal{X} \times \mathcal{Y}$. On note $h(X,Y) = -\log(p_{XY}(X,Y))$. On appelle entropie conjointe :*

$$H(X,Y) = \mathbb{E}\{h(X,Y)\} = -\sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p_{XY}(x,y) \log_2(p_{XY}(x,y))$$

Définition 4.9 (Entropie conditionnelle) *Soit X, Y deux variables aléatoires discrètes à valeurs dans $\mathcal{X} \times \mathcal{Y}$ de loi conjointe $p_{XY}(x,y) = \mathbb{P}\{X = x, Y = y\}$ où $(x,y) \in \mathcal{X} \times \mathcal{Y}$. On note $h(X|Y) = -\log(p_{X|Y}(X,Y))$. On appelle entropie conditionnelle de X sachant Y :*

$$H(X|Y) = \mathbb{E}\{h(X|Y)\} = -\sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p_{XY}(x,y) \log_2(p_{X|Y}(x,y))$$

On rappelle que $p_{X|Y}(x,y) = p_{XY}(x,y)/p_Y(y)$ et que $p_Y(y) = \sum_{x \in \mathcal{X}} p_{XY}(x,y)$. D'une certaine façon $H(X|Y)$ mesure ce qui reste d'incertitude sur X lorsque Y a été observé. Il est raisonnable de penser que si X et Y sont indépendants alors $H(X|Y) = H(X)$. En effet cette propriété est vraie. De même si $H(X|Y) = 0$ on est en droit de penser que $X = g(Y)$.

Définition 4.10 (Information mutuelle) *Soit X, Y deux variables aléatoires discrètes à valeurs dans $\mathcal{X} \times \mathcal{Y}$ de loi conjointe $p_{XY}(x,y) = \mathbb{P}\{X = x, Y = y\}$ où $(x,y) \in \mathcal{X} \times \mathcal{Y}$. On appelle information mutuelle :*

$$I(X,Y) = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p_{XY}(x,y) \log_2 \left(\frac{p_{XY}(x,y)}{p_X(x) p_Y(y)} \right)$$

$I(X,Y)$ s'interprète comme une mesure d'indépendance entre X et Y. On remarque en effet que si X et Y sont indépendantes, $I(X,Y) = 0$. La réciproque est vraie : en effet, partant du lemme fondamental appliqué aux deux lois de probabilité $p(u) = p_{XY}(x,y)$ et $q(u) = p_X(x)p_Y(y)$ définies sur le même espace produit $\mathcal{X} \times \mathcal{Y}$, on déduit que $I(X,Y) \geq 0$ et que l'égalité a lieu si et seulement si $p(u) = q(u)$, c'est-à-dire si les deux variables sont indépendantes.

Propriétés 4.3
Pour des sources discrètes on a :

1. $H(X,Y) \geq 0$
2. $H(X|Y) \geq 0$
3. $H(g(X)|X) = 0 \Rightarrow H(g(X)) \leq H(X)$
4. $H(X,Y) = H(X) + H(Y|X) = H(Y) + H(X|Y)$
5. $H(X,Y) = H(X) + H(Y) - I(X,Y)$
6. $I(X,Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)$
7. $I(X,Y) \geq 0$, *l'égalité ayant lieu ssi X et Y sont indépendantes,*
8. $H(X,Y) = H(X) + H(Y) \Leftrightarrow X$ *et Y sont indépendantes,*
9. $H(X) = H(X|Y)$ *(ou $H(Y) = H(Y|X)$)* $\Leftrightarrow X$ *et Y sont indépendantes.*
10. $H(X,Y|Z) = H(X|Z) + H(Y|X,Z) = H(Y|Z) + H(X|Y,Z)$

Certaines des propriétés données ci-dessus peuvent se retrouver en utilisant le diagramme représenté figure 4.13. De façon imagée, la surface de l'ensemble X mesure l'incertitude $H(X)$, la surface de $X-(X\cap Y)$ mesure $H(X|Y)$, c'est-à-dire ce qu'il reste d'incertitude sur X une fois que Y a été observé. La surface de $(X\cap Y)$ mesure $I(X,Y)$, la surface de $X\cup Y$ l'incertitude $H(X,Y)$ conjointe du couple (X,Y). Partant de la, on retrouve simplement que $H(X,Y)=H(X)+H(Y)-I(X,Y)$, ou encore $I(X,Y)\leq H(Y)$. Attention dans ce diagramme les objets ne représentent pas des événements. Ainsi $I(X,Y)=0$ si et seulement si X et Y sont *indépendants*, soit dans le diagramme si et seulement si $X\cap Y=\emptyset$, ce qui correspond, en termes d'événements, à des événements *incompatibles*.

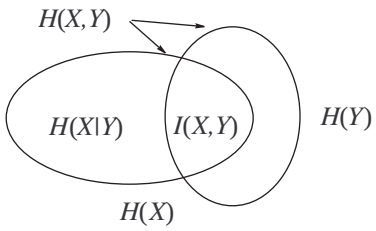

FIG. 4.13 – *Diagramme de Venn*

Exemple 4.4 ($H(g(X))\leq H(X)$) *On suppose que X est une variable aléatoire à valeurs discrètes dans $\mathcal{X}$ et soit $Y=g(Y)$.*

1. *Déterminer la loi conditionnelle de Y sachant X. En déduire que $H(g(X)|X)=0$.*
2. *En utilisant la propriété 4 des propriétés 4.3 à $H(X,g(X))$, déduire que $H(g(X))\leq H(X)$.*

Réponse :

1. On note $\mathcal{Y}$ l'ensemble des valeurs prises par Y. On note $g^{-1}(y_j)=\{x_k\in\mathcal{X}:g(x_k)=y_j\}$ (si g n'est pas bijective, $g^{-1}(y_j)$ peut contenir plus d'un élément. Alors pour tout couple $(x_i,y_j)\in\mathcal{X}\times\mathcal{Y}$, on a :

$$\begin{aligned}\mathbb{P}\{Y=y_j|X=x_i\} &= \frac{\mathbb{P}\left\{X\in g^{-1}(y_j),X=x_i\right\}}{\mathbb{P}\{X=x_i\}}\\ &= \begin{cases}1 & \text{si} \quad x_i\in g^{-1}(y_j)\\ 0 & \text{sinon}\end{cases}\end{aligned}$$

   Par conséquent :

$$H(Y|X=x_i)=-\sum_{y_j\in\mathcal{Y}}\mathbb{P}\{Y=y_j|X=x_i\}\log(\mathbb{P}\{Y=y_j|X=x_i\})=0$$

   Partant alors de la définition de $H(Y|X)$, on obtient :

$$\begin{aligned}H(Y|X) &= -\sum_{x_i\in\mathcal{X}}\mathbb{P}\{X=x_i\}\sum_{y_j\in\mathcal{Y}}\mathbb{P}\{Y=y_j|X=x_i\}\log(\mathbb{P}\{Y=y_j|X=x_i\})\\ &= \sum_{x_i\in\mathcal{X}}\mathbb{P}\{X=x_i\}H(Y|X=x_i)=0\end{aligned}$$

   En conclusion $H(g(X)|X)=0$.
2. D'après la propriété 4 des propriétés 4.3, on a $H(X,g(X))=H(X)+H(g(X)|X)=H(g(X))+H(X|g(X))$. Et donc $H(X)+0=H(g(X))+H(X|g(X))\geq H(g(X))$.

La définition de l'entropie conjointe se généralise à un nombre n de variables :

Définition 4.11 (Entropie conjointe) *Soit $X_1,\ldots,X_n$, n variables aléatoires discrètes, définies sur le même espace de probabilité, à valeurs respectivement dans $\mathcal{X}_1\ldots\mathcal{X}_n$. Pour $(x_1,\cdots,x_n)\in\mathcal{X}_1\times\cdots\times\mathcal{X}_n$, on note :*

$$p_{X_1\ldots X_n}(x_1,\ldots,x_n)=\mathbb{P}\{X_1=x_1,\ldots,X_n=x_n\}$$

et $h(X_1,\cdots,X_n) = -\log_2(p_{X_1\ldots X_n}(X_1,\ldots,X_n))$ *On appelle entropie conjointe :*

$$\begin{aligned} H(X_1,\ldots,X_n) &= \mathbb{E}\{h(X_1,\cdots,X_n)\} \\ &= -\sum_{\mathcal{X}_1}\cdots\sum_{\mathcal{X}_n} p_{X_1\ldots X_n}(x_1,\ldots,x_n)\log_2(p_{X_1\ldots X_n}(x_1,\ldots,x_n)) \end{aligned}$$

Avec des notations évidentes, on montre aisément que :

$$H(X_1,\ldots,X_n) = \sum_{i=1}^{n} H(X_i|X_{i-1},\cdots,X_1)$$

où

$$\begin{aligned} &H(X_i|X_{i-1},\cdots,X_1) \\ &\quad = -\sum_{\mathcal{X}_1}\cdots\sum_{\mathcal{X}_i} p_{X_1\ldots X_i}(x_1,\ldots,x_i)\log_2(p_{X_i|X_{i-1},\cdots,X_1}(x_1,\ldots,x_i)) \end{aligned}$$

Propriétés 4.4 *On a* $H(X_1,X_2,\ldots,X_n) = \sum_{i=1}^{n} H(X_i)$ *si et seulement si les* n *variables aléatoires* $(X_1,X_2,\ldots,X_n)$ *sont indépendantes.*

Ce résultat est la généralisation immédiate de la propriété 8 des propriétés 4.3.

4.2.3 Théorème du traitement de l'information

Définition 4.12 (Cascade markovienne) *Soit* X*,* Y *et* Z *trois variables aléatoires définies sur le même espace de probabilité. On dit que* $X \to Y \to Z$ *forment une cascade markovienne si et seulement si la loi conditionnelle de* Z *sachant* X *et* Y *coïncident avec la loi conditionnelle de* Z *sachant* Y*.*

On rappelle le résultat général suivant : si A, B, C désignent trois événements quelconques alors :

$$\mathbb{P}\{A,B,C\} = \mathbb{P}\{C|A,B\}\,\mathbb{P}\{B|A\}\,\mathbb{P}\{C\}$$

Dans un cascade markovienne $\mathbb{P}\{C|A,B\} = \mathbb{P}\{C|B\}$ et, par conséquent, si X, Y et Z désignent trois variables aléatoires à valeurs dans les ensembles discrets $\mathcal{X}$, $\mathcal{Y}$ et $\mathcal{Z}$, on a pour tout $x \in \mathcal{X}$, $y \in \mathcal{Y}$ et $z \in \mathcal{Z}$:

$$\mathbb{P}\{X=x,Y=y,Z=z\} = \mathbb{P}\{Z=z|Y=y\}\,\mathbb{P}\{Y=y|X=x\}\,\mathbb{P}\{X=x\}$$

La définition 4.12 est aussi équivalente à dire que, conditionnellement à Y, X et Z sont indépendantes, ce qui s'écrit :

$$\mathbb{P}\{X=x,Z=z|Y=y\} = \mathbb{P}\{X=x|Y=y\}\,\mathbb{P}\{Z=z|Y=y\}$$

Propriétés 4.5

- *Si* $X \to Y \to Z$*, alors* $Z \to Y \to X$*,*
- *Si* $Z = g(Y)$*, alors* $X \to Y \to Z$*.*

En effet on a simultanément :

$$\begin{aligned} \mathbb{P}\{X=x,Y=y,Z=z\} &= \mathbb{P}\{Z=z|Y=y\}\,\mathbb{P}\{Y=y,X=x\} \\ &= \mathbb{P}\{X=x|Y=y,Z=z\}\,\mathbb{P}\{Z=z|Y=y\}\,\mathbb{P}\{Y=y\} \end{aligned}$$

et donc $\mathbb{P}\{X=x|Y=y,Z=z\} = \mathbb{P}\{Y=y,X=x\}/\mathbb{P}\{Y=y\} = \mathbb{P}\{X=x|Y=y\}$. Pour démontrer la seconde propriété, il suffit de remarquer que, si $Z=g(Y)$, $\mathbb{P}\{Z=z|Y=y\} = \mathbb{P}\{Z=z,Y=y\}/\mathbb{P}\{Y=y\}$ est égal à 1 ou 0 suivant que $z=g(y)$ ou $z \neq g(y)$.

Propriétés 4.6 *Soit X, Y et Z trois variables aléatoires définies sur le même espace, à valeurs respectivement dans $\mathcal{X}$, $\mathcal{Y}$ et $\mathcal{Z}$. On note :*

$$I((X,Y),Z) = \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}}\sum_{z\in\mathcal{Z}} \mathbb{P}\{X=x,Y=y,Z=z\} \log\left(\frac{\mathbb{P}\{Z=z|X=x,Y=y\}}{\mathbb{P}\{Z=z\}}\right)$$

Alors :

$$I((X,Y),Z) \geq I(Y,Z)$$

l'égalité ayant lieu si et seulement si $X \to Y \to Z$.

En effet :

$$I((X,Y),Z) - I(Y,Z) = \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}}\sum_{z\in\mathcal{Z}} \mathbb{P}\{X=x,Y=y,Z=z\} \log\left(\frac{\mathbb{P}\{Z=z|X=x,Y=y\}\,\mathbb{P}\{Z=z\}}{\mathbb{P}\{Z=z\}\,\mathbb{P}\{Z=z|Y=y\}}\right)$$

$$= \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} \mathbb{P}\{X=x,Y=y\}\left(\sum_{z\in\mathcal{Z}} \mathbb{P}\{Z=z|X=x,Y=y\} \log\left(\frac{\mathbb{P}\{Z=z|X=x,Y=y\}}{\mathbb{P}\{Z=z|Y=y\}}\right)\right)$$

D'après le lemme fondamental le terme dans la dernière parenthèse est positif et est nul si et seulement si $\mathbb{P}\{Z=z|X=x,Y=y\}$ coïncide avec $\mathbb{P}\{Z=z|Y=y\}$. Notons ici que, dans la propriété 4.6, X et Y jouent un rôle symétrique et que, par conséquent, on a toujours :

$$I((X,Y),Z) \geq \max\{I(X,Z), I(Y,Z)\}$$

Théorème 4.5 (Traitement de l'information) *Soit $X \to Y \to Z$ une cascade markovienne. Alors :*

$$I(X,Z) \leq \min\{I(X,Y), I(Y,Z)\}$$

En effet, d'après la propriété 4.6, on a à la fois $I((X,Y),Z) \geq I(X,Z)$ et $I((X,Y),Z) = I(Y,Z)$.

On en déduit que, si $X \to Y \to g(Y)$, alors on a :

$$I(X,g(Y)) \leq I(X,Y)$$

4.2.4 Cas de variables aléatoires "continues"

Dans ce paragraphe on s'intéresse à des sources sans mémoire, dont la loi de probabilité possède une densité $p_X(x)$ par rapport à la mesure de Lebesgue.

Définition 4.13 (Entropie différentielle) *Soit X une variable aléatoire à valeurs dans $\mathbb{R}^k$ dont la loi possède une densité de probabilité $p_X(x)$ par rapport à la mesure de Lebesgue dans $\mathbb{R}^k$. On appelle entropie différentielle :*

$$H_d(X) = -\int_{\mathbb{R}^k} p_X(x)\log_2(p_X(x))dx$$

Contrairement au cas discret où l'entropie est une quantité toujours positive, l'entropie différentielle $H_d(X)$ peut être *négative*.

Théorème 4.6 (Lemme fondamental) *Soit deux mesures de probabilité $\mathbb{P}$ et $\mathbb{Q}$ définies sur $\mathbb{R}^k$ et possédant des densités, par rapport à la mesure de Lebesgue dans $\mathbb{R}^k$, que l'on note respectivement $p(x)$ et $q(x)$. Alors on a :*

$$K(\mathbb{P},\mathbb{Q}) = \int_{\mathbb{R}^k} p(x)\log_2\left(\frac{p(x)}{q(x)}\right)dx \geq 0 \tag{4.11}$$

(4.11) est une conséquence directe de l'inégalité de Jensen appliquée à la fonction $f(X) = q(X)/p(X)$, l'égalité ayant lieu si et seulement si les deux lois coïncident. Soulignons que les lois $p(x)$ et $q(x)$ peuvent être multi-dimensionnelles.

Définition 4.14 (Information mutuelle) *Soit (X, Y) deux variables aléatoires définies sur le même espace de probabilité à valeurs dans $\mathbb{R}^2$ dont la loi conjointe possède une densité de probabilité $p_{XY}(x, y)$ par rapport à la mesure de Lebesgue dans $\mathbb{R}^2$. On appelle information mutuelle :*

$$I(X,Y) = \int_{\mathbb{R}} \int_{\mathbb{R}} p_{XY}(x,y) \log_2 \left(\frac{p_{XY}(x,y)}{p_X(x) p_Y(y)} \right) dxdy$$

Propriétés 4.7 *on a :*

$$I(X,Y) \geq 0$$

l'égalité ayant lieu si et seulement si X et Y sont indépendantes.

Il suffit d'appliquer le lemme fondamental aux densités $p(u) = p_{XY}(x, y)$ et $q(u) = p_X(x)p_Y(y)$ définies sur $\mathbb{R}^2$.

Propriétés 4.8
On a aussi :
- $H(X,Y) = H(X) + H(Y|X) = H(Y) + H(X|Y)$
- $H(X,Y) = H(X) + H(Y) - I(X,Y)$
- $I(X,Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)$

Exemple 4.5 (Maximum d'entropie : loi gaussienne) *Soit X une variable aléatoire à valeurs dans $\mathbb{R}$, centrée, de variance σ_X^2 fixée. Déduire de l'inégalité* (4.11) *que la loi qui maximise l'entropie, sous contrainte de moyenne nulle et de variance σ_X^2 fixée, est la loi gaussienne.*

Réponse : Appliquons l'inégalité (4.11) aux densités $p(x) = p_X(x)$ et $q(x) = (2\pi\sigma_X^2)^{-1/2} e^{-x^2/2\sigma_X^2}$. Il vient :

$$\begin{aligned} H(X) &= -\int_{\mathbb{R}} p_X(x) \log_2(p_X(x)) dx \\ &\leq -\int_{\mathbb{R}} p_X(x) \log_e(q(x)) \log_2(e) dx = \frac{1}{2} \log_2(2\pi\sigma_X^2 e) \end{aligned}$$

Par conséquent, quelle que soit la loi centrée de variance σ_X^2, $H(X) \leq \log_2(\sigma_X \sqrt{2\pi e})$. Cette borne, *indépendante* de la loi, peut être atteinte en prenant pour loi de X la loi gaussienne. C'est donc la loi gaussienne qui maximise l'entropie sous contraintes que la moyenne est nulle et la variance égale à σ_X^2.

Exemple 4.6 (Capacité du canal gaussien) *Soit $Y = X + B$ où X et B sont deux variables aléatoires indépendantes. Montrer que $I(X,Y) = H(Y) - H(B)$. En déduire la capacité du canal gaussien.*

Réponse : En effet $I(X,Y) = H(Y) - H(Y|X)$. Mais, comme X et B sont indépendantes, la loi $p_{Y|X}(y) = p_B(y - x)$ (dans le cas où X et B sont indépendantes, il suffit de "fixer" $X = x$ dans $Y = X + B$.). Par conséquent :

$$H(Y|X) = \int_{\mathbb{R}} \int_{\mathbb{R}} p_X(u) p_B(v-u) \log_2(p_B(v-u)) du dv = \int_{\mathbb{R}} p_X(u) g(u) du$$

où

$$\begin{aligned} g(u) &= \int_{\mathbb{R}} p_B(v-u) \log_2(p_B(v-u)) dv = \int_{\mathbb{R}} p_B(\theta) \log_2(p_B(\theta)) d\theta \\ &= -\frac{1}{2} \log_2(2\pi e \sigma_B^2) = H(B) \end{aligned}$$

et donc $I(X,Y) = H(Y) - H(B)$. Sous contrainte que $\mathbb{E}\{X\} = 0$ et $\mathbb{E}\{X^2\} = \sigma_X^2$, on a $\mathbb{E}\{Y\} = 0$ et $\mathbb{E}\{Y^2\} = \sigma_X^2 + \sigma_B^2$. Par conséquent pour maximiser $I(X,Y)$ il faut maximiser $H(Y)$ sous contrainte que $\mathbb{E}\{Y\} = 0$ et $\mathbb{E}\{Y^2\} = \sigma_X^2 + \sigma_B^2$. On sait, d'après l'exercice 4.5, que le maximum est atteint par la loi gaussienne. Cette solution est acceptable : il suffit que X soit une variable gaussienne. On a alors :

$$\max_{\{\mathcal{P}_X \text{ t.q. } \mathbb{E}\{X\}=0,\ \mathbb{E}\{X^2\}=\sigma_X^2\}} I(X,Y) = \frac{1}{2} \log_2 \left(\frac{\sigma_X^2 + \sigma_B^2}{\sigma_B^2} \right) = \frac{1}{2} \log_2 \left(1 + \frac{\sigma_X^2}{\sigma_B^2} \right)$$

4.3 Exercices

Exercice 4.1 (Canal en Z) *On considère le canal dont les alphabets d'entrée et de sortie sont binaires et dont les probabilités de transition sont :*

$$\mathbb{P}\{Y=1|X=1\}=q \text{ et } \mathbb{P}\{Y=0|X=0\}=1$$

Ce canal est dit canal en Z. Il peut modéliser un canal physique optique dans lequel l'absence de lumière (entrée 0) ne peut pas donner lieu à une détection lumineuse.

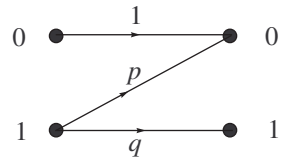

FIG. 4.14 – *Canal en Z. $p+q=1$*

1. *Déterminer l'expression de sa capacité.*
2. *On adopte le protocole suivant : si on observe $Y=1$, on décide $\hat{X}=1$, par contre, si on observe $Y=0$, on demande, au moyen d'un canal auxiliaire dont on négligera la capacité (ce qui est justifié si p est supposé petit), à l'émetteur de retransmettre le symbole d'entrée mais en commutant sa valeur. Déterminer les expressions respectives $\bar{\ell}_0$ et $\bar{\ell}_1$ du nombre moyen d'utilisation du canal, si le symbole à émettre est respectivement 0 et 1.*
3. *En déduire le nombre moyen de bits transmis par utilisation du canal. Comparer à la capacité.*

Exercice 4.2 *On considère le canal dont l'entrée X est à valeurs dans $\{0,1\}$ et dont la sortie s'écrit $Y=X+Z$ où Z est une variable aléatoire à valeurs dans $\{0,a\}$ avec $\mathbb{P}\{Z=0\}=\mathbb{P}\{Z=a\}=1/2$. On suppose que X et Z sont indépendantes. Déterminer l'expression de la capacité en fonction de a.*

Exercice 4.3 *Soit les deux canaux $\{\mathcal{X}_1, p(y_1|x_1), \mathcal{Y}_1\}$ et $\{\mathcal{X}_2, p(y_2|x_2), \mathcal{Y}_2\}$. On note C_1 et C_2 leurs capacités respectives. On considère à présent le canal $\{\mathcal{X}_1\times\mathcal{X}_2, p(y_1|x_1)p(y_2|x_2), \mathcal{Y}_1\times\mathcal{Y}_2\}$. Déterminer l'expression de sa capacité en fonction de C_1 et C_2.*

Exercice 4.4 *On considère la mise en cascade de n canaux binaires symétriques de même probabilité d'erreur p. Montrer que le canal obtenu est un canal binaire symétrique dont on déterminera l'expression de sa probabilité d'erreur en fonction de p. Vers quelle valeur tend sa capacité quand n tend vers l'infini ?*

Exercice 4.5 *On considère le canal dont les alphabets d'entrée et de sortie comportent 5 symboles notés $\{0,1,2,3,4\}$ et dont les probabilités de transition sont :*

$$\mathbb{P}\{Y=i|X=j\}=\begin{cases} 1/2 & si \quad i=j\pm 1 \bmod 5 \\ 0 & sinon \end{cases}$$

1. *Déterminer l'expression de sa capacité.*
2. *Clairement ce canal transmet sans erreur au moins 1 bit. Trouver un code en blocs qui montre que la capacité est supérieure à 1.*

Exercice 4.6 *Soit X une variable aléatoire prenant un nombre fini de valeurs. On note $H(X)$ son entropie. Déterminer, en fonction de $H(X)$, l'expression de l'entropie $H(Y)$ de la variable aléatoire Y dans les deux cas suivants :*

1. *$Y=e^X$.*
2. *$Y=\cos(X)$.*

Exercice 4.7 *Trouver la loi de probabilité de la variable aléatoire X à valeurs dans $\mathbb{N}$ qui maximise l'entropie sous la contrainte que $\mathbb{E}\{X\}=A>0$.*

Exercice 4.8 *Montrer que l'entropie de la distribution $\{p_1, \ldots, p_N\}$ est inférieure à celle de la distribution $\{p_1, \ldots, p_N\}$ où on a remplacé p_i et p_j par $q = (p_i + p_j)/2$.*

Exercice 4.9 *On considère les variables aléatoires X et Y à valeurs dans $\{0, 1\}$ dont la loi conjointe est donnée par :*

$$\mathbb{P}\{X = 0, Y = 0\} = 1/4, \ \mathbb{P}\{X = 0, Y = 1\} = 1/4, \ \mathbb{P}\{X = 1, Y = 0\} = 1/2$$

Calculer $H(X)$, $H(Y)$, $H(X, Y)$, $H(Y|X)$, $H(X|Y)$, $I(X, Y)$.

Exercice 4.10 (Inégalité de Fano) *On considère deux variables aléatoires discrètes X et Y. On note $M = |\mathcal{X}|$ la cardinalité de l'ensemble des valeurs de X. On considère la variable aléatoire $\hat{X} = g(Y)$ obtenue à partir de Y par la fonction (mesurable) g. On pose $E = \mathbb{1}(\hat{X} \neq X)$ et on note $P_e = \mathbb{P}\{E = 1\}$.*

1. *Montrer que $H(E|X, Y) = 0$ (indication : utiliser le résultat de l'exemple 4.4).*
2. *En déduire que $H(X|Y) = H(E|Y) + H(X|E, Y)$.*
3. *Montrer que $H(E|Y) \leq -P_e \log(P_e) - (1 - P_e) \log(1 - P_e)$.*
4. *Montrer que $H(X|Y, E) = H(X|Y, E = 0)\mathbb{P}\{E = 0\} + H(X|Y, E = 1)\mathbb{P}\{E = 1\} \leq P_e \log(M - 1)$.*

En déduire l'inégalité de Fano :

$$H(P_e) + P_e \log(M - 1) \geq H(X|Y)$$