UTILISATION OPTIMALE D'UN DÉTECTEUR. APPLICATION AU SONAR.

par

Hubert DEBART *

INTRODUCTION

Il est bien connu que la détection constitue une opération de choix ; pratiquement, un détecteur choisit entre un nombre fini de signaux possibles. Quand on reçoit un signal, on ne peut se passer de cette opération. Il peut exister des traitements ultérieurs, en particulier pour utiliser la redondance statistique entre plusieurs signaux détectés indépendamment. Mais les traitements en question ne peuvent évidemment que détériorer la quantité d'information qui sort du détecteur proprement dit. On a donc intérêt à la rendre maximale.

Si on appelle m le nombre de signaux entre lesquels le détecteur peut choisir, ce nombre peut être qualifié de pouvoir séparateur. Nous verrons que ce pouvoir séparateur définit, dans des conditions données, une quantité d'information maximale pouvant être extraite du détecteur en un temps donné ; il ne sert à rien d'élargir la bande spectrale transmise par le canal si on n'augmente pas en même temps ce pouvoir séparateur. Nous verrons ensuite comment ce point de vue peut se raccorder à celui de Shannon par un passage à la limite.

I. TAUX D'ERREUR FOURNI PAR UN DÉTECTEUR OPTIMAL DE SIGNAUX DE POUVOIR SÉPARATEUR DONNÉ.

On peut optimiser un détecteur, de pouvoir séparateur m, c'est-à-dire construire en minimisant le taux d'erreur à sa sortie, quand on se donne :

— la puissance de bruit par Hz, σ^2 du bruit (en répartissant l'énergie sur les fréquences positives seulement) ; quantité homogène à une énergie ;

— les énergies affectées aux signaux $A_1(t) \ldots A_m(t)$ soient :

$$\int_T A_1^2\, dt \ldots \int_T A_m^2\, dt\,;$$

— les produits scalaires des signaux entre eux, soient :

$$\int_T A_i\, A_j\, dt\,;$$

— les probabilités de présence des signaux.

La théorie de la détection optimale a été exposée, en particulier par V. A. Kotelnikov (*), pour des signaux orthogonaux entre eux ; elle se généralise facilement au cas de signaux non orthogonaux. Nous supposons que toutes les distances des signaux entre eux sont égales, les signaux étant aussi de même énergie. C'est-à-dire que :

$$\int_T A_i^2\, dt = Q^2 \quad \int_T A_i\, A_j\, dt = \lambda Q^2, \; (i \neq j)\,;$$

le nombre λ est inférieur ou égal à 1 en module. Il doit d'ailleurs remplir une autre condition, car on peut écrire :

$$\int_T \left[\sum_1^m A_k(t)\right]^2 dt = \sum_{i,j}\sum \int_T A_i\, A_j\, dt = Q^2\,[m + (m^2 - m)\,\lambda] \geqslant 0$$

donc

$$\lambda \geqslant -1/(m-1).$$

On doit donc choisir λ de telle sorte que

$$-1/(m-1) \leqslant \lambda \leqslant 1.$$

Ceci posé, on peut calculer la probabilité de réception correcte fournie par le détecteur optimal.

Ce détecteur compare les quantités :

$$T\overline{[X(t) - A_1(t)]^2} \quad \text{et} \quad T\overline{[X(t) - A_i(t)]^2}$$
pour $i = 2 \ldots m$.

Si le signal $A_1(t)$ a été envoyé, le signal $X(t) = A_1(t) + n(t)$ a été reçu par le détecteur.

Il donnera une indication correcte si toutes les inégalités

$$T\overline{[X(t) - A_1(t)]^2} < T\overline{[X(t) - A_i(t)]^2},$$

sont satisfaites simultanément.

Ce système d'inégalités est équivalent au suivant :

$$2T\,\overline{n(t)\,[A_i(t) - A_1(t)]} < T\overline{[A_i(t) - A_1(t)]^2},$$

ou encore au suivant :

$$\sigma\sqrt{2T}\sqrt{\overline{A_i^2}(t)}\,\theta_i - \sigma\sqrt{2T}\sqrt{\overline{A_1^2}(t)}\,\theta_1 < T\overline{[A_i(t) - A_1(t)]^2},$$

* Ingénieur à la société ALCATEL

(*) The theory of maximum noise immunity. *Wiley, N.Y*, 1960.

où les quantités $\theta_1 \ldots \theta_n$ sont des variables aléatoires gaussiennes centrées mais corrélées, avec

$$E(\theta_i \theta_j) = \lambda, \quad (i \neq j).$$

En effet, les grandeurs $\overline{n(t)\, A_i(t)}$ sont des variables aléatoires gaussiennes centrées, dont on peut évaluer les corrélations :

$$E[n(t)\, A_i(t)\, n(t)\, A_j(t)]$$

$$= \frac{1}{T^2} \iint_T E[n(t)\, n(\tau)]\, A_i(t)\, A_j(\tau)\, dt\, d\tau,$$

et comme

$$E[n(t)\, n(\tau)] = (\sigma^2/2)\, \delta(t - \tau),$$

l'expression $\left(\frac{\sigma^2}{2T^2}\right) . Q^2 \lambda$ est le coefficient de corrélation de deux variables distinctes, dont la variance est $\left(\frac{\sigma^2}{2T^2}\right) Q^2$.

On peut alors écrire :

$$\overline{n(t)\, A_i(t)} = \frac{\sigma}{2T}.$$

Dans ces conditions, le système d'inégalités :

$$\theta_i - \theta_1 < (1 - \lambda) \sqrt{2}\, \frac{Q}{\sigma},$$

doit être vérifié pour qu'il y ait réception correcte du message $A_1(t)$.

Pour résoudre ce problème, remarquons qu'à partir du système de variables aléatoires $\theta_1 \ldots \theta_m$ on peut en définir un autre, soit : $a, \lambda_1 \ldots \lambda_m$, gaussiennes, centrées, indépendantes, telle que :

$$E(a^2) = \lambda, \quad \theta_1 = a + \lambda_1, \quad \theta_2 = a + \lambda_2 \ldots \theta_m = a + \lambda_m.$$

On est alors amené à considérer le système d'inégalités :

$$\lambda_i - \lambda_1 < (1 - \lambda) \sqrt{2}\, Q/\sigma \,;$$

par ailleurs, il est évi ent que $E(\lambda_i^2) = 1 - \lambda$.

On peut introduire un nouveau système de variables aléatoires de variance unité, si on pose :

$$\lambda_i = \sqrt{1 - \lambda}\, \mu_i,$$

et notre système d'inégalités se transforme finalement en :

$$\mu_i - \mu_1 < \sqrt{1 - \lambda}\, \sqrt{2}\, \frac{Q}{\sigma},$$

où les variables $\mu_1 \ldots \mu_m$ sont gaussiennes, centrées et de variance unité.

L'événement :

$$\mu_i < \sqrt{2}\, \frac{Q}{\sigma} \cdot \sqrt{1 - \lambda} + y,$$

est affecté d'une probabilité :

$$1 - V\left(\sqrt{2}\, \frac{Q}{\sigma} \cdot \sqrt{1 - \lambda} + y\right),$$

expression où :

$$V(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-x} e^{-z^2/2}\, dz.$$

La variable aléatoire μ_1 ayant le domaine de valeurs indiquées, toutes les inégalités du système sont satisfaites simultanément avec une probabilité :

$$\frac{dy}{\sqrt{2\pi}}\, e^{-y^2/2} \left[1 - V\left(\frac{\sqrt{2}\, Q}{\sigma} \sqrt{1 - \lambda} + y\right)\right]^{m-1}.$$

La probabilité de recevoir correctement le message $A_1(t)$, ou tout autre message de la collection $A_1 \ldots A_m$ est donc :

$$p(m) = \frac{1}{\sqrt{2\pi}} \times$$

$$\int_{-\infty}^{\infty} \left[1 - V\left(\sqrt{2}\, \frac{Q}{\sigma} \sqrt{1 - \lambda} + y\right)\right]^{m-1} e^{-y^2/2}\, dy.$$

On constate qu'en cas de disparition du message ($Q = 0$), la probabilité se réduit à $1/m$, ce qui est évident.

Par ailleurs, m signaux d'énergie Q^2 et de coefficient de corrélation λ se comportent exactement comme m signaux orthogonaux, d'énergie :

$$(1 - \lambda)\, Q^2.$$

Si les m signaux ont toujours même énergie et sont toujours équiprobables *a priori*, mais ont des distances mutuelles différentes, on peut poser $\int A_i\, A_j\, dt = Q^2 \lambda_{ij}$ et définir ainsi $m\,(m - 1)/2$ grandeurs λ_{ij}. Parmi celles-ci, il y en a une qui est plus grande que toutes les autres. Nous l'appellerons λ_μ et nous supposerons que les deux signaux qu'elle intéresse gardent les numéros 1 et 2.

Appelons $p_{12}, \ldots p_{1m}$, les probabilités de voir le détecteur décider en faveur des signaux respectifs 2, $\ldots$ m quand le signal 1 est envoyé. On en déduit facilement la probabilité p_1 de réception correcte du signal 1 :

$$p_1 = (1 - p_{12}) \ldots (1 - p_{1m}),$$

et cette probabilité est supérieure à $(1 - p_{12})^{m-1}$, qui serait la probabilité de réception correcte si tous les produits scalaires avaient pour valeur λ_m. La probabilité p_1 est donc supérieure à la probabilité qu'on aurait si les m signaux avaient pour produits scalaires deux à deux λ_μ ; il en est de même pour la probabilité de recevoir correctement n'importe quel signal de l'ensemble.

II. QUANTITÉ D'INFORMATION FOURNIE PAR UN DÉTECTEUR ET SON OPTIMALISATION.

II. 1. Calcul.

Un détecteur, de pouvoir séparateur m, fait une opération de choix parmi les m signaux qui forment l'ensemble considéré, avec une probabilité $p(m)$ de choix correct, calculée plus haut. Il apporte ainsi une quantité d'information facile à évaluer si les signaux jouent des rôles identiques. Il suffit de dresser le tableau des probabilités de présence des m signaux $A_1(t) \ldots A_m(t)$ avant et après l'opération.

Probabilités de :	$A_1(t)$	$A_2(t)$	$A_m(t)$
Avant	$1/m$	$1/m$	$1/m$
Après	p/m	$\dfrac{1-p(m)}{m-1}$	$\dfrac{1-p(m)}{m-1}$

La quantité d'information apportée est :

$$I = p(m) \log_2 [p(m)] + [1 - p(m)] \times \log_2 \left[\frac{1-p(m)}{m-1}\right] - \log_2 \left[\frac{1}{m}\right] \text{ bits,}$$

ou encore :

$$I = \log_2 m + p(m) \log_2 [p(m)] + [1 - p(m)] \log_2 \left[\frac{1-p(m)}{m-1}\right] \text{ bits.}$$

II. 2. Fréquence de récurrence optimale pour puissance moyenne donnée.

Les considérations précédentes nous permettent d'aborder le problème suivant :

— en présence d'une puissance moyenne donnée pour le signal, et d'une puissance de bruit par Hz donnée, ainsi que d'un détecteur de pouvoir séparateur m, à quelle cadence doit-on envoyer les messages, c'est-à-dire à quelle cadence doit opérer le détecteur, pour que l'information effective reçue par seconde soit la plus grande possible ?

Nous supposerons égale à l'unité la puissance de bruit par Hz, ce qui n'introduit aucune restriction ; nous appellerons W la puissance moyenne de signal et f la fréquence de récurrence des messages. L'énergie de chaque signal est alors W/f, la quantité d'information reçue par seconde est $I F$ bits/s C'est cette dernière grandeur que nous chercherons à maximaliser.

Il faut, pour cela, calculer I avec la formule du paragraphe précédent, où :

$$p(m) = \frac{1}{\sqrt{2\pi}} \times \int_{-\infty}^{\infty} \left[1 - V\left(\sqrt{\frac{2W}{f}}\sqrt{1-\lambda} + y\right)\right]^{m-1} e^{-y^2/2}\, dy.$$

Le problème ne peut se traiter que numériquement, l'intégrale $p(m)$ étant tabulée à l'aide d'une machine. On constate que la quantité d'information reçue par seconde est maximale pour une certaine valeur de f.

Par ailleurs, la grandeur $(I f)$ est de la forme $f\varphi m \left(\frac{W}{f}\right)$. Donc si m et λ sont donnés et si on multiplie la puissance moyenne par un facteur k, on trouve comme fréquence optimale la valeur kf et le taux maximal d'information est simplement multiplié par k. On peut remarquer que la fréquence optimale est obtenue en résolvant l'équation :

$$\varphi_m \left(\frac{W}{f}\right) = \left(\frac{W}{f}\right) \varphi'_m \left(\frac{W}{f}\right).$$

On peut tracer la courbe indiquant, pour m donné, la variation de $\pm f$ en fonction de f. Il suffit de la tracer pour un jeu particulier de valeurs de W et de λ. Si on veut la connaître pour un autre jeu, soit W_1 et λ_1, il suffit de multiplier l'abscisse et l'ordonnée par $\frac{W_1}{W} \cdot \frac{1-\lambda_1}{1-\lambda}$, donc de faire une simple homothétie.

La courbe tracée a donc, pour chaque valeur du pouvoir séparateur m, un caractère de courbe universelle.

Voici, par exemple, le cas de $m = 2^{10} = 1\,024$. Si on pose : $\sqrt{\frac{2W}{f}}(\sqrt{1-\lambda}) = U$, on peut dresser le tableau de $p(1024)$ en fonction de U :

U	$p(1024)$
0	1/1024
1	0,0163
2	0,1190
3	0,4092
4	0,7624
5	0,9505
6	0,9950
7	0,9998
8	0,999979
9	0,999992

et on peut dresser le tableau de If en fonction de f, pour $W(1 - \lambda) = 1$; par exemple :

f(Hz)	2	0,5	0,222	0,125	0,080	0,056
If(bits/s)	0,0887	0,3326	0,6926	0,7932	0,7317	0,5503

On obtient un taux d'information maximal pour une fréquence de récurrence de 0,15 Hz environ, le taux maximal étant 0,8 bit/s environ.

II. 3. Interprétation des résultats.

Physiquement, envoyer les messages à une cadence très lente signifie qu'on affecte beaucoup d'énergie à chaque demi-message. La réception est bonne, mais la quantité d'information envoyée est faible. Les envoyer à une cadence très rapide signifie mauvaise réception, mais beaucoup d'information envoyée. Il n'est pas étonnant qu'on trouve un optimum entre les deux pour l'information reçue.

On peut constater ainsi, qu'en présence d'un détecteur donné, on doit utiliser une certaine bande de fréquence utile, celle qui correspond à la transmission de la cadence f optimale et que tout accroissement de la bande ne présente aucun avantage.

D'autres problèmes se posent, non encore résolus. Par exemple, supposons qu'on ait à mesurer la distance d'un but par un écho sonar. On dispose d'une certaine puissance moyenne et d'un détecteur de précision donnée, c'est-à-dire qu'en fait, on a à choisir entre m signaux orthogonaux. Faut-il émettre toute la puissance dans un seul canal ou faut-il utiliser plusieurs canaux différents, avec des détecteurs différents, l'un assurant un dégrossissage et l'autre une détermination fine par exemple ?

III. RACCORDEMENT AVEC LE POINT DE VUE DE SHANNON : VALEUR ASYMPTOTIQUE DU TAUX D'ERREUR.

Supposons alors que nous avons un détecteur dont le pouvoir séparateur peut croître indéfiniment. Envisageons un message formé d'une suite de signaux à détermination binaire : $+ A(t)$, $- A(t)$. Un message quelconque est formé par une séquence $S(t)$ de la forme $+A, -A, -A, +A, +A, +A, \ldots$ etc. Si nous considérons une suite de n signaux, il y aura 2^n messages possibles, soit $S_1 \ldots S_2\, n(t)$. Supposons ensuite n très grand, et prenons l'un des messages, que nous appellerons $S_1(t)$. Les distances de ce message aux autres, sont :

$$\frac{\int S_1(t)\, S_i(t)\, dt}{\left(\int S_1^2(t)\, dt \int S_i^2(t)\, dt\right)^{1/2}} = \lambda_{1i},$$

Classons alors les messages $S_i(t)$ en deux catégories :

— la première, telle que $\lambda_{1i} > \alpha$ (α donné à l'avance avec $0 < \alpha < 1$) ;

— la seconde, telle que $\lambda_{1i} < \alpha$.

La première catégorie renferme un nombre $2^n\, \eta$, la seconde un nombre $2^n (1 - \eta)$ de messages.

Si α est donné, on peut trouver n assez grand pour que η soit inférieur à tout nombre positif ε si petit qu'on veut. Pour le voir, on peut supposer, sans restreindre les généralités, que $A(t) = 1$ sur un intervalle de temps fini pris comme unité et $A(t) = 0$ sur le reste de l'axe de temps, c'est-à-dire que :

$$\int_i S_i^2\, dt = n\ ;$$

On supposera aussi que S_1 est le message fourni par une répétition de n signaux $(+ 1)$. Donc :

$$\lambda_{1i} = \frac{\text{Nombre de } (+1) \text{ dans } S_i(t) - \text{Nombre de } (-1) \text{ dans } S_i(t)}{n},$$

$$= 1 - \frac{\text{Nombre de } (-1) \text{ dans } S_i(t)}{n/2}.$$

Si on veut que $\lambda_{1i} > \alpha$, il faut donc que le rapport $\dfrac{\text{Nombre de } (-1) \text{ dans } S_i(t)}{n/2}$ soit inférieur à $1 - \alpha$. Or, le nombre de messages qui répondent à cette condition est de la forme $2^n \dfrac{k}{\sqrt{n}}$ quand n est très grand. On peut donc bien, si α est donné, choisir n assez grand pour que $\dfrac{k}{\sqrt{n}} < \varepsilon$. Ceci est vrai quel que soit α ; on peut en particulier choisir $\alpha < \mu$, μ étant un nombre donné à l'avance si petit qu'on veut.

En résumé, on se donne deux nombres ε et μ. On prend n assez grand pour pouvoir classer les séquences $S_i(t)$ en deux catégories :

— la première, pour laquelle $\lambda_{1i} > \mu$, comprenant 2^n séquences ;

— la seconde, pour laquelle $\lambda_{1i} < \mu$, comprenant $2^n (1 - \varepsilon)$ séquences.

Pour chacune de ces deux catégories, nous majorerons la quantité λ_{1i}, augmentant ainsi les probabilités de confusion du message $S_1(t)$ avec les autres :

— pour la première catégorie, nous prendrons $\lambda_{1i} = 1$;

— pour la seconde catégorie, nous prendrons $\lambda_{1i} = \mu$.

Les messages de la première catégorie sont alors considérés comme indiscernables d'avec S_1. La probabilité de ne pas confondre S_1 avec un message de la seconde catégorie est, d'après ce que nous avons vu :

$$p[2^n (1 - \varepsilon)] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-v^2/2} \times$$

$$\left[1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-(\sqrt{2\pi}(Q/\sigma)\sqrt{1-\mu}+v)} e^{-v^2/2}\, dx\right]^{2^{(1-\varepsilon)}-1} dy.$$

Cette expression se présente sous la forme :

$$p[2^n (1 - \varepsilon)] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-v^2/2}\, \varphi(y)\, dy,$$

où le facteur $\varphi(y)$ est compris entre 0 et 1.

On peut alors séparer l'intégrale en deux et écrire :

$$p[2^n(1-\varepsilon)] = \frac{1}{\sqrt{2\pi}} \times \left[\int_{-\infty}^{-k} e^{-v^2/2}\,\varphi(y)\,dy + \int_{-k}^{+\infty} e^{-v^2/2}\,\varphi(y)\,dy\right],$$

et on peut choisir k assez grand pour que :

$$\int_{-\infty}^{-k} e^{-v^2/2}\,\varphi(y)\,dy,$$

soit inférieur à tout nombre donné à l'avance.

On peut écrire alors avec une approximation aussi bonne qu'on veut :

$$p[2^n(1-\varepsilon)] = \frac{1}{\sqrt{2\pi}} \int_{-k}^{+\infty} e^{-v^2/2}\,\varphi(y)\,dy.$$

Une transformation élémentaire permet d'écrire :

$$p[2^n(1-\varepsilon)] = \frac{1}{\sqrt{2\pi}} \int_{-k}^{\infty} e^{-v^2/2} \times \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+(\sqrt{2n}(Q/\sigma)\sqrt{1-\mu}+v)} e^{-v^2/2}\,dx\right]^{2^n(1-\varepsilon)-1} dy.$$

Comme $y > -k$, on peut considérer :

$$\log_e(\varphi(y)) = \log_e\left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(\sqrt{2n}(Q/\sigma)/\overline{1-\mu}+v)} e^{-v^2/2}\,dx\right]^{2^n(1-\varepsilon)-1} dy,$$

et le développement asymptotique de cette expression, valable pour n assez grand, soit :

$$[2^n(1-\varepsilon)-1] \times \log_e\left[1 - \frac{\exp -\frac{1}{2}(2n(Q/\sigma)\sqrt{1-\mu}+y)^2}{\sqrt{2\pi}(\sqrt{2n}(Q/\sigma)\sqrt{1-\mu}+y)^2}(1+\ldots)\right].$$

Si n grandit indéfiniment, le logarithme qui figure dans la dernière expression tend vers 0 et a pour partie principale :

$$-\frac{\exp\left[-\frac{1}{2}(\sqrt{2n}(Q/\sigma)\sqrt{1-\mu}+y)^2\right]}{\sqrt{2\pi}(\sqrt{2n}(Q/\sigma)\sqrt{1-\mu}+y)^2},$$

qu'on peut réduire à :

$$-\frac{\exp[-nQ^2\sqrt{1-\mu}/\sigma^2]}{2\sqrt{2\pi}nQ^2(1-\mu)/\sigma^2}.$$

La partie principale de $\log_e \varphi(y)$ est alors :

$$-[2^n(1-\varepsilon)-1]\,\frac{\exp[-nQ^2(1-\mu)/\sigma^2]}{2\sqrt{2\pi}nQ^2(1-\mu)/\sigma^2},$$

qu'on peut réduire à :

$$-(1-\varepsilon)\,\frac{\exp[n\log_e 2 - nQ^2(1-\mu)/\sigma^2]}{2\sqrt{2\pi}nQ^2(1-\mu)/\sigma^2}.$$

Quand n grandit indéfiniment, $\log_e \varphi(y)$ tend vers 0 pour toute valeur de $y > -k$ si, et seulement si $n \log_e 2 < n\,Q^2\,(1-\mu)/\sigma^2$ ou :

$$\frac{Q^2}{\sigma^2} > \frac{\log_e 2}{1-\mu};$$

$\log_e \varphi(y)$ tend alors uniformément vers zéro, c'est-à-dire indépendamment de la valeur de y, si $y > -k$;
alors, $p\,[2^n\,(1-\varepsilon)]$ tend vers :

$$\frac{1}{\sqrt{2\pi}} \int_{-k}^{+\infty} e^{-v^2/2}\,dy,$$

et comme k peut être choisi aussi grand qu'on veut, $p\,[2^n\,(1-\varepsilon)]$ tend vers 1.

Comme ε, dt, φ peuvent être choisis arbitrairement petits, on peut écrire que $p\,(2^n)$ tend vers 1 si, et seulement si : $Q^2/\sigma^2 > \log_e 2$.

Ce résultat est à rapprocher de celui de Shannon. Mais il ne fait pas intervenir explicitement la largeur de bande du canal, sur les limitations de laquelle on n'a fait aucune hypothèse.

Si on reprend la formule fondamentale de Shannon, qui fournit le débit limite d'information d'un canal où les conditions imposées sont : largeur de bande limitée à F Hz et puissance moyenne de signal donnée W,

$$C = F \log_2\left[1 + \frac{W}{\sigma_F^2}\right] \text{ bit/s};$$

on peut avoir une valeur asymptotique de cette capacité quand on élargit indéfiniment la bande :

$$C \sim \left(\frac{W}{\sigma^2}\right) \log_2 e.$$

Quand le message est constitué par une suite de signaux binaires orthogonaux d'énergie Q^2 et de fréquence de récurrence $1/\tau$, la puissance moyenne de signal est $W = Q^2/\tau$. Donc :

$$C = \left(\frac{Q^2}{\sigma^2\,\tau}\right) \log_2 e;$$

pour une transmission sans erreur, le débit d'information $1/\tau$ doit être au plus égal à la capacité du canal ; alors :

$$\frac{1}{\tau} \leqslant \frac{Q^2}{\sigma^2\,\tau} \log_2 e \quad \text{ou} \quad \frac{Q^2}{\sigma^2} \geqslant \log_e 2.$$

Enfin, si on considère une succession de signaux binaires de durée τ et d'énergie finie Q^2 :

a) on peut faire tendre le taux d'erreur vers 0, si $\dfrac{Q^2}{\sigma^2} > \log_e 2$: le débit d'information est alors de $1/\tau$ bit/s ;

b) si on considère les signaux comme des signaux à spectre borné et uniforme dans la bande $(0, 1/2\tau)$, ce débit limite donné par la formule de Shannon est :

$$\frac{1}{2\tau}\log_2\left(1+\frac{W}{N}\right)=\frac{1}{2\tau}\log_2\left(1+\frac{Q^2/\tau}{\sigma^2/2\tau}\right)$$

$$=\frac{2\tau}{1}\log_2(1+2\log_e 2),$$

soit :

$$\frac{1}{\tau}\left[\frac{1}{2}\log_2(1+2\log_e 2)\right]=\frac{1}{\tau}+0{,}65.$$

Ce résultat donne une idée de l'erreur commise en remplaçant les signaux réels par des signaux à spectre borné.

CONCLUSION

On a vu comment est limitée la capacité d'un canal d'information quand on lui connaît la puissance moyenne de signal, la puissance moyenne de bruit par Hz et le pouvoir séparateur du détecteur utilisé. On est conduit à utiliser une bande de fréquences optimale pour extraire du canal la quantité d'information maximale ; cette fréquence est, toutes choses égales d'ailleurs, proportionnelle à la puissance de signal. Ce point de vue se rapproche de celui de Shannon à condition de faire croître indéfiniment le pouvoir séparateur du détecteur et la largeur du canal.