# LA PROBABILITÉ DES CAUSES DANS LA TECHNIQUE DES TÉLÉCOMMUNICATIONS ET DES MESURES

par Julien LOEB
Ingénieur en Chef des Télécommunications

SOMMAIRE. — *Le problème des télécommunications peut se poser de la façon suivante : on a affaire à un ensemble de causes possibles (par exemple signaux émis) et à un certain nombre d'observations faites (par exemple signaux reçus). Il s'agit, une de ces observations étant acquise, d'en déterminer la cause la plus probable.*

*La « probabilité des causes » dépend non seulement de l'action aléatoire de la voie de communication ou de l'instrument de mesure, mais encore de la probabilité* a priori *de la cause.*

*On applique le théorème de* BAYES *au calcul de l'ambiguïté* $H_y(x)$ *de* SHANNON, *et à celui de l'adaptation d'une source à une « voie avec bruit ».*

## 1. INTRODUCTION

Depuis quelques années, on a fait de réels progrès dans l'interprétation de l'opération — mentale ou matérielle — qui permet de passer, par voie d'*inférence,* d'un signal reçu au signal effectivement transmis correspondant.

Notamment, dans la technique du radar, MM. P. M. WOODWARD [1] et D. MIDDLETON [2] ont montré la vraie position du problème : étant donné un aspect du « scope A », quelle est la probabilité pour que, dans tel intervalle de distance, il y ait un écho ?

Nous allons essayer d'appliquer ce moyen d'investigation au domaine des signaux discrets de télécommunications, tel que SHANNON l'étudie dans son mémoire fondamental [3] dont on est encore loin d'avoir tiré toutes les conséquences.

Les lois de la physique permettent, connaissant le symbole appliqué à l'entrée et les conditions de bruit de la ligne, de calculer les probabilités respectives des symboles reçus. Mais le vrai probleme des télécommunications est tout autre, ainsi que les auteurs précités l'ont montré. La réalité sur laquelle on travaille, c'est l'*observation* (le signal reçu, en télécommunications). Le travail de l'observateur consistera à attribuer à chacun des symboles émis une probabilité : le système sera bon si l'un des symboles possibles à l'émission se voit attribuer une probabilité nettement supérieure à celles des autres, parfait si cette probabilité *a posteriori* est égal à 1.

Nous négligerons, bien entendu, les erreurs systématiques. Si à l'observation $Ob_k$ correspond une distribution de probabilités *a posteriori* avantageant nettement l'événement $E_j$ nous appellerons cette observation $Ob_j$, et nous en serons quittes, si le tableau des correspondances est déjà convenu à l'extrémité réceptrice, à publier un code de corrections bi-univoque et exempt de tout caractère aléatoire.

C'est là, au fond, une limitation de la théorie de l'information : nous sommes capables, devant un texte imprimé — un article de journal par exemple — de calculer la « quantité d'information » que comporte le tirage au sort des symboles qui auraient pu le composer. Cela, répétons-le, ne nécessite qu'une connaissance préalable de la liste des mots possibles avec leurs probabilités *a priori.* Mais notre théorie ne nous rend pas capables de discerner si l'article du journal est tout à fait conforme à la vérité !

Le théorème central de la théorie des probabilités des causes est celui de BAYES. Bien qu'il s'agisse d'un théorème classique, nous allons en reprendre ici la démonstration, qui éclairera les définitions des diverses probabilités mises en jeu.

## 2. LE THÉORÈME DE BAYES

### 2.1. Le cas discret.

Soit une série d'événements $E_i$ dont les probabilités $P(E_i)$ sont connues *a priori* et constituent ainsi une donnée du problème. Ces événements, dans le problème qui nous occupe, ne sont pas directement observables mais peuvent conduire à un certain nombre d'observations $Ob_k$ sans qu'il y ait entre les événements $E_i$ et $Ob_k$ de relations strictement bi-univoques. On est ainsi amené à définir :

$P(Ob_k|E_i)$, la probabilité pour que, $E_i$ étant, $Ob_k$ soit celle des observations possibles qui a été réellement effectuée, et

$P(E_i|Ob_k)$, la probabilité pour que $Ob_k$ ayant été constatée, la cause en soit $E_i$. C'est là le problème bien connu : « Mon adversaire au jeu de l'écarté tire dix fois de suite un roi. Quelle est la probabilité pour que ce soit un tricheur ? »

Le théorème de BAYES établit une relation entre $P(E_i)$, $P(Ob_k|E_i)$ et $P(E_i|Ob_k)$. On l'obtient en calculant de deux façons différentes la probabilité $\Pi_{ik}$ pour que l'on ait *à la fois* $E_i$ et $Ob_k$.

1° $\Pi_{ik}$ est égale au produit de $P(E_i|Ob_k)$ par la probabilité globale de $Ob_k$. $Ob_k$ peut, nous le savons, être réalisée en partant de tous les événements $E_j$ et sa probabilité est :

$$P(Ob_k) = \sum_j P(E_j)\, P(Ob_k|E_j).$$

[ ] Pour tout renvoi *entre crochets,* se reporter à la bibliographie *in fine.*

Donc

$$\Pi_{ik} = P(E_i|Ob_k) \sum_j P(E_j)\, P(Ob_k|E_j).$$

2° $$\Pi_{ik} = P(E_i) . P(Ob_k|E_i).$$

D'où le théorème :

$$(1) \qquad P(E_i|Ob_k) = \frac{P(E_i)\, P(Ob_k|E_i)}{\sum_j P(E_j)\, P(Ob_k|E_j)}.$$

Dans la formule (1) les deux catégories de quantités figurant au second membre sont connues : celles du premier membre sont les inconnues qu'on a ainsi réussi à calculer. On peut d'ailleurs remarquer que la formule (1) ne permet guère de résoudre le problème du tricheur, car elle suppose connues les probabilités *a priori* $P(F_i)$. Il faudrait ainsi connaître, *avant de commencer à jouer*, combien de chances on a de se trouver en présence d'un tricheur.

Dans les applications techniques que nous allons envisager, nous aurons toujours à faire une hypothèse, parfois arbitraire, sur les probabilités *a priori* des causes possibles $E_i$. Nous chercherons cependant à voir comment varie le résultat en fonction de $P(E_i)$.

## 2.2. Le cas continu.

Soit $X$ la variable aléatoire et $Y$ l'observation. Soient :

$P(x)dx$ la probabilité (*a priori*) pour que la variable aléatoire $X$ soit entre $x$ et $x + dx$,

$P_x(y)dy$ la probabilité pour que $X$ étant donné et égal à $x$, $Y$ soit compris entre $y$ et $y + dy$,

$Q_x(x)dx$ la probabilité des causes cherchées : $Y$ étant donné et égal à $y$, quelle est la probabilité pour que $X$ soit entre $x$ et $x + dx$ ?

$P(y)dy$ la probabilité que $Y$ soit entre $y$ et $y + dy$ (en tenant compte de toutes les possibilités de $X$).

Utilisant la même méthode que ci-dessus, nous exprimons de deux facons la probabilité pour que, à la fois

$$x < X < x + dx,$$
$$y < Y < y + dy.$$

Nous obtenons :

$$P(y)\, dy\, Q_y(x)\, dx = P(x)\, dx\, P_x(y)\, dy.$$

Or

$$P(y)\, dy = \int P(x)\, dx\, P_x(y)\, dy$$

(prise dans tout le domaine de variation de $x$).

D'où la nouvelle formule de BAYES.

$$(2) \qquad Q_y(x) = \frac{P(x)\, P_x(y)}{\int P(x)\, P_x(y)\, dx},$$

# 3. LES CAS DE SIMPLIFICATION

On est, *a priori*, tenté de confondre $P(E_i|Ob_k)$ et $P(Ob_k|E_i)$. En général, ce n'est pas permis, mais il y a des cas particuliers très importants où cette identification devient licite.

## 3.1. Cas discret.

Examinons d'abord les relations globales auxquelles satisfont les probabilités mises en jeu.

On a d'abord :

$$(3) \qquad \sum_i P(E_i) = 1.$$

Puis, l'équation

$$(4) \qquad \sum_k P(Ob_k|E_j) = 1$$

indique que la cause $E_j$ donne forcément lieu à l'une des observations $Ob_k$ (ces observations sont exclusives l'une de l'autre).

Inversement, il n'est pas obligatoire que l'observation $Ob_k$ ait pour cause l'un des événements $E_i$ qui s'excluent. Cependant il sera intéressant d'étudier un cas assez fréquent où nous aurons affaire à ce que nous appellerons un *groupe causal.*

Un nombre fini de causes $F_i$ est associé à un nombre fini d'observations $Ob_k$, de telle sorte que

$$(5) \qquad \sum_j P(Ob_k|E_j) = 1.$$

### *3.1.1. Les deux conditions de réversibilité.*

L'équation (1) montre que, si (5) est vraie, et *si de plus les probabilités $P(E_i)$ sont égales,*

$$(6) \qquad P(E_i) = P(E_j) = \ldots = 1/M$$

($M$ étant le nombre total des $E_i$),

on obtient la loi de réversibilité :

$$(7) \qquad P(Ob_k|E_j) = P(E_j|Ob_k).$$

Il faut bien remarquer que l'équation (5) est une hypothèse supplémentaire, et ne provient pas du théorème des probabilités complémentaires. Ce n'est qu'*a posteriori* que l'application de l'équation (7) permettrait d'écrire :

$$\sum_j P(E_j|Ob_k) = 1,$$

qui, intuitivement, signifie que l'une des causes de $Ob_k$ se situe obligatoirement parmi les $E_j$.

Il y a un autre cas où cette réversibilité existe, quelles que soient les $P(E_i)$. C'est celui où à chaque $E_i$ ne correspond qu'une seule observation possible, que nous notons $Ob_i$. Dans l'équation (1), en effet, toutes les probabilités $P(Ob_k|E_i)$ sont nulles, sauf $P(Ob_i|E_i)$ qui est égale à 1. On a alors :

$$(8) \qquad P(E_i|Ob_k) = P(Ob_k|E_i) = \delta_{ik}.$$

Nous avons ainsi défini les cas où la réversibilité est légitime. On la voit parfois invoquer dans les textes sans qu'elle soit accompagnée des justifications nécessaires.

Lorsque la réversibilité n'a pas lieu, on constate une sorte d'« attraction » vers les valeurs les plus probables des $P(E_i)$. Prenons par exemple la lecture du voltage du secteur 110 V. Même si, par suite d'une accumulation presque impossible d'erreurs, on lit 80 V sur le voltmètre, c'est tout de même une tension voisine de 110 V qui sera la cause la plus probable au sens de BAYES.

#### 3.1.2. Exemples d'application des résultats ci-dessus.

On peut citer un cas idéal dans lequel les deux hypothèses (6) et (7) ci-dessus sont vérifiées.

Reprenons l'exemple de codage, donné par SHANNON [3], d'un « langage » composé de 4 symboles A, B, C, D de probabilités inégales, et codé en binaire de telle sorte que les nouveaux symboles 0 et 1 soient équiprobables.

Supposons que ces nouveaux symboles soient transmis sur une ligne où l'on ne peut discerner que deux niveaux :

0 sera le niveau d'un signe indiscernable du bruit,

1 sera le premier niveau discernable du bruit.

On a bien affaire à un groupe causal, car un 0 ou un 1 reçus ne peuvent avoir d'autres causes qu'un 0 ou un 1 transmis.

La pratique ordinaire du télétype donne un exemple qui s'écarte déjà du modèle simplifié ci-dessus. A chaque symbole émis correspond bien l'un des symboles reçus, et réciproquement, mais ici les $P(E_i)$ ne sont plus égales.

### 3.2. Le cas continu.

Le schéma du cas continu sera le suivant :

Une grandeur à mesurer $X$ peut avoir toutes les valeurs comprises entre 0 et 1. Pour chacune de ces valeurs, la quantité observée $Y$ a une loi de répartition :

(9) $P_x(y)\,dy$ = probabilité pour que,

$X$ étant donné et égal à $x$, $Y$ soit entre $y$ et $y + dy$.

On a en général

(10) $$\int P_x(y)\,dy = 1.$$

Nous allons supposer maintenant, comme précédemment, que la répartition des $x$ est équiprobable :

$$P(x) = 1.$$

L'équation 2 devient :

$$Q_y(x) = \frac{P_x(y)}{\int P_x(y)\,dx}.$$

L'intégrale du dénominateur ne pourra être prise égale à 1 que si on en fait expressément l'hypothèse. Là encore, il y a par définition, un « groupe causal ». On en déduit la loi de réciprocité.

(11) $$Q_y(x) = P_x(y).$$

L'équation (11) est vérifiée en particulier si $P_x(y)$ garde la même valeur si on remplace $x$ par $y$. Prenons par exemple le cas où la voie introduit une erreur à répartition gaussienne :

$$P_x(y) = (1/\sigma\sqrt{2\pi})\,e^{-(y-x)^2/2\sigma^2} = Q_y(x)\,;$$

c'est là le type d'un « genre causal ».

#### 3.2.1. Exemple de cas continu.

$x$ est un voltage $v$ à mesurer, supposé ici constant, $y$ est la déviation du voltmètre.

C'est seulement si $v$ a une répartition équiprobable que l'on peut intervertir la probabilité directe, donnée par l'application des lois de la physique, et la probabilité des causes.

En particulier, si, comme dans l'article de l'auteur [4], le voltage inconnu a une autre loi de répartition, elle-même gaussienne dans le cas traité, la valeur du voltage $v$ à déduire d'une lecture $y$ est différente de celle qu'on aurait inférée dans le cas d'une répartition équiprobable des $x$.

## 4. APPLICATION AU CALCUL DE LA CAPACITÉ D'UNE VOIE SELON SHANNON (CAS DISCRET)

Si (voir [3]) l'ambiguïté $H_x(y)$ se calcule directement en fonction des probabilités $P(\mathrm{Ob}_k|\mathrm{E}_i)$, en fait, on utilise surtout l'ambiguïté $H_y(x)$ construite, elle, à partir des probabilités des causes $P(\mathrm{E}_i|\mathrm{Ob}_k)$. En effet, c'est cette dernière qu'il faut retrancher de $H(x)$, l'autre donnée du problème, pour avoir la capacité d'une voie avec bruit.

Dans son mémoire [3], C. E. SHANNON définit comme capacité « le maximum de $H(x) - H_y(x)$ pour toutes les répartitions en probabilités des $\mathrm{E}_i$ côté $x$ ». On sait bien que $H(x)$ est maximum pour une répartition équiprobable mais, jusqu'ici, SHANNON n'a pas indiqué comment $H_y(x)$ varie en fonction des probabilités $P(\mathrm{E}_i)$. C'est ce que nous allons chercher à calculer maintenant.

### 4.1. Cas de deux symboles (voir fig. 1).

Pour commencer, nous allons examiner le cas de deux événements $\mathrm{E}_1$ et $\mathrm{E}_2$ pouvant chacun produire

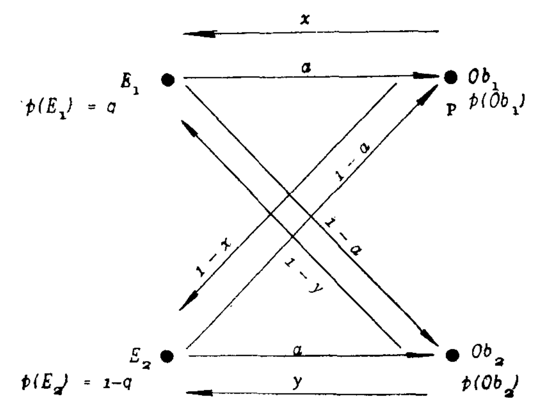


FIG. 1. — Tableau des probabilités relatives à deux événements et à deux observations.

l'une de deux observations $\mathrm{Ob}_1$ et $\mathrm{Ob}_2$. Les données du problème sont :

(12) $$\begin{cases} a = P(\mathrm{Ob}_1|\mathrm{E}_1) = P(\mathrm{Ob}_2|\mathrm{E}_2) \\ 1 - a = P(\mathrm{Ob}_2|\mathrm{E}_1) = P(\mathrm{Ob}_1|\mathrm{E}_2). \end{cases}$$

#### 4.1.1. Calcul des probabilités des causes.

Nous devons d'abord déterminer les probabilités des causes :

$$x = P(\mathrm{E}_1|\mathrm{Ob}_1),$$
$$y = P(\mathrm{E}_2|\mathrm{Ob}_2),$$
$$z = P(\mathrm{E}_2|\mathrm{Ob}_1),$$
$$t = P(\mathrm{E}_1|\mathrm{Ob}_2).$$

Le paramètre variable est $q$ :

$$(13) \quad \begin{cases} q = P(\mathrm{E}_1), \\ 1 - q = P(\mathrm{E}_2). \end{cases}$$

Comme nous supposons tout de suite qu'il s'agit d'un groupe causal, conforme au schéma de la figure 2, nous tirons de (1) les équations (14) et (15) :

$$(14) \quad x = \frac{qa}{qa + (1-q)(1-a)}.$$

$$(14') \quad P(\mathrm{Ob}_1) = qa + (1-q)(1-a),$$

$$(15) \quad y = \frac{(1-q)q}{(1-q)a + q(1-a)},$$

$$(15') \quad P(\mathrm{Ob}_2) = q(1-a) + (1-q)a,$$

$$(16) \quad z = 1 - x,$$

$$(17) \quad t = 1 - y.$$

Nos inconnues $x$, $y$, $z$, $t$ dépendent :

— des lois de la physique : probabilité $a$ de voir la cause $\mathrm{E}_1$ produire l'observation $\mathrm{Ob}_1$,

— de la répartition en probabilité *a priori* des causes $\mathrm{E}_1$ et $\mathrm{E}_2$, cette répartition étant souvent subjective.

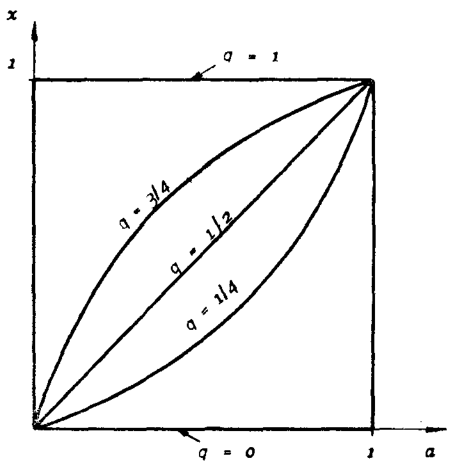


FIG. 2. — Variation de $x = P(\mathrm{E}_1|\mathrm{Ob}_1)$ en fonction de $a = P(\mathrm{Ob}_1|\mathrm{E}_1)$ pour diverses valeurs de $q = p(\mathrm{E}_1)$.

La figure 2 donne $x$ en fonction de $a$ pour les diverses valeurs de $q$.

Les courbes à $q$ constant sont des hyperboles équilatères passant par O et par B ($a = 1$, $x = 1$). Elles ne dégénèrent en la droite OB que pour la valeur remarquable $q = 1/2$.

On a alors, *mais alors seulement* :

$$x = y = a,$$
$$z = t = 1 - a,$$
$$P(\mathrm{Ob}_1) = P(\mathrm{Ob}_2) = 1/2.$$

### *4.1.2. Calcul de l'ambiguïté selon Shannon.*

L'ambiguïté $H_y(x)$ sert, en définitive à mesurer, statistiquement, le nombre des suites composées des symboles émis qui restent discernables malgré le bruit dont la voie est le siège. Si $N$ est le nombre des tirages au sort, on a

$$2^{N[H(x) - H_y(x)]}$$

suites discernables parmi les $2^{NH(x)}$ suites possibles à l'émission.

Pour calculer $H_y(x)$, refaisons avec SHANNON le raisonnement dans le cas de deux symboles. Prenons d'abord une suite reçue qui ne comporte que des observations $\mathrm{Ob}_1$.

Le nombre des suites du côté émission qui lui correspondent est

$$2^{-N[P(\mathrm{E}_1|\mathrm{Ob}_1)\log P(\mathrm{E}_1|\mathrm{Ob}_1) + P(\mathrm{E}_2|\mathrm{Ob}_1)\log P(\mathrm{E}_2|\mathrm{Ob}_1)]};$$

cette expression doit être construite au moyen des *probabilités des causes*. Le reste du calcul se poursuivra de la manière indiquée par SHANNON. On pourrait, il est vrai, construire $H_x(y)$ avec les probabilités directes, mais cela ne ferait que déplacer le problème, car il faudrait alors introduire $H(y)$ qui se construit au moyen des $P(\mathrm{E}_i)$ et des $P(\mathrm{Ob}_k|\mathrm{E}_i)$.

Reprenons, en le ramenant au cas de 2 symboles, l'exemple du § 15, p. 415, [3] où les quantités $p$ et $q$ correspondent, suivant nos notations, à $a$ et $1 - a$, et où $P$ et $Q$ doivent être remplacées par $q$ et $1 - q$ (ici, il y en a une de moins). SHANNON a supposé implicitement que les probabilités de transition du groupe des causes $\mathrm{E}_i$ à celui des observations $\mathrm{Ob}_k$ pouvaient être renversées.

Nous venons de voir que ce n'est vrai que dans la mesure où $q = 1/2$ (équiprobabilité *a priori* des causes).

Or il y a des problèmes — et celui, capital, de l'adaptation en est un — où on doit calculer les variations des entropies avec $q$. La réversibilité, admise implicitement par SHANNON, ne nous paraît pas devoir être conservée dans les calculs qui vont suivre.

En introduisant ainsi les probabilités des causes, on obtient pour $H_y(x)$ l'expression :

$$(18) \quad H_y(x) = -\{P(\mathrm{Ob}_1)[x \log x + (1-x)\log(1-x)] + P(\mathrm{Ob}_2)[y \log y + (1-y)\log(1-y)]\}.$$

Pour $q = 1/2$, on trouve :

$$H_y(x) = -[a \log a + (1-a)\log(1-a)],$$

à ce moment $H(x) = 1$.

La capacité de la voie est, par unité de bande passante :

$$(19) \quad C = H(x) - H_y(x).$$

Nous vérifions immédiatement que si, à son tour, $a = 1/2$, ce qui donne à $H_y(x)$ la valeur maximum égale à 1, la capacité ($q = 1/2$) de la voie devient nulle.

Il s'agit maintenant de résoudre le problème suivant :

La quantité $a$ étant donnée, comment va varier $C$ avec $q$ ?

Nous savons que, en l'absence de bruit, la capacité passe par un maximum pour $q = 1/2$.

Est-ce encore vrai pour la quantité $C$ de l'équation (19) ?

Cherchons à tracer, pour une valeur donnée de $a$, la courbe des variations de $H_y(x)$ en fonction de $q$.

Nous pourrons d'ailleurs, sans restreindre la généralité, supposer $a > 1/2$. Cette courbe est symétrique par rapport à l'axe $q = 1/2$. Pour $q = 0$, on a

$$x = 0, \quad y = 1, \quad z = 1, \quad t = 0,$$
$$P(\mathrm{Ob}_1) = 1 - a, \quad P(\mathrm{Ob}_2) = a,$$

d'où

$$H_y(x) = 0.$$

Comme la courbe est symétrique par rapport à l'axe $q = 1/2$, et comme il n'y a pas de point anguleux pour cette valeur, $H_y(x)$ passe par une valeur stationnaire en M (fig. 3).

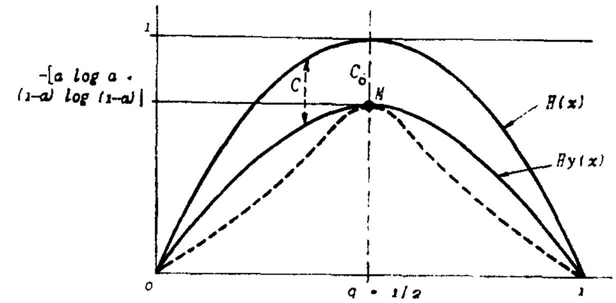


Fig. 3. — Variation de $H(x)$ et de $H_y(x)$ en fonction de $q = p(\mathrm{E}_1)$ pour une valeur de $a = P(\mathrm{Ob}_1|\mathrm{E}_1)$.

La courbe marquée $H_y(x)$ sur la figure 3 a d'ailleurs été tracée arbitrairement.

Nous n'avons pas démontré, en effet, que $C$ passe effectivement par un maximum pour $q = 1/2$. Tout ce que nous savons au point actuel de l'étude, c'est que :

1° La courbe $H_y(x)$ ne passe pas au-dessus de la courbe $H(x)$. En effet la multiplicité des suites appartenant au domaine des signaux émis et causes possibles des suites reçues, ne saurait en aucun cas déborder la multiplicité de toutes les suites de ce domaine.

2° Les diverses courbes $H_y(x)$ correspondant aux valeurs croissantes de $a$ se placent l'une au-dessus de l'autre, car l'ambiguïté ne peut qu'augmenter lorsque $a$ approche de $1/2$.

3° $H_y(x)$ est une fonction croissante de $q$ « pour $0 < q < 1/2$ », car l'ambiguïté ne peut que croître lorsque, à l'incertitude sur l'effet de la voie (probabilité $a$, se joint l'incertitude sur les $\mathrm{E}_i$ (probabilité $q$).

Si l'affirmation jusqu'ici intuitive « $C$ passe par un maximum effectif lorsque $q = 1/2$ » reçoit un jour une confirmation mathématiquement établie, on pourra donner une réponse complète à la question suivante :

Dans le § 13 de [3], on définit la capacité $C$ comme étant « le maximum de l'expression $H(x) - H_y(x)$, pour toutes les sources qui pourraient être appliquées à l'entrée de la voie ».

Quelle est effectivement la répartition en probabilités qui rend $C$ maximum ?

Jusqu'ici nous pouvons juste dire que pour la répartition équiprobable, $C$ est stationnaire.

Mais rien de ce que nous savons de façon certaine n'empêche l'existence de courbes telles que la courbe en pointillé de la figure 3, selon lesquelles $C_0$ ne serait pas la valeur maximum de $H(x) - H_y(x)$.

### 4.2. Cas général.

Un même raisonnement donne pour $- H_y(x)$ la valeur

$$(20) \qquad -H_y(x) = \sum_i P(\mathrm{Ob}_i) \sum_k P(\mathrm{E}_k|\mathrm{Ob}_i) \log P(\mathrm{E}_k|\mathrm{Ob}_i),$$

avec

$$P(\mathrm{Ob}_i) = \sum_j P(\mathrm{Ob}_i|\mathrm{E}_j)\, P(\mathrm{E}_j),$$
$$P(\mathrm{E}_k|\mathrm{Ob}_i) = \frac{P(\mathrm{E}_k)\, P(\mathrm{Ob}_i|\mathrm{E}_k)}{P(\mathrm{Ob}_i)}.$$

Donc

$$-H_y)(x = \sum_{ik} P(\mathrm{E}_k)\, P(\mathrm{Ob}_i|\mathrm{E}_k) \log \frac{P(\mathrm{E}_k)\, P(\mathrm{Ob}_i|\mathrm{E}_k)}{P(\mathrm{Ob}_i)}.$$

Cette expression est invariante devant le groupe des permutations entre les $\mathrm{E}_k$. Cherchons les valeurs des $P(\mathrm{E}_k)$ qui rendent stationnaire $- H_y(x)$. Ces $P(\mathrm{E}_k)$ sont soumises à la liaison :

$$\sum_k P(\mathrm{E}_k) = 1$$

La méthode classique du calcul des variations indique les valeurs de $P(\mathrm{E}_k)$ vérifiant la condition de stationnarité : ce sont celles pour lesquelles

$$(21) \qquad \frac{\partial H_y(x)}{\partial P(\mathrm{E}_1)} = \frac{\partial H_y(x)}{\partial P(\mathrm{E}_2)} = \ldots\ldots = \frac{\partial H_y(x)}{\partial P(\mathrm{E}_M)}.$$

Un « point » dans l'expace à $M$ dimensions des $P(\mathrm{E}_i)$ satisfaisant à (20) est justement celui où une « surface » $H_y(x) = \mathrm{C}^{\mathrm{te}}$ vient à être tangente à l'hyperplan $\Sigma P(\mathrm{E}_i) = 1$, car (21) exprime que la « normale » à la « surface » a, comme la « droite » D d'équation $P(\mathrm{E}_i) = P(\mathrm{E}_2) = \ldots = P(\mathrm{E}_M)$, tous ses cosinus directeurs égaux.

Comme $H_y(x)$ ne change pas par la permutation des $P(\mathrm{E}_i)$, ce point est forcément sur la « droite » D.

Cette valeur stationnaire est-elle un vrai maximum ? Est-elle la seule ?

Des recherches effectuées dans ce sens permettraient de consolider le travail fondamental [3] de Shannon, en confirmant (ou en infirmant) l'opinion intuitive selon laquelle la répartition équiprobable, qui donne un maximum pour $H(x)$ donne aussi un maximum pour $H(x) - H_y(x)$.

*Manuscrit reçu le* 8 *octobre* 1953.

BIBLIOGRAPHIE

[1] Woodward (P. M.). Information theory and the design of radar receivers. (Utilité de la théorie de l'information pour l'établissement des plans de récepteurs de radar.) *Proc. Inst. Radio Engrs*, U. S. A. (déc. 1951), **59**, n° 12, pp. 1521-1524.

[2] Middleton (D.). Statistical criteria for the detection of pulsed carriers in noise. (Critères statistiques pour la détection de courants porteurs impulsionnels dans le bruit.) *J. appl. Phys.*, U. S. A. (avril 1953), **24**, n° 4, pp. 371-378.

[3] Shannon (C. E.). A mathematical theory of communication. (Une théorie mathématique des télécommunications.) *Bell Syst. techn. J.*, U. S. A. (juil. 1948), **27**, n° 3, pp. 379-423.

[4] Loeb (J. M.). Une théorie « informationnelle » de la mesure et de la télémesure. *Ann. Télécommunic.*, Fr. (avril 1951), **6**, n° 4, pp. 90-92.