Michèle Dehen

Application des mathématiques polonaises et des capacités à l'étude de l'information sans probabilité

Séminaire Choquet. Initiation à l'analyse, tome 9, n° 2 (1969-1970), exp. n° 23, p. 1-11
<http://www.numdam.org/item?id=SC_1969-1970__9_2_A13_0>

Séminaire CHOQUET
(Initiation à l'Analyse)
9e année, 1969/70, n° 23, 11 p.

4 juin 1970

APPLICATION DES MATHÉMATIQUES POLONAISES ET DES CAPACITÉS À L'ÉTUDE DE L'INFORMATION SANS PROBABILITÉ

par Michèle DEHEN

avec la collaboration d'Yves DERMENJIAN et Jean SAINT-RAYMOND

L'étude suivante se place dans un cadre analogue à celui de la théorie de l'information axiomatisée par Bruno FORTE et Nicolo PINTACUDA ([4] et [9]). J'expose les axiomes utilisés, et j'étudie des conditions permettant de prolonger la fonction d'information lorsque celle-ci est donnée sur une classe d'expériences.

1. Définitions et axiomes.

1.1. Information sans probabilité associée aux évènements. - On considère un ensemble Ω, qui sera appelé évènement certain. On note E un sous-ensemble de $\mathcal{P}(\Omega)$, stable par réunion et intersection finie, qui sera appelé ensemble des évènements.

Si F est un sous-ensemble de $\mathcal{P}(\Omega)$, on note $\sigma(F)$ la tribu engendrée par F ; si $(G_i)_{i\in I}$ est une famille de tribus dans $\mathcal{P}(\Omega)$, elle sera dite algébriquement indépendante (resp. σ-algébriquement indépendante) si, pour tout sous-ensemble J, fini (resp. dénombrable), contenu dans I, pour toute famille $(a_j)_{j\in J}$, où $a_j \in G_j$ et $a_j \neq \emptyset$ pour tout j, on a $\bigcap_{j\in J} a_j \neq \emptyset$.

Une famille K d'évènements est appelée algébriquement indépendante, si la famille de tribus $(\sigma(\{a\}))_{a\in K}$ est algébriquement indépendante. D'après un théorème de Banach-Marczewski, l'indépendance algébrique est équivalente à l'universelle indépendance stochastique [8]. L'indépendance de K signifie qu'aucun des évènements de K n'entraîne une conjonction d'autres évènements, ni le complémentaire de celle-ci.

Une information sur E sera une application de E dans les réels permettant, en termes intuitifs, de mesurer l'apport de caractérisation des points (dans le cas discret) dû à la connaissance de l'appartenance de ceux-ci à un évènement. Elle sera adaptée à une classe K, si la connaissance de l'appartenance à $a \cap b$ apporte une caractérisation égale à la somme de celles obtenues pour a et pour b (appartenant à K), séparément. Cette définition sera précisée dans la suite ; le choix de K est souvent suggéré empiriquement, et on peut voir, dans les cas suivants, que la famille K n'est, souvent, pas canonique. Si Ω est un ensemble discret composé de 2^n points, l'information associée à un de ces points sera n, nombre de symboles binaires nécessaires pour le caractériser. Lorsque les points sont équiprobables, on peut remarquer que

$$j(a) = n = - \log_2 p(a) .$$

Exemple, exposé par R. S. INGARDEN et K. URBANIK [7] : si p et q sont deux probabilités sur Ω ,

$$j(a) = \int_a \log \frac{q(dx)}{p(dx)} p(dx) ,$$

lorsque cette formule a un sens.

Exemple de l'entropie (= information) en mécanique statistique : On note P et Q les points de variétés de dimension s $(s \in \underset{\sim}{N})$ définissant respectivement le moment et la position du système mécanique (à " s degrés de liberté"), h est la constante de Planck, ρ la densité d'une mesure par rapport à la mesure de Lebesgue (dans le cas classique, ρ est constante et égale à h^{-1}), et on pose :

$$j(a) = - \int_a \rho(P , Q) \log [h^s \rho(P , Q)] \, dP \, dQ .$$

Un autre exemple physique peut être obtenu en optique en remplaçant, dans cette formule, h par la longueur d'onde λ .

On dira qu'une application j est une <u>mesure d'information</u> sur E , adaptée à la famille algébriquement indépendante K , si :

1° $j : E \to \underset{\sim}{\bar{R}}_+$;

2° $a \in E$, $b \in E$, $a \subset b \Longrightarrow j(a) \geq j(b)$;

3° $a \in K$, $b \in K \Longrightarrow j(a \cap b) = j(a) + j(b)$;

4° On dira que j est $\top$-compositive, où $\top$ est une application de

$$\Gamma = \{(j(a) , j(b)) : (a , b) \in E \times E , \ a \cap b = \emptyset\} \text{ dans } \underset{\sim}{\bar{R}}_+ ,$$

si

$$(a , b) \in E \times E , \ a \cap b = \emptyset \Longrightarrow j(a \cup b) = j(a) \top j(b) .$$

Le quadruplet (Ω , E , K , j) sera appelé <u>espace informationnel</u>.

Si E est une tribu, une partie n de Ω sera dite <u>j-négligeable</u>, s'il existe un $b \in E$ tel que $n \subset b$, $j(b) = j(\emptyset)$. N. PINTACUDA a démontré le théorème de complétion suivant [9] :

<u>Si</u> E <u>est une tribu, et si</u> j <u>est compositive et vérifie, pour toute suite</u> $(a_n)_{n \in \underset{\sim}{N}}$ <u>d'évènements disjoints,</u>

$$j(\cup a_n) = \lim_{p \to \infty} j(a_1) \top j(a_2) \top \dots \top j(a_p) ,$$

<u>alors la classe</u> $M = \{a \cup n ; \ a \in E , \ n \ \text{j-négligeable}\}$ <u>est une tribu</u> ; M <u>est égale à</u> E <u>si, et seulement si,</u> E <u>est complète</u> ; j <u>se prolonge de manière</u>

unique en une mesure d'information j_c sur M , et M est complète pour j_c .

Dans le même exposé, l'auteur indique comme probable l'existence de résultats plus forts, constitués par des prolongements obtenus en construisant des capacités de Choquet. Dans la suite, cette étude sera faite dans le cadre plus général des expériences, en procédant de deux manières :

- on ne construira pas de capacités, mais on appliquera directement les techniques de G. CHOQUET ;
- on utilisera les procédés de rabotage de W. SIERPINSKI.

1.2. Information associée aux expériences. - On conserve les notations précédentes ; au début, on étudiera le cas où $\mathcal{P}(\Omega)$ est l'ensemble des évènements, et, lorsqu'on fera des opérations sur les expériences, on étudiera si celles-ci peuvent être définies lorsque l'ensemble des évènements est une sous-tribu E de $\mathcal{P}(\Omega)$.

Une expérience $a = \{x_i\}_{i\in I}$ est un ensemble d'évènements non vides et disjoints deux à deux. On note X l'ensemble des expériences. Si $a \in X$, $a = \{x_i\}_{i\in I}$, on pose $u(a) = \bigcup_{i\in I} x_i$. On munit X de la relation d'ordre partielle :

$a \leqslant b$ (a moins fine que b) si, et seulement si, $\forall\, y \in b$, $\exists\, x \in a$, $y \subset x$.

On étudiera ensuite des axiomes permettant de définir l'information sur les expériences. Lorsque a est une information formée d'un nombre fini d'évènements disjoints, on appellera information de a , un réel permettant, intuitivement, de mesurer l'apport de caractérisation des points (dans le cas discret) dû à la classification associée à l'expérience.

On dira qu'une classe L d'expériences est algébriquement indépendante lorsque tout évènement de l'une des expériences de L est indépendant de toute conjonction d'évènements des autres expériences de L .

On imposera que l'information associée aux expériences soit additive sur L , c'est-à-dire, de manière heuristique, que l'apport de caractérisation des points associé à la classification de l'expérience borne supérieure de a et $b \in L$, dont les évènements sont les conjonctions de deux évènements appartenant respectivement aux expériences a et b , soit la somme des apports dus à a et b .

Ces notions sont axiomatisées dans la suite.

Si Ω est un espace muni d'une probabilité p , on peut définir une information sur l'ensemble des expériences finies $a = (x_1 , x_2 , \ldots , x_n)$, telles que $\bigcup_{i=1}^{n} x_i = \Omega$, par la formule de SHANNON

$$h(a) = - [\sum_{k=1}^{n} p(x_i) \log_2 p(x_i)] .$$

Au paragraphe 4, on verra un exemple d'information à laquelle on ne peut associer de probabilité.

Si $(a_i)_{i\in I}$ est un ensemble d'expériences, $a_i = \{x_j^i\}_{j\in J_j}$, on définit :

$$\bigvee_{i\in I} a_i = \{\bigcap_{i\in I} a_{j_i}^j\}_{(j_i)\in\prod_i J_i} \setminus \{\emptyset\} .$$

Il est immédiat que $\bigvee_{i\in I} a_i$ est la borne supérieure, dans X , des a_i ; on voit de plus que, si I est dénombrable, et si les évènements des a_i appartiennent tous à une même tribu E , il en est de même pour les évènements de $\bigvee_{i\in I} a_i$.

Sur les éléments de $\bigcup_{i\in I} a_i \subset \mathcal{P}(\Omega)$, on définit la relation d'équivalence

$x \sim y \iff \exists\, x_1 , x_2 , \dots , x_n \in \bigcup a_i : x \cap x_1 \neq \emptyset \dots x_p \cap x_{p+1} \neq \emptyset \dots x_n \cap y \neq \emptyset .$

Si $x \in \bigcup a_i$, on note

$$\dot{x} = \bigcup \{y , \ y \sim x , \ y \in \bigcup a_i\}$$

et

$$\bigwedge_{i\in I} a_i = \{\dot{x} : \ x \in \bigcup a_i\} .$$

On voit que $\bigwedge_{i\in I} a_i$ est la borne inférieure, dans X , des a_i . Si I est dénombrable, ainsi que chaque a_i , si les évènements des a_i appartiennent tous à une tribu E , les évènements de la borne inférieure appartiennent encore à E .

Si $u(a_i) = U$ est indépendante de l'indice $i \in I$, on vérifie que

$$u(\bigvee a_i) = u(\bigwedge a_i) = U .$$

Le sous-ensemble X_c des expériences complètes (i. e. $X_c = \{a ; \ u(a) = \Omega\}$) est donc un sous-treillis de X .

On note Y l'ensemble des expériences constituées d'un seul évènement au plus (et on les identifie à l'évènement qui les compose) ; pour l'ordre $\leqslant$, Y est un treillis, et

$$\bigvee a_i = \{\bigcap a_i\} \setminus \{\emptyset\} , \quad \bigwedge a_i = \{\bigcup a_i\} .$$

La borne supérieure dans Y coïncide avec celle dans X , mais non la borne inférieure. De plus, le treillis Y est distributif.

Un sous-ensemble L de X sera dit algèbriquement indépendant, si les tribus engendrées par les éléments de deux expériences quelconques de L forment un ensemble algébriquement indépendant. Soient S un sur-ensemble de l'ensemble d'indépendance L , et A un sous-ensemble de X dont on précisera les propriétés en cours d'étude.

On appellera information sur (Ω, S, A) adaptée à L, une application :

1° $h : S \to \underset{\sim}{\bar{R}}_+$;

2° h croissante pour la finesse, c'est-à-dire $(a \leqslant b) \implies (h(a) \leqslant h(b))$;

3° h additive sur L, $((x, y) \in L \times L) \implies (h(x \vee y) = h(x) + h(y))$;

4° h continue à droite, relativement à A,

$$\forall \varepsilon > 0, \forall x \in S, \exists y \in A, y \geqslant x : z \in S, x \leqslant z \leqslant y \implies h(z) - h(x) < \varepsilon .$$

On dira que (Ω, A, L, h) est un espace expérimental. La trace de cet espace sur l'ensemble des expériences constituées d'un seul évènement constitue un espace informationnel.

2. Information intérieure et extérieure ; expérience informative.

On peut employer une terminologie analogue à celle définie par G. CHOQUET pour les capacités ; si $a = \{x_i\}_{i \in I}$ est une des expériences de X, on définit :

- son information intérieure : $h_*(a) = \sup \{h(s), a \geqslant s \in S\}$;
- son information extérieure : $h^*(a) = \inf \{h(b), a \leqslant b \in A\}$;

en suivant la convention $\inf \emptyset = +\infty$, $\sup \emptyset = 0$.

On dira que a est une expérience informative si $h_*(a) = h^*(a)$. Il est immédiat que les éléments de A sont informatifs.

PROPOSITION 2.1. - *Soit (Ω, A, S, h) un espace expérimental, où S soit absorbant, c'est-à-dire*

$$(x_1, x_2 \in S, y \in A, x_1 \vee x_2 \leqslant y \implies \exists x \in S : x_1 \vee x_2 \leqslant x \leqslant y) ;$$

alors toute expérience de S est informative.

Démonstration. - Si $a \in S$, si $h_*(a) = \infty$, on voit que $h^*(a) = h_*(a)$. Si $h_*(a) < \infty$, d'après le 4°,

$$\forall \varepsilon > 0, \exists y \in A, y \geqslant a : \forall x \in S, a \leqslant x \leqslant y \implies h(x) - h(a) < \varepsilon .$$

On voit alors que

$$\forall s \in S, s \leqslant y \implies \exists u \in S : s \vee x \leqslant u \leqslant y ;$$

par conséquent $h(s) \leqslant h(u) \leqslant h(a) + \varepsilon$, et donc $h(y) \leqslant h(a) + \varepsilon$, ce qui entraîne $h^*(a) = h(a)$. On peut remarquer que si S est stable par $\vee$ fini, la condition d'absorption est vérifiée. Dans la suite, on supposera que S est toujours formé d'expériences informatives.

Définition. - Une classe M, contenue dans X, sera dite additive (resp. multiplicative, resp. σ-additive, resp. σ-multiplicative) si elle est stable par borne

...

supérieure finie (resp. inférieure finie, resp. dénombrable).

PROPOSITION 2.2. - Soient S et A deux ensembles additifs. On suppose que :

1° $x \in S$, $(y_1 , y_2) \in A \times A$, $x \leqslant y_1 \vee y_2$

$$\Longrightarrow \exists (x_1 , x_2) \in S \times S , \quad x_i \leqslant y_i \quad (i = 1 , 2) , \quad x \leqslant x_1 \vee x_2 \quad ;$$

2° $\forall \varepsilon > 0$, $\exists \eta > 0$; $((a_1 , a_2 , b_1 , b_2) \in S^4$, $a_i \leqslant b_i$, $h(b_i) - h(a_i) < \eta)$

$$\Longrightarrow h(b_1 \vee b_2) - h(a_1 \vee a_2) < \varepsilon \ ;$$

alors toute borne supérieure finie d'expériences informatives l'est aussi.

Démonstration. - On montre d'abord que, si $(c_1 , c_2) \in A \times A$, si $(a_1 , a_2) \in S \times S$, et $h(c_i) - h(a_i) < \eta$, alors $h(c_1 \vee c_2) - h(a_1 \vee a_2) \leqslant \varepsilon$. En effet,

$$h(c_1 \vee c_2) = \sup \{h(d) : d \leqslant c_1 \vee c_2 , \ d \in S\} = \sup \{h(d_1 \vee d_2) , \ c_i \geqslant d_i \in S\} ,$$

$$h(c_1 \vee c_2) = \sup \{h(d_1 \vee d_2) \ldots (< h(a_1 \vee a_2) + \varepsilon)\} \leqslant h(a_1 \vee a_2) + \varepsilon .$$

Il suffit ensuite de montrer que, si a_1 et a_2 sont deux expériences informatives, $a_1 \vee a_2$ l'est aussi. Si $h(a_1)$ ou $h(a_2) = + \infty$, alors

$$h_*(a_1 \vee a_2) = + \infty = h^*(a_1 \vee a_2) \quad \text{si} \quad \sup [h(a_1) , h(a_2)] < \infty ,$$

$\forall \varepsilon > 0$, $\exists \eta > 0$, $\exists (b_1 , b_2) \in A.A$, $\exists (c_1 , c_2) \in S.S$,

$$c_i \leqslant a_i \leqslant b_i , \quad h(b_i) - h(c_i) < \eta ,$$

alors $b_1 \vee b_2 \geqslant a_1 \vee a_2 \geqslant c_1 \vee c_2 \in S$ et $h(b_1 \vee b_2) - h(c_1 \vee c_2) < \varepsilon$. Par conséquent, $a_1 \vee a_2$ est informative.

PROPOSITION 2.3. - Si (Ω , A , S , h) est un espace expérimental vérifiant les hypothèses 2.2, si A est σ-additive, si :

1° $(c_i , d_i)_{i \leqslant n} \in A^{2n}$, $\forall i$, $c_i \leqslant d_i \Longrightarrow h(\vee d_i) - h(\vee c_i) \leqslant \sum_{i \leqslant n} h(d_i) - h(c_i)$,

2° $\forall (a_n)_{n \in \underset{\sim}{N}}$ suite croissante, dans A , $\Longrightarrow h(\vee a_n) = \lim h(a_n)$;

alors, pour toute suite (b_n) croissante, dans X , $h(\vee b_n) = \sup h^*(b_n)$, et toute borne supérieure d'une suite d'expériences informatives l'est aussi.

Démonstration. - Soit (b_n) une suite croissante contenue dans X ; s'il existe n_0 tel que $h^*(b_{n_0}) = + \infty$, on voit que $h^*(\vee b_n) = \lim h^*(b_n)$; sinon,

$$\forall \varepsilon > 0 , \ \forall n \in \underset{\sim}{N} , \ \exists a_n \in A , \ a_n \geqslant b_n , \ h(a_n) - h^*(b_n) < \varepsilon . 2^{-n} ;$$

par suite,

$$\forall p \in \underset{\sim}{N} , \ h(\bigvee_{n \leqslant p} a_n) - h^*(\bigvee_{n \leqslant p} b_n) \leqslant \varepsilon .$$

Comme A est σ-additive,

$$a = \bigvee_{p \in \underset{\sim}{N}} \left(\bigvee_{n \leq p} a_n \right) \in A$$

et

$$h\left(\bigvee_{p \in \underset{\sim}{N}} b_p \right) \leqslant h(a) = \lim h\left(\bigvee_{n \leq p} a_n \right) \leqslant \lim h^*(b_p) + \varepsilon .$$

Si $(c_n)_{n \in \underset{\sim}{N}}$ est une suite d'expériences informatives, on peut, d'après 2.2, se ramener à une suite croissante, et conclure que la borne supérieure des (c_n) est informative.

Remarque 2.4. - Lorsque S et A sont contenues dans l'ensemble Y des expériences constituées d'un seul évènement, on obtient des résultats analogues en faisant l'étude dans le treillis Y, c'est-à-dire en considérant Y indépendamment de son plongement dans X.

Il existe des cas où la condition imposée dans 2.3 au prolongement de h à A est une conséquence de propriétés intrinsèques de S et A ; par exemple, si, pour toute suite croissante d'expériences (a_n) de A et toute $b \in S$ telle que $b \leqslant \bigvee a_n$, il existe un entier n_0 tel que $b \leqslant a_{n_0}$.

Les propositions 2.2 et 2.3 sont des transpositions du théorème 28.1 de [1], démontré par G. CHOQUET pour l'étude des capacités sur un ensemble.

3. Capacitances, rabotages, expériences polies (ou lisses).

On pourrait faire une étude directe de l'informativité d'expériences obtenues par bornes supérieures ou inférieures, dénombrables à partir d'une classe d'expériences informatives ; cependant, on appliquera plutôt l'étude systématisée des treillis, faite en [2].

Pour tout cardinal α, l'ensemble X est un α-treillis (c'est-à-dire, toute partie $z \subset X$, de cardinal $\leqslant \alpha$, admet une borne supérieure $\bigvee z$ et une borne inférieure $\bigwedge z$). On rappelle qu'un pavage E est un sous-ensemble de X, additif et multiplicatif. Si $(a_n)_{n \in \underset{\sim}{N}}$ est une suite décroissante d'éléments de X, on dit que a est une E-enveloppe des $(a_n)_{n \in \underset{\sim}{N}}$, s'il existe une suite $(b_n) \subset E$, telle que $\{a_n\}$ soit cofinale dans $\{a_n , b_n\}$ pour $\geqslant$ et $a \geqslant \bigwedge b_n$. Une classe C est une famille semi-extérieure (ou capacitance) si $a \in X$, si (b_n) est une suite croissante contenue dans X,

$$a \wedge (\bigvee b_n) \in C \implies \exists\, n_0 : a \wedge b_{n_0} \in C .$$

On suppose que, dans tout ce paragraphe, pour toute suite (b_n) croissante dans X et tout $a \in X$:

$$h^*((\bigvee b_n) \wedge a) = \lim_{n \to \infty} h^*(b_n \wedge a) .$$

On peut voir alors que, pour tout réel $\lambda \in \underset{\sim}{R}$, $C_\lambda = \{a \in X : h^*(a) > \lambda\}$ est une capacitance.

L'étude de la stabilité des éléments polis par l'opération de Souslin faite en [2] a permis de démontrer le théorème suivant :

THÉORÈME. - Si a est le noyau d'un système déterminant

$$a = \bigvee_{\sigma \in \underset{\sim}{N}^{\underset{\sim}{N}}} \Big(\bigwedge_{\substack{s \in S \\ s \prec \sigma}} a_s \Big) ,$$

où les a_s soient polis et vérifient $r \trianglelefteq s \Longrightarrow a_r \leq a_s$, alors a est un élément poli.

On rappelle que $r \trianglelefteq s$ signifie que les longueurs de r et s sont égales et que chaque terme de r est inférieur ou égal au terme de même rang de s .

PROPOSITION 3.1. - Si (Ω , A , S , h) est un espace expérimental et si E est un pavage de X , formé d'expériences informatives, tel que, pour toute suite décroissante $(e_n) \subset E$, $h_*(\bigwedge e_n) = \lim h(e_n)$, alors tout élément poli de X est informatif.

Démonstration. - Soient a un élément poli de X , et F un rabotage global compatible avec a . Si $\varepsilon \in \underset{\sim}{R}_+^*$, on note C_ε la capacitance :

$$C_\varepsilon = \{b \in X : h^*(b) > h^*(a) - \varepsilon\} .$$

Soit $(x_n)_{n \in \underset{\sim}{N}}$ la suite dérivée de a . Comme a en est une enveloppe, il existe une suite (a_n) , contenue dans E , décroissante, vérifiant

$$\forall n , \quad x_n \leq e_n \quad \text{et} \quad \wedge e_n \leq a ;$$

par conséquent,

$$h_*(a) \geq h^*(a) - \varepsilon .$$

L'élément poli a est donc informatif.

PROPOSITION 3.2. - Si S (resp. A) est additif et multiplicatif dans X , si :

(a) Pour toute suite décroissante $(a_n)_{n \in \underset{\sim}{N}} \subset S$ (resp. A),

$$h_*(\bigwedge a_n) = \lim h(a_n) .$$

(b) Pour toute suite croissante $(b_n) \subset X$, et tout $a \in X$,

$$h^*[(\bigvee b_n) \wedge a] = \lim h^*(b_n \wedge a) ,$$

alors toute expérience appartenant au stabilisé de S (resp. A), par borne supérieure de suite croissante et par borne inférieure dénombrable, est informative.

PROPOSITION 3.3. - Si S (resp. A) est additif et multiplicatif dans Y , si (a) est vérifiée et si, pour toute suite croissante $(b_n) \subset Y$, $h^*(\bigvee b_n) = \lim h^*(b_n)$,

<u>alors toute expérience du stabilisé de</u> S (resp. A), <u>par borne supérieure de suite croissante et par borne inférieure, dans Y, dénombrable, est informative.</u>

Les démonstrations de 3.2 et 3.3 se déduisent de la stabilité des éléments polis par opération de Souslin, et de 3.1.

4. L'information shannonienne ; une généralisation sans probabilité.

Dans trois notes aux Comptes Rendus [6], B. FORTE et J. KAMPÉ de FÉRIET étudient l'information, portée par les évènements (boréliens) d'un groupe abélien, invariante par translation. Dans ce paragraphe, on étudiera l'information portée par des expériences complètes sur un groupe abélien.

Soit Ω un groupe abélien localement compact non compact (il n'existe pas alors de probabilité de Haar) ; on ne peut pas munir l'ensemble Z des expériences complètes constituées d'un nombre fini d'évènements (boréliens ou universellement mesurables) d'une information de Wiener-Shannon. Soit ρ une mesure de Haar sur Ω ; si θ est une application de $\bar{\underset{\sim}{R}}_+$ dans lui-même, décroissante, et vérifiant $\theta(0) = +\infty$, $\theta(+\infty) = 0$, la fonction définie sur la tribu E des ensembles universellement mesurables par $j(a) = \theta \circ \mu(a)$ est alors une mesure d'information sur les évènements, invariante par translation. Lorsque θ est continue, il est immédiat que tout ensemble μ-mesurable est j-informatif.

Par analogie avec l'information shannonienne, on cherche s'il existe une information h, définie sur Z, telle que, si $a = \{x_1, x_2, \dots, x_n\}$, $h(a)$ soit une fonction de $(\mu(x_1) \dots \mu(x_n))$ de la forme

$$h(a) = \sum_{i=1}^{n} \mu(x_i)\, \rho[\mu(x_i)] ,$$

où ρ soit une application continue de $\bar{\underset{\sim}{R}}_+$ dans lui-même. Pour que h soit croissante, il suffit alors que, si x et y sont deux évènements disjoints,

$$\mu(x \vee y) \times \rho \circ \mu(x \vee y) \leq \mu(x) \times \rho \circ \mu(x) + \mu(y) \times \rho \circ \mu(y) .$$

En particulier, on peut choisir $\rho = \theta$, on obtient l'information sur Z :

$$h(\{x_1, x_2, \dots, x_n\}) = \sum_{i=1,2,\dots,n} \mu(x_i)\, j(x_i) ,$$

avec les conventions $\infty.0 = 0$ (pour que $h(\{\Omega\}) = 0$), $\sum_{\emptyset} = \infty$ (pour que $h(\emptyset) = \infty$, car $\emptyset = \vee z$), $0.\infty = \infty$ (car, si x est un évènement de mesure nulle,

$$h(\{x, Cx\}) = \infty) .$$

On peut alors appliquer l'étude du paragraphe 3 dans le treillis X_c des expériences complètes. Soit S la classe des expériences de X_c, constituées d'un nombre fini d'évènements, tous ouverts, sauf un au plus, ce dernier étant un fermé de mesure infinie. Si $s = \{x_1 , x_2 , \dots , x_n\} \in S$ (x_i ouvert, y fermé), on pose

$$h(s) = \sum \rho(x_i)\, j(x_i) \quad \text{(avec la convention } \infty.0 = 0 \text{)} .$$

On voit que S est additif et $\aleph_0$-multiplicatif. On définit ensuite la classe A des expériences de X_c , constituées d'un nombre quelconque de compacts ; A est multiplicative et $\aleph_0$-additive.

Si $(b_n)_{n\in\underset{\sim}{N}}$ est une suite croissante d'expériences de A , si $a \in S$, si $\forall b_n \geqslant a$, on peut montrer qu'il existe un entier n_0 pour lequel $h(b_{n_0}) \geqslant h(a)$ dès que la fonction $(t \longmapsto t\theta(t))$, de $\underset{\sim}{\bar{R}}_+$ dans lui-même, est décroissante ; on déduit alors de 2.3 que la borne supérieure d'une suite d'expériences informatives l'est aussi. En particulier, une expérience, constituée d'un ensemble dénombrable d'égènements, qui soient des G_δ ou des fermés, est informative.

Si la forme explicite de θ permet de conclure que les hypothèses 3.2 sont vérifiées, on voit en particulier que, si $a = \{x_n\}_{n\in\underset{\sim}{N}}$ est une expérience complète formée de boréliens, alors a est informative ; si $h(a) < \infty$, $\forall \varepsilon > 0$, il existe $b \in S$,

$$b = \{z_k = \bigcup_{n\in\underset{\sim}{N}} x_n^k \ , \ \ (x_n^k) \subset (x_n)\}_{k\leqslant p} \ ,$$

les z_k étant tous ouverts sauf un au plus, et il existe une expérience $c \in A$, $c \geqslant a$, telle que $h(b) - h(c) < \varepsilon$; si $h(a) = +\infty$, $\forall n \in \underset{\sim}{N}$, $\exists b \in S$, $b \leqslant a$, $h(b) > n$.

BIBLIOGRAPHIE

[1] CHOQUET (Gustave). - Theory of capacities, Ann. Inst. Fourier, Grenoble, t. 5, 1953/54, p. 131-295.

[2] DEHEN (Michèle), DERMENJIAN (Yves) et SAINT-RAYMOND (Jean). - Rabotages sur un treillis, Séminaire CHOQUET : Initiation à l'analyse, 9e année, 1969/70, nº 23.

[3] DELLACHERIE (C.). - Capacitances et rabotages, Exposé fait au Séminaire de probabilités, III, Strasbourg, 1967/68 (non publié).

[4] FORTE (B.) et PINTACUDA (N.). - Information fournie par une expérience, C. R. Acad. Sc. Paris, t. 266, 1968, Série A, p. 242-245.

[5] FORTE (B.) e PINTACUDA (N.). - Sull'informazione associata alle esperienze incomplete, Annali di Mat. pura ed appl., 4e série, t. 80, 1968, p. 215-234.

[6] KAMPÉ de FÉRIET (J.) et FORTE (B.). - Information et probabilité, C. R. Acad. Sc. Paris, t. 265, 1967, Série A, p. 110-114, 142-146, 350-353.

[7] INGARDEN (R. S.) and URBANIK (K.). - Information without probability, Coll. Math., Warszawa, t. 9, 1962, p. 131-150.

[8] KAPPOS (D. A.). - Strukturtheorie der Wahrscheinlichkeitsfelder und -räume. - Berlin, Springer-Verlag, 1960 (Ergebnisse der Mathematik. Neue Folge, 24).

[9] PINTACUDA (Nicolo). - Quelques développements récents de la théorie de l'information (à paraître).

[10] PINTACUDA (Nicolo). - Prolongement des mesures d'information (à paraître).

[11] RENYI (A.). - On measures of entropy and information, Proceedings of the 4th Berkeley symposium on mathematical statistics and probability [1960, Berkeley], vol. 1, p. 547-561. - Berkeley, University of California Press, 1961.

[12] SIERPINSKI (W.). - Sur la puissance des ensembles mesurables (B) ; Fund. math., Warszewa, t. 5, 1924, p. 166-171.

[13] ZERNER (Martin). - Quelques propriétés des espace nucléaires, d'après un article de Mitiagin, Séminaire Leray : Equations aux dérivées partielles, 1963/64, fascicule 2, p. 1-24.

(Texte remis le 1er octobre 1970)

Michèle DEHEN
Ass. Fac. Sc. Paris
11 bis rue du Lion
93 - BONDY