T H E S E

présentée

A LA FACULTE DES SCIENCES

DE L'UNIVERSITE DE GRENOBLE

pour obtenir

LE GRADE DE DOCTEUR ES-SCIENCES PHYSIQUES

par

Daniel THOULOUZE

Deuxième THESE

L'ENTROPIE EN THEORIE DE L'INFORMATION

Soutenue le 8 Mars 1968 devant la Commission d'Examen

MM. L. NEEL Président

B. DREYFUS
J. BOK Examinateurs
A. LACAZE

TABLE DES MATIERES

I N T R O D U C T I O N

Qu'est-ce que l'information ?

Le mot est suffisamment vague pour être interprété de façons tellement différentes qu'à certain colloque international historiens, sociologues, biologistes et physiciens entre autres n'ont pu trouver de terrain d'antente. Et encore, n'y avait-il pas de représentants de toutes les branches scientifiques concernées par ce sujet.

On peut poser que l'information consiste en une collection de données résultant de l'étude de la nature, de symboles représentant des classes d'êtres naturels ou le comportement de ces êtres.

Pour faire une théorie scientifique de l'information, il faut considérer le problème tel qu'il se pose au chercheur, problème qui comporte un certain nombre de réponses possibles lorsqu'on ne possède pas d'informations particulières sur la situation présente. Si l'on parvient à obtenir quelques informations sur le problème, le nombre de réponses possibles se trouve diminué : une information totale peut même conduire à une seule réponse possible. L'information est une fonction du rapport des réponses possibles après et avant qu'on l'ait reçue. On a ainsi une mesure de l'information absolue.

La description physique et mathématique a conduit à découvrir un lien entre l'information et l'entropie. Puisque l'information est une mesure de l'ordre et le concept d'entropie une mesure du désordre, il semble logique qu'il y ait une relation entre les deux. L'entropie d'un système physique serait considérée comme une mesure de l'incertitude dans laquelle on se trouve sur la structure de ce dernier. Il est cependant difficile de dire s'il y a identité ou analogie de forme et de fond, ce qui rend la distinction difficile.

CHAPITRE I

DEFINITION DE L'INFORMATION

1°) DEFINITION :

La définition de l'information dérive de données statistiques. Considérons par exemple (1) une expérience X dont le résultat est un message valable. Supposons que l'expérience ait N résultats également probables à priori. L'information contenue dans le message sur l'expérience X sera une fonction f(N). Supposons que X soit une expérience combinée consistant en deux expériences indépendantes Y et Z qui ont N_1 et N_2 résultats également probables. Le nombre total de résultats de l'expérience combinée est $N_1 \; N_2$. La transmission du résultat de X est équivalente à la transmission des résultats de Y et de Z séparément. Par conséquent, l'information sur X doit être la somme des informations sur Y et Z, c'est-à-dire :

$$f(N) = f(N_1) + f(N_2)$$

$$\text{avec} \quad N = N_1 \; N_2$$

f(N) doit être de plus une fonction croissante de N. Seule f(N) = C log N satisfait à ces conditions.

L'information peut donc être définie par la relation :

$$I = C \log N$$

De façon générale (2) on peut dire que E étant un évènement quelconque se produisant avec la probabilité P(E), si nous savons que l'évènement E est arrivé, nous avons reçu une quantité d'information

$$I = C \, Log \frac{1}{P(E)} = -C \, Log \, P(E)$$

2°) UNITES :

Le choix de l'unité revient à définir C. En prenant $C = \log_2 e$, c'est-à-dire en se plaçant en logarithme à base 2, l'unité d'information est le bit.

Si P(E) = 1/2, c'est-à-dire si N = 2, I(E) = 1 bit.

Le bit est donc la quantité d'information obtenue quand l'un de deux états également possibles se réalise.

Pour une comparaison ultérieure avec la thermodynamique, on peut définir I en unités d'entropie et poser par analogie $C \simeq k \simeq 10^{-16}$ cgs.

Pour l'instant, nous ne spécifierons plus C ; le logarithme sera pris de façon formelle, dans une base arbitraire.

3°) GENERALISATION :

- S'il existe No possibilités dans l'état initial et N_1 dans l'état final, l'information reçue s'écrit :

$$I = Log \frac{N_0}{N_1} = Log \, N_0 - Log \, N_1 = I_0 - I_1$$

- Si maintenant nous considérons l'exemple d'un nombre de G chiffres dans la base N ou d'un message de G lettres dans un alphabet en contenant N, la quantité d'information contenue dans ce nombre ou ce message sera :

$$I = Log\, N^{G} = G\, Log\, N$$

quelle que soit la base ou l'alphabet.

4°) INFORMATION CONTENUE DANS UN ENSEMBLE DE SYMBOLES AYANT DIFFERENTES PROBABILITES A PRIORI - SOURCE SANS MEMOIRE

Supposons que nous utilisions une séquence de symboles choisis dans un alphabet $S = \{ s_1, s_2, \ldots\ldots s \}$; les symboles successifs sont choisis selon une loi de probabilité fixée. Nous allons encore supposer que les symboles émis sont statistiquement indépendants. La source qui émettrait les séquences ci-dessus définies, dite "sans mémoire" (zero memory source) serait complètement décrite par l'alphabet S et les probabilités avec lequelles le symbole se produit, $P\ (s_1, s_2 \ldots\ldots s\ )$. Si le symbole est émis, nous obtenons une quantité d'information : $I(si) = -\log P\ (si)$. Ceci se produisant avec la probabilité P(si), la <u>quantité moyenne d'information obtenue par symbole</u> sera :

$$I = \sum_i P(s_i)\ I(s_i) = -\sum_i P(s_i)\ Log\, P(s_i)$$

Cette expression ressemble beaucoup à la quantité appelée entropie en mécanique statistique. C'est pour cette raison que SHANNON (3), qui l'a formulée le premier, l'a appelée "entropie" de la source. C'est une mesure de la diversité,

de la variété des messages produits par la source, une mesure de ce que BONSACK (4) appelle la "variabilité" potentielle d'un certain type de message.

Cette entropie H(S) peut être interprétée, soit comme la quantité moyenne d'information par symbole fourni par la source, soit comme mesure de l'incertitude de l'observateur connaissant les pj et essayant de prévoir le prochain évènement.

Il est possible de donner une justification mathématique de ce point de vue, (FEINSTEIN (5)). Appelons Ø cette incertitude à laquelle on impose les trois conditions :

1) Ø est une fonction continue des pj seulement ;
2) L'ordre dans lequel sont énumérés les évènements Aj est indifférent ;
3) Deux façons différentes d'envisager la situation pour évaluer l'incertitude doivent donner le même résultat.

Ceci suffit à déterminer pour Ø la forme :

$$\phi = K \sum_j p_j \operatorname{Log} p_j$$

5°) PROPRIETES FONDAMENTALES DE L'ENTROPIE

L'entropie H est maximale lorsque les n probabilités sont égales.
Si $p_1 \ldots\ldots p_n$ et $u_1 \ldots\ldots u_n$ sont deux distributions de probabilités des x_i, c'est-à-dire si

$$\sum_{i=1}^{n} p_i = \sum_{i=1}^{n} u_i = 1$$

alors, en utilisant le fait que $\operatorname{Log} x \geqslant 1 - \frac{1}{x}$, on trouve que :

$$\sum_{i=1}^{n} p_i \operatorname{Log} \frac{p_i}{u_i} \geqslant \sum_{i=1}^{n} p_i \left(1 - \frac{u_i}{p_i}\right) = 0$$

soit $$\sum_{i=1}^{n} p_i \operatorname{Log} p_i \geqslant \sum_{i=1}^{n} p_i \operatorname{Log} u_i$$

avec égalité si et seulement si $p_i = u_i$

6°) <u>SOURCE D'INFORMATION DE MARKOV</u>

Les spécialistes de la transmission de l'information trouvent trop restrictive la source sans mémoire et introduisent une source dans laquelle l'occurence d'un symbole s_i dépend d'un nombre fini m de symboles précédents. Une telle source appelée source de MARKOV du m^{ieme} ordre est déterminée par la donnée de l'alphabet et de l'ensemble des probabilités conditionnelles

$$P\left(s_i / s_{j_1} \cdots s_{j_m}\right)$$

$$i = 1, \ldots q$$
$$j = 1, \ldots q$$

La quantité d'information par symbole, pendant que l'on est dans l'état $(sj_1, \ldots\ldots sj_m)$ est :

$$H\left(S / s_{j_1} \cdots s_{j_m}\right) = -\sum_{S} P\left(s_i / s_{j_1} \cdots s_{j_m}\right) \operatorname{Log} P\left(s_i / s_{j_1} \cdots s_{j_m}\right)$$

Si nous faisons la moyenne de cette quantité sur les q^m états possibles, nous obtenons l'entropie de la source de MARKOV du $n^{ième}$ ordre

$$H(S) = \sum_{S^m} P\left(s_{j_1} \cdots s_{j_m}\right) H\left(S / s_{j_1} \cdots s_{j_m}\right)$$

Source adjointe : Considérons une source de MARKOV du 1er ordre S. La source adjointe à S, soit $\overline{S}$, est la source d'information sans mémoire ayant le même alphabet et les mêmes probabilités de symboles.

On démontre que $H(S) \leq H(\overline{S})$

Nous voyons ici un fait important : la connaissance des m symboles précédents impose à s_j une contrainte qui a pour effet de diminuer la quantité d'information dans le choix de s_j. L'égalité ne se rencontre que dans le cas de symboles indépendants.

Rôle des liaisons : D'une façon très générale, les liaisons dimininuent l'entropie. Nous avons déjà vu que l'entropie est maximale si les signes sont équiprobables. Elle est également maximale s'ils sont indépendants. Ceci est intuitivement compréhensible : les liaisons sont des entraves, elles diminuent le nombre de degrés de liberté. Il faut remarquer qu'à propos des liaisons l'entropie n'est diminuée qu'en moyenne. Si on calcule la quantité d'information transmise par un message selon la formule $I = -\log p$, il peut arriver qu'un message particulier transmette une plus grande information.

CHAPITRE II

LES PRINCIPES DE CARNOT

1°) EN THERMODYNAMIQUE :

Rappelons le principe de Carnot de la Thermodynamique.

L'entropie totale d'un système isolé fermé ne peut que rester constante (si ce système est le siège de phénomènes reversibles) ou croître (s'il est le siège de phénomènes irréversibles). Elle ne peut en aucun cas diminuer.

2°) EN INFORMATION :

Passons maintenant à l'information. Dans l'exposé de SHANNON, on trouve un théorème n° 7 dont l'énoncé est le suivant :

La production (out put) d'un traducteur à nombre fini d'états alimenté par une source statistique à nombre fini d'états correspond à celle d'une source statistique à nombre fini d'états dont l'entropie est inférieure ou égale à celle de l'alimentation. Si le traducteur est univoque, (ou sans bruit), ces deux entropies sont égales ce qui, d'après BRILLOUIN (6), signifie qu'un système de codage peut diminuer ou au plus maintenir constante la valeur de l'information. Ce théorème a une ressemblance indiscutable

avec le principe de Carnot ; d'abord, parce qu'il s'agit d'entropie ; ensuite, parce qu'il y a constance de l'entropie quand la traduction est reversible, variation de l'entropie si non ; enfin l'interprétation paraît très semblable : le principe de Carnot dit qu'aucune transformation ne peut augmenter l'énergie libre, la "néguentropie", pour utiliser le terme créé par M. BRILLOUIN ; le théorème de SHANNON dit qu'aucune traduction ni aucune transmission ne peut augmenter la quantité d'information.

3°) PRINCIPE DE NEGUENTROPIE DE L'INFORMATION ET GENERALISATION DU PRINCIPE DE CARNOT DE M. BRILLOUIN

a) Principe de néguentropie de l'information

M. BRILLOUIN pense qu'il s'agit du même principe et formule un principe de néguentropie de l'information appliqué à l'information "liée" que l'on rencontre lorsque les cas possibles à priori peuvent être interprétés comme les divers aspects d'un système physique, sinon l'information I_f est dite libre. La raison de cette distinction est que seule l'information liée sera considérée comme apparentée à l'entropie. Pour établir la relation entre l'information et l'entropie, nous allons considérer les cas également probables comme des complexions au sens de PLANCK. Dans ces conditions, on a :

	Information liée	Nombre de complexions	Entropie
Au départ	$I_{\ell_0} = 0$	P_0	$S_0 = k \log P_0$
A la fin	$I_{\ell_1} \neq 0$	$P_1 < P_0$	$S_1 = k \log P_1$

Il nous faut préciser les conditions dans lesquelles l'information a pu être obtenue et ceci nous oblige à décrire le schéma général d'une opération d'observation et de mesure. A côté du système considéré, nous devons avoir un instrument de mesure relié à diverses sources d'énergie dans le laboratoire à température T. Au début, le système à étudier et l'appareil de mesure sont isolés l'un de l'autre ; le premier a une entropie S_0, le second une entropie S'_0. Pour procéder à la mesure, il faut établir un couplage entre le système physique étudié et l'appareil de mesure. Une fois les opérations de mesure terminées, on sépare à nouveau les deux parties et l'on peut déterminer séparément l'entropie S_1 du système et celle S'_1 de l'instrument. L'entropie totale de l'ensemble, système observé et instrument de mesure, doit avoir augmenté, d'après Carnot :

$$S_0 + S'_0 \leq S_1 + S'_1$$

Pendant la mesure le système n'est pas isolé ; ceci permet à l'entropie du système de décroître lorsqu'on obtient de l'information, car cette information réduit le nombre de complexions. Elle a été fournie par l'appareil de mesure dont l'entropie a augmenté.

Nous avons vu au chapitre I que :

$$I = k(\mathrm{Log}\, P_0 - \mathrm{Log}\, P_1)$$

$$I_{\ell_1} = S_0 - S_1 \qquad\qquad S_1 = S_0 - I_{\ell_1}$$

L'information liée se manifeste par un terme négatif dans l'entropie totale du système physique et l'on peut conclure :

Information liée = décroissance de l'entropie S
= croissance de la néguentropie N

La néguentropie étant définie comme l'opposé de l'entropie.

Si maintenant nous isolons le système, le principe de Carnot conduit à poser que $\Delta S_1 \geqslant 0$

$$\text{ou} \quad \Delta(S_0 - I_{\ell_1}) \geqslant 0$$

Si S_0 correspond à la structure générale de sorte qu'il n'y ait pas surabondance de détermination pour S_0, S_0 reste constant et l'on a alors :

$$\Delta I_{\ell_1} \leqslant 0$$

et $\mathcal{N} = -S$ $\quad \Delta \mathcal{N}_1 \leqslant 0 \quad \quad \Delta(\mathcal{N}_0 + I_{\ell_1}) \leqslant 0$

qui exprime symboliquement le principe de dégradation de l'énergie. Le fait que ΔI_{ℓ_1} soit négatif correspond bien aux résultats antérieurs relatifs à l'information libre

$$\Delta I_f \leqslant 0$$

obtenus lorsque nous avons établi que l'information libre est maximale pour le cas d'égales probabilités.

L'une ou l'autre forme de l'information peut être transformée en néguentropie et l'information, libre ou non, ne peut être obtenue qu'en empruntant de la néguentropie à un système physique.

Supposons un gaz placé dans une enceinte volume A qui se trouve brusquement mise en communication avec une autre de volume B. L'entropie initiale S' est inférieure à

l'entropie S après détente et en diffère d'une quantité donnée par l'équation

$$S - S' = kn \left(\mathrm{Log}\, B - \mathrm{Log}\, A \right) = I$$

I mesurant l'information. L'accroissement d'entropie et la perte d'information vont de pair.

b) Généralisation du principe de Carnot

Un tel fait implique une généralisation du principe de Carnot si l'on fait intervenir l'un et l'autre type d'information :

Si $\Delta I_f \leqslant 0 \qquad \Delta N_1 \leqslant 0$

$$\Delta N_1 + \Delta I_f \leqslant 0$$

Soit $\Delta \left(S_1 - I_f \right) = \Delta \left(S_0 - I_{\ell_1} - I_f \right) = \Delta \left(S_0 - I \right) \geqslant 0$

que de l'information libre ou liée se présentent dans le problème. On peut donc écrire :

$$\Delta \left(N_0 + I \right) \leqslant 0$$

C'est le principe de la dégradation de l'énergie.

La somme de la néguentropie et de l'information doit rester constante dans le cas d'une transformation reversible et décroître dans le cas contraire. Ces inégalités s'appliquent seulement aux <u>valeurs moyennes</u>, comme la thermodynamique.

En résumé :

--- La néguentropie N correspond à l'information I

--- La température T implique un bruit thermique perturbant la transmission de l'information

--- L'énergie conserve son sens habituel

--- $\Delta Q = T\Delta S$ donne la chaleur engendrée par un certain processus

--- $\Delta W = T\Delta N - T\Delta I$ correspond au travail mécanique disponible

L'acquisition d'information au sujet d'un système physique correspond donc à un état inférieur de l'entropie de ce dernier. Une faible entropie implique un état instable qui, plus ou moins tard, évolue normalement vers la stabilité et une entropie élevée.

L'entropie est, en général, considérée comme exprimant l'état de désordre d'un système physique. On peut dire de façon plus précise que l'entropie mesure le manque d'information sur la véritable structure du système.

C) EXEMPLE : LE DEMON DE MAXWELL

Le démon de MAXWELL se présente comme un excellent exemple d'application de la théorie de l'information et montre très nettement le lien qui existe entre l'information et l'entropie. C'est, d'après MAXWELL "un être dont les facultés seraient si organisées qu'il puisse suivre chaque molécule dans sa course...". "Supposons qu'une enceinte soit divisée en deux parties A et B par une cloison dans laquelle est pratiquée une ouverture et qu'un être capable de voir individuellement les molécules ouvre et ferme cette ouverture de manière à permettre aux molécules les plus rapides de passer de A vers B et aux plus lentes

d'aller de B vers A. Il élèvera ainsi, sans dépense de travail, la température de B et diminuera celle de A, ce qui est en contradiction avec le second principe de la thermodynamique".

M. VON SMOLUCHOVSKI (7) a, le premier, signalé l'influence possible de l'agitation brownienne sur la porte de l'ouverture envisagée, agitation qui pourrait occasionner une ouverture et une fermeture désordonnées de cette dernière et perturber l'opération. Mais ce mouvement brownien ne constitue qu'une faille apparente dans le second principe en raison de la nature désordonnée imprévisible et de la très courte durée du mouvement.

Le problème de l'action du clapet est semblable à celui d'un redresseur idéal agissant sur les électrons individuels. Il redresserait l'agitation thermique des électrons, ce qui seraiten contradiction avec le second principe. Mais en fait il n'existe pas de redresseur idéal et le seul résultat que l'on peut obtenir est une agitation thermique dissymétrique. La tension redressée est précisément égale à la chute de tension aux bornes du redresseur.

Une très importante contribution a été apportée par SZILARD (8) qui a montré que le démon agit sur l'information relative au mouvement intime du gaz et transforme réellement l'information en néguentropie.

SLATER (9) a posé la question de savoir si le principe d'incertitude intervenait dans de tels problèmes. Le démon de MAXWELL doit estimer simultanément la position et la vitesse d'un atome donné, grandeurs qui ne peuvent être

mesurées en même temps avec une précision infinie à cause de l'inégalité

$$\Delta p \, . \, \Delta q \leq \hbar$$

Cette incertitude ne peut pas jouer de rôle pour les atomes lourds et les faibles pressions, ainsi que l'a noté DEMERS (10). Mais une question essentielle a été posée par DEMERS et BRILLOUIN : est-il réellement possible au démon de <u>voir</u> les atomes individuellement ? Le démon se trouve dans une enceinte en équilibre à une température constante et le rayonnement ne peut y être que celui du corps ; or, il est impossible de voir quoique ce soit à l'intérieur du corps noir. Il ne sert à rien de varier la température.

- Le démon, ne pouvant pas voir les molécules, peut chercher à les distinguer par les forces de VAN DER WAALS ou les forces dues aux dipôles électriques ou aux moments magnétiques. Tous ces champs ont leur intensité décroissant en r^{-2}. Le démon détecterait les molécules trop tard pour pouvoir ouvrir la porte sans fournir de travail. De plus, ces forces agiraient sur la porte et il faudrait fournir du travail supplémentaire pour la manoeuvrer.

La détection ne peut donc se faire que par l'intermédiaire du rayonnement. On peut munir le démon d'une torche électrique. La torche est une source de rayonnement qui n'est pas en équilibre ; elle fournit de la néguentropie au système et, de cette dernière, le démon tire de l'information. Utilisant cette information, il fait fonctionner la porte et reconstruit l'entropie négative.

Etudions ce processus de plus près. Le système envisagé comporte les éléments suivants :

- une batterie et une ampoule électrique
- un gaz à température T_0 renfermé dans l'enceinte
- le démon

La batterie porte la lampe à la température $T_1 > T_0$, telle que la lumière visible puisse être distinguée dans l'enceinte à la température T_0

La batterie fournit une énergie totale E et pas d'entropie.

La variation d'entropie du filament est :

$$S_f = -\frac{E}{T_1}$$ = néguentropie transmise au gaz

Si le démon n'intervient pas, l'énergie E est absorbée par le gaz à la température T_0 et l'on observe une augmentation d'entropie

$$S = \frac{E}{T_0} + S_f > 0$$

Le démon décèle une molécule lorsqu'au moins un quantum d'énergie $h\nu_1$ est diffracté par cette dernière et est absorbé par ses yeux. Ceci représente un accroissement final d'entropie

$$S_1 = \frac{h\nu_1}{T_0} = kb \qquad \text{où} \qquad \frac{h\nu_1}{kT_0} = b > 1$$

On obtient ainsi de l'information qui sera utilisée pour diminuer l'entropie du système. Celle-ci est initialement :

$$S_0 = k \operatorname{Log} P_0$$

P_0 étant le nombre de complexions de PLANCK. Après obtention de l'information, le système est mieux défini. P_0 est diminué de p

$$P_1 = P_0 - p$$

$$\Delta S_i = S - S_0 = k\,\Delta(\mathrm{Log}\,P) = -k\,\frac{p}{P_0}$$

avec $p < P_0$.

La compensation totale d'entropie ressort à :

$$\Delta S_1 + \Delta S_i = k\left(b - \frac{p}{P_0}\right) > 0$$

puisque $b > 1$ et $p/P_0 < 1$. En définitive on constate un accroissement de l'entropie de tout le système isolé, conformément au second principe.

Supposons maintenant qu'après un certain temps le démon a été capable d'obtenir une différence de température ΔT :

$$T_B > T_A \qquad\qquad T_B = T + \frac{1}{2}\,\Delta T$$

$$T_B - T_A = \Delta T \qquad\qquad T_A = T - \frac{1}{2}\,\Delta T$$

Le démon choisit une molécule rapide dans A, d'énergie cinétique $3/2\ kT\ (1 + \varepsilon_1)$ et la laisse pénétrer dans B, puis une molécule lente dans B, d'énergie $3/2\ kT(1 - \varepsilon_2)$ et la laisse entrer dans A. Pour observer ces deux molécules, le démon a besoin de deux quanta de lumière, d'où un accroissement d'entropie

$$\Delta S_d = 2\,k\,b$$

L'échange des molécules correspond à un transfert d'énergie

$$\Delta Q = \frac{3}{2}\,kT\,(\varepsilon_1 + \varepsilon_2)$$

soit une diminution d'entropie totale

$$\Delta S_i = \Delta Q \left(\frac{1}{T_B} - \frac{1}{T_A} \right) = -\frac{3}{2} k (\varepsilon_1 + \varepsilon_2) \frac{\Delta T}{T}$$

ε_1 et ε_2 sont faibles

$$\frac{\Delta T}{T} << 1 \qquad \Delta S_i = -\frac{3}{2} k \eta \qquad \eta << 1$$

$$\Delta S_d + \Delta S_i = k \left(2b - \frac{3}{2} \eta \right) > 0$$

De même si le démon est à température T_2 plus basse que T_0

$$h\nu > kT_2 \qquad \text{avec } T_2 < T_0$$

et la discussion se poursuit de la même manière. Quoiqu'il en soit, qu'on utilise une température T_1 supérieure à T_0 ou une température T_2 inférieure, il est toujours nécessaire d'avoir une différence de température. Un moteur thermique fera alors aussi bien que le démon.

Le principe de Carnot généralisé est donc satisfait.

M. RODD (11), formulant lui aussi le principe de néguentropie de l'information, corrige la forme de la démonstration de M. BRILLOUIN. La lampe doit émettre un rayonnement dont l'énergie $h\nu$ est plus grande que kT_0, où T_0 est la température de l'enceinte. Pour produire ce rayonnement, le filament de la lampe doit être chauffé à une température $T_1 > T_0$. Supposons qu'aucune énergie ne soit rayonnée par le filament pendant son échauffement ; alors l'augmentation d'entropie du filament pendant le changement de température est $m_{fil}\ C_{fil} \log T_1/T_0$. Quand le filament a atteint la température T_1 et commence à rayonner, il n'y a plus de changement d'entropie, puisque le filament ne change pas d'état.

(Le flux d'entropie est - E/T_1 où E est l'énergie totale de la lumière rayonnée, et la production d'entropie est + E/T_1). Supposons, pour que le démon reçoive de l'entropie, qu'un quantum d'énergie $h\nu$ soit diffracté par une molécule et absorbé par le démon. Comme

$$\frac{h\nu}{T_0} > k \qquad \frac{h\nu}{T_0} = kb \qquad b > 1$$

L'augmentation d'entropie du démon due à son absorption de ce quantum d'énergie est :

$$\Delta S_d = \frac{h\nu}{T_0} = kb$$

L'entropie du gaz dans son état original peut être donnée comme :

$$S = k \operatorname{Log} P$$

L'action du démon modifie le nombre d'états microscopiques accessibles au gaz de P à P - p_1. Donc

$$\Delta S_{gaz} = k \left[\operatorname{Log}(P - p_1) - \operatorname{Log} P\right]$$

$$= k \operatorname{Log}\left(1 - \frac{p_1}{P}\right)$$

puisque $\frac{p_1}{P} << 1$ $\qquad \operatorname{Log}\left(1 - \frac{p_1}{P}\right) \simeq - \frac{p_1}{P}$

Le changement d'entropie de la batterie fournissant le courant à la lampe dans ce processus est nul.

Définissant le système isolé comme le filament, le démon, le gaz et la batterie, nous obtenons pour ce système :

$$\Delta S = \Delta S_{fil} + \Delta S_{demon} + \Delta S_{gaz} + \Delta S_{batterie}$$

$$= m_{fil}\, c_{fil} \operatorname{Log} \frac{T_1}{T_0} + k\left(b - \frac{p_1}{P}\right) + 0$$

$$b > 1 \quad \text{et} \quad \frac{p_1}{p} < 1$$

$$\Delta S > 0$$

Nous voyons qu'en réalité le démon de MAXWELL ne viole pas le second principe, puisque l'entropie du système isolé augmente. Notons aussi que le chauffage du filament de la lampe est un processus irréversible qui rend irréversible le processus tout entier. A cause de cette irréversibilité l'entropie a augmenté même si l'information sur le système a augmenté.

M. BRILLOUIN a donc essayé de construire une théorie qui synthétiserait la théorie de l'information et la thermodynamique, l'entropie étant, selon lui, le manque d'information, l'ignorance quant à l'état microscopique du système. Cette théorie a le grand avantage de rendre concrets les rapports existant entre information et thermodynamique ; il semble cependant que l'analogie ait été poussée trop loin, qu'il existe des distinctions importantes à faire entre les deux.

4°) ENTROPIE - VARIABILITE ET ENTROPIE SPECIFICITE

Ainsi M. BONSACK ne veut pas considérer l'information comme une néguentropie. Sa théorie est basée sur deux notions : la variabilité et la spécificité. Du fait que l'entropie augmente lors d'une transmission d'information avec bruit alors que l'information diminue, le bruit étant une cause d'erreurs, il cherche à préciser la notion d'information.

a) En Thermodynamique, la notion centrale est celle d'entropie-variabilité, et non celle de spécificité.

L'entropie absolue, sa valeur nulle au zéro absolu, son augmentation avec la température et avec le volume ne se comprennent que s'il s'agit d'une variabilité. Et lorsqu'on a passé de la définition probabiliste de l'entropie à une définition statistique, en prenant tout simplement le logarithme du nombre de complexions possibles dans un certain état macroscopique, on a passé d'une spécificité à une variabilité.

b) En théorie de l'information, par contre, la notion centrale est celle de spécificité. Si l'on veut faire de la variabilité une mesure de la quantité d'information on aboutit au paradoxe cité ci-dessus de la transmission avec bruit.

L'entropie représente la variété des messages qu'un expéditeur peut envoyer. L'information concerne au contraire le message particulier qu'il a choisi, la spécificité qui est le rapport entre deux classes, celle des états possibles après le choix, soient n, et celle des états possibles avant le choix, soient N. Dans le cas de l'équiprobabilité

$$S = -\log p = -\log \frac{n}{N} = \log N - \log n$$

Si n = 1, il y a identité entre le concept défini précédemment et la spécificité présente. L'entropie est la spécificité d'un seul message et non pas d'une classe de messages parmi lesquels on choisirait au hasard. C'est une moyenne des spécificités des messages particuliers, ce qui nous ramène au fait qu'elle caractérise l'ensemble des formes qu'un système peut prendre. La spécificité, elle, est beaucoup plus individuelle.

La spécificité moyenne apparaît comme égale à la diminution de variabilité au cours du choix. Elle est une différence de deux entropies, de l'entropie du référentiel, log N, et de l'entropie après choix, log n. Comme l'ont souligné YAGLOM et YAGLOM (12), l'information est équivalente à une différence de potentiel, le potentiel étant l'incertitude. Cette distinction entre spécificité et variabilité permet de préciser la question du signe (entropie ou néguentropie). Une variation de l'entropie après choix se traduit par une variation inverse de la spécificité ; une diminution de cette entropie correspond à une augmentation de spécificité (l'entropie du référentiel étant maintenue constante). Cette relation montre les raisons de l'identification de la quantité d'information à une néguentropie, mais également les limites de cette identification.

La spécificité individuelle se calcule, dans le cas où les probabilités ne sont pas égales, à l'aide de la formule :

$$S = - N \sum f_i \operatorname{Log} p_i$$

f_i étant la fréquence relative du symbole X_i dans cette complexion particulière.

Au contraire, si on calcule la spécificité moyenne ou l'entropie, on peut substituer à la fréquence relative la probabilité et l'on obtient :

$$S = - N \sum p_i \operatorname{Log} p_i$$

Dans le cas d'un grand nombre d'états possibles, f_i se rapproche de p_i et la spécificité se rapproche de l'entropie-variabilité.

Reprenons l'exemple du gaz placé dans une enceinte de volume A qui se trouve brusquement mise en communication avec une autre de volume B. L'entropie du gaz est plus grande dans le second état que dans le premier. La spécificité par contre est plus grande dans le premier état. Les dispositions rares dans certaines conditions ont une spécificité individuelle plus élevée que les positions fréquentes. C'est ainsi que l'on pourra étudier la cinétique de la diffusion du gaz de A en A + B, alors qu'en thermodynamique on saute d'une situation à l'autre. La théorie de l'information permettra de traiter des états hors d'équilibre.

c) Le principe de Carnot

Le principe de Carnot de la thermodynamique s'énonce alors : la variabilité d'un système physique fermé tend à augmenter ; si les transformations dont il est le siège sont toutes reversibles, elle reste constante. Si par contre, on veut appliquer ce principe tel quel à la théorie de l'information, on aboutit encore une fois à des paradoxes : certains "traducteurs" diminuent la variabilité. Il faut donc formuler un autre principe de Carnot qui dit que la spécificité ne peut que diminuer ou au plus rester constante. Ce principe de Carnot s'applique sans difficultés aux structures spécifiques autres que les messages ; l'évolution qu'il autorise se traduit par l'altération, l'usure qui correspondent en général à une perte de spécificité.

Quant au principe de Carnot généralisé énoncé par M. BRILLOUIN, il a l'avantage de pouvoir s'appliquer tel quel à la thermodynamique, mais il met sur le même pied la variabilité et la spécificité. Il contient certainement une part de vérité. L'augmentation de la spécificité peut être due à deux causes :

- Le resserrement de l'ensemble spécifié
- L'extension du référentiel auquel il est rapporté.

Dans le premier cas, le principe généralisé est applicable sous la forme suivante :

Toute diminution de la variabilité dans une certaine partie d'un système fermé se fait aux dépens d'une augmentation de la variabilité dans une autre partie du même système. Il y a une sorte de compensation au sens de Clausius.

d) Le démon de MAXWELL

Dans le cas du démon de MAXWELL, il y a également compensation : le démon ne peut créer de la néguentropie sans en dépenser par ailleurs, pour voir les molécules ou immobiliser son volet. Au lieu d'étudier la perception du démon, M. BONSACK étudie son action et cherche à montrer que toute influence déterminante sur un système exige des processus irréversibles qui provoqueront une augmentation de l'entropie. Pour qu'un démon puisse remplir son office, il faut qu'il y ait couplage entre le mouvement de la molécule incidente et l'état de réaction du démon. Or, pour que tel mouvement de la molécule mette le démon dans tel état déterminé il faut ou bien que celui ci soit dans un état déterminé avant l'action de la molécule, ou bien que le mouvement de la molécule le force à se trouver dans un certain état, quel qu'ait été son état antérieur. Ceci est possible si l'on fait intervenir des phénomènes irréversibles, si l'on détruit de la néguentropie.

Encore ne s'agit-il ici que d'ordre et non d'information. Nous avons vu que la spécificité pouvait être définie

comme la diminution de la variabilité au cours du choix. On peut de même la considérer comme la diminution subjective de la variabilité par l'information. L'information R apportée par un message est mesurée à l'aide de la différence d'entropie de deux incertitudes :

$$R = H(x) - H_y(x)$$

où H(x) est l'entropie de la source donnée, et Hy(x) l'entropie sachant que y s'est produit. L'information reçue restreint les possibilités à envisager et réduit l'entropie.

M. BRILLOUIN pense que cette diminution d'entropie est contraire au principe de Carnot et postule finalement qu'elle doit être compensée, que l'acquisition d'information doit être payée par une augmentation de l'entropie. C'est ce postulat supplémentaire, principe de Carnot généralisé, ou principe de néguentropie de l'information, que M. BRILLOUIN applique au démon de MAXWELL et que M. BONSACK réfute, non dans l'idée, qui contient quelque chose de juste, mais dans la forme et la démonstration.

e) Généralisation

En théorie de l'information le problème se pose différemment : il est possible d'augmenter la spécificité d'un système au prix de l'augmentation de la variabilité physique. Cette augmentation de la spécificité correspond à un choix, à une sélection, sélection qui n'est pas un privilège de l'homme : même un crible, un filtre peuvent sélectionner un ensemble d'une plus grande spécificité au sein

d'un référentiel de spécificité moindre.

A cette augmentation de spécificité seront ramenées les innovations au cours de l'évolution biologique ou au cours de l'évolution intellectuelle. La théorie proposée comprend trois mécanismes fondamentaux : la variation, la sélection et la reproduction.

La variation est aléatoire, elle est provoquée par un "bruit". Elle ne pose pas de problème : la spécificité diminue conformément au principe de Carnot de l'information.

Nous avons parlé de la sélection.

Quant à la reproduction, elle n'augmente pas la spécificité : la copie étant liée à l'original, le référentiel des couples original-copie est moins variable que ne le seraient deux originaux indépendants.

Ces trois mécanismes peuvent se combiner pour donner des mécanismes dérivés : reproduction variée, variation liée,.. Cette théorie peut être appliquée d'une part à l'évolution, où elle correspond à la théorie darwinienne ou à ses dérivés récents (théorie synthétique de SIMPSON), d'autre part à l'innovation intellectuelle, à l'invention, à la découverte, à la création artistique.

CHAPITRE III

LE FORMALISME D'ENTROPIE MAXIMALE - APPLICATION AUX TRANSFORMATIONS IRREVERSIBLES

Nous avons jusqu'à maintenant examiné les rapports existant entre la théorie de l'information et la thermodynamique. Nous avons été obligé de passer par l'intermédiaire de probabilités, de nombres de complexions et, par le même, de nous rapprocher de la mécanique statistique.

JAYNES a discuté l'utilisation de la fonction de partition de la mécanique statistique en théorie de l'information. Il propose que la probabilité en mécanique statistique soit interprétée non comme des fréquences relatives objectives, mais, au sens "subjectif", comme un moyen de décrire la force avec laquelle nous croyons que tel évènement va arriver. Ceci pour raisonner le mieux possible à propos des cas individuels.

Au début de tout problème en théorie des probabilités, il faut assigner une distribution de probabilités qui corresponde à ce que nous connaissons du problème. La distribution P_n qui décrit le plus honnêtement ce que nous savons est la plus étendue, la plus "plate" possible. Mais c'est aussi celle

qui ne permet aucune conclusion nette qui ne soit pas garantie par les données. Il faut une mesure de l'étendue de la distribution que nous puissions rendre maximale.

La mesure de l'information de SHANNON

$$S_I = -\sum_i p_i \operatorname{Log} p_i$$

est, avons-nous vu, la mesure de l'incertitude dans une distribution de probabilité.

Pour JAYNES cette expression ne représente l'entropie que pour quelques distributions et dans quelques cas physiques. L'entropie d'information S_I et l'entropie thermodynamique expérimentale S_e sont des concepts entièrement différents. Contrairement à M. BRILLOUIN, il ne postule aucune relation entre elles, mais veut déduire tous les rapports possibles des faits mathématiques et physiques.

1°) LE FORMALISME GENERAL D'ENTROPIE MAXIMALE

On connait en mécanique statistique une formule identique à celle de l'entropie d'information. Elle peut être déduite de :

$$S = k \operatorname{Log} W$$

en utilisant un ensemble. Mais cette procédure n'est pas nécessaire ici. Nous traitons d'un système avec l'information que nous en avons. L'obligation usuelle de trouver les distributions canoniques ou grand-canoniques devient ici la maximisation de l'information manquante de façon consistante avec l'information donnée, l'information donnée étant l'énergie moyenne dans l'ensemble canonique ou l'énergie moyenne et le nombre moyen de particules dans le grand ensemble.

Supposons en effet que la quantité x puisse prendre les valeurs $(x_1, x_2 \ldots\ldots x_n)$ où n peut être fini ou infini et que les valeurs moyennes de plusieurs fonctions $f_1(x), \ldots\ldots f_n(x)$ soient données, avec $m < n$. Le problème est de trouver la probabilité

$$p_i = p(x_i) \qquad \text{qui satisfait les données}$$

$$p_i \geqslant 0 \qquad \sum_{i=1}^{n} p_i = 1$$

$$\sum_{i=1}^{n} p_i \, f_k(x_i) = \langle f_k(x) \rangle = F_k \qquad k = 1, 2, \ldots m$$

et rend maximale l'entropie d'information

$$S_I = - \sum_{i=1}^{n} p_i \operatorname{Log} p_i$$

En utilisant la propriété fondamentale de l'entropie

$$\sum_{i=1}^{n} p_i \operatorname{Log} \frac{p_i}{u_i} \geqslant \sum_{i=1}^{n} p_i \left(1 - \frac{u_i}{p_i} \right) = 0$$

avec égalité si pi = ui

et en choisissant

$$u_i = \frac{1}{Z(\lambda_1, \ldots \lambda_n)} \exp \left\{ - \lambda_1 f_1(x_i) \ldots - \lambda_m f_m(x_i) \right\}$$

on montre que S_I atteint la valeur maximale possible

$$(S_I)_{max} = \operatorname{Log} Z + \sum_{k=1}^{m} \lambda_k \langle f_k \rangle$$

si et seulement si pi est la distribution canonique généralisée ni. Il faut alors choisir les constantes λ_k telles que

$$\lambda_k = \frac{\partial S}{\partial \langle f_k \rangle} \qquad k = 1, \ldots m$$

Les fonctions log Z (λ_1, ... λ_n) et S ($<f_1>$ .... $<f_n>$) sont équivalentes, c'est à dire que chacune donne une information entière sur la probabilité de distribution.

Si, en plus de leur dépendance en x, les fonctions f_k dépendent de paramètres α, la fonction de partition a aussi une dépendance explicite en α. Si α_j varie de $\delta\ \alpha_j$, l'entropie maximale que l'on puisse atteindre est changée de :

$$\delta S = \sum_{k=1}^{m} \lambda_k \, \delta Q_k$$

avec $$\delta Q_k = \delta < f_k > - < \delta f_k >$$

où $$< \delta f_k > = < \sum_{j=1}^{r} \frac{\partial f_k}{\partial \alpha_j} \partial \alpha_j >$$

λ_k est un facteur intégrant de Q_k tel que $\sum_{k=1}^{m} \lambda_k \, \delta Q_k$ soit la différentielle exacte d'une fonction d'état $S(\lambda_1, \ldots \lambda_m ; \alpha_1 \ldots\ldots \alpha_r)$.

Toutes les relations qu'on peut ainsi obtenir sont des conséquences élémentaires d'une maximisation de l'entropie d'information sujette à des restrictions sur les valeurs moyennes de certaines quantités. Quoiqu'ayant une ressemblance formelle avec la mécanique statistique, elles ne se réfèrent pas à la physique et peuvent être appliquées aussi bien à "l'engineering" qu'à l'économie. Il suffit que la situation puisse être décrite par :

1°) l'énumération d'un ensemble discret de possibilités

2°) la spécification des valeurs moyennes des diverses quantités.

Dans la plupart des problèmes, il est intéressant de faire les meilleures prévisions possibles pour un cas particulier. Le formalisme d'entropie maximale ne peut pas décrire un état existant, mais plutôt donner les meilleures prévisions que nous puissions faire avec l'information donnée.

On ne peut pas certifier que les prévisions soient justes, on peut dire seulement que ce sont les meilleures que l'on puisse faire avec les données que l'on possède.

Cependant, dans les cas où on peut imaginer x_i comme étant le résultat d'expériences aléatoires que l'on peut répéter de nombreuses fois, il est possible de donner une interprétation plus "objective" de ce formalisme dans le même sens que ce qu'a fait BOLTZMANN. Si l'expérience est répétée N fois, le résultat x_i sera obtenu m_i fois. Nous voulons faire la meilleure estimation des nombres m_i. Raisonnant alors comme pour l'entropie de BOLTZMANN, on retrouve l'entropie d'information

$$S = - \sum_{i=1}^{n} g_i \, \mathrm{Log}\, g_i$$

où $g_i = \frac{m_i}{N}$ est la fréquence relative

La probabilité p_i que la théorie de l'information attribue à l'évènement x_i lors d'une expérience isolée est numériquement égale à l'estimation de la fréquence relative g_i de ce résultat lors d'expériences nombreuses et les plus diverses.

Les fluctuations peuvent décrire soit notre incertitude quant à la valeur vraie de $f_k(x)$ lors d'une expérience isolée, soit la meilleure estimation que nous puissions faire de la moyenne des écarts de $\langle f_k \rangle$ lorsque l'expérience est répétée de nombreuses fois.

2°) APPLICATION A LA THERMODYNAMIQUE DES ETATS D'EQUILIBRE

On peut bien sûr appliquer ceci à la thermodynamique des états d'équilibre. On a alors :

$$< E > = -\frac{\partial}{\partial \lambda} \operatorname{Log} Z = U \qquad < P > = \frac{1}{\lambda} \frac{\partial}{\partial V} \operatorname{Log} Z$$

Mais aucune quantité de mathématique ne peut prouver quoi que ce soit quant aux faits expérimentaux. JAYNES propose de lier ces probabilités subjectives aux résultats de la physique expérimentale grâce à l'hypothèse physique suivante :

"Les propriétés d'un système en équilibre thermodynamique mesurées expérimentalement correspondent aux résultats calculés par les méthodes usuelles de mécanique statistique, c'est-à-dire à partir de l'ensemble canonique ou du grand ensemble approprié au système".

On peut alors poser :

$$\lambda = \frac{1}{kT'}$$

$$S'_{exp} = k \left(S_I \right)_{max}$$

Si nous acceptons l'hypothèse précédente, l'identification de l'entropie est complète et la liaison entre l'entropie de la théorie de l'information et l'entropie expérimentale pour le problème considéré peut être établie en théorème :

Soit $p_i \equiv \text{prob}\ (E_i)$ toute attribution de probabilité qui se conforme aux données dans le sens que $<E> = \sum_i p_i E_i$ est l'énergie mesurée. Soit $S_I \equiv -\sum_i p_i \log p_i$ l'entropie d'information correspondante, et S_e l'entropie du système mesurée expérimentalement. La constante additive est choisie

de sorte qu'à température nulle S_e = log n où n est la dégénérescence de l'état fondamental, et S_e est exprimée en unités telles que la constante de BOLTZMANN k ≡ 1. Alors :

$$S_I \leq S_e$$

avec égalité si et seulement si p_i est choisi comme la distribution canonique

$$p_i = \frac{1}{Z} \exp \left\{ - \lambda E_i(V) \right\}$$

Si non seulement <E> est connu, mais aussi la précision de la mesure, par exemple $<E^2>$, on peut écrire :

$$f_1(x_i, \alpha) = E_i(V) \qquad f_2(x_i, \alpha) = E_i^2(V)$$

En pratique la variance $<E^2> - <E>^2$ de tout système macroscopique ayant des propriétés thermodynamiques reproductibles est faible devant les erreurs expérimentales. C'est ce qui fait le succès du formalisme de l'ensemble canonique de GIBBS.

<u>Généralisation</u>

On peut généraliser ce formalisme :

--- aux systèmes décrits par une matrice de densité

--- aux distributions de probabilité continues que l'on trouve en théorie classique.

a) <u>Matrice de densité</u>

La valeur "attendue" d'un opérateur F_k d'un système décrit par la matrice de densité ρ est :

$$< F_k > = Tr \left(\rho F_k \right)$$

L'entropie d'information est :

$$S_I = -\,\mathrm{Tr}\,(\rho \,\mathrm{Log}\, \rho)$$

Maximiser S_I soumis à des contraintes imposées par la connaissance de F_k donne :

$$\rho = \frac{1}{Z(\lambda_1, \dots \lambda_m)} \exp\{-\lambda_1 F_1 \;----\; -\lambda_m F_m\}$$

où $$Z(\lambda_1, \dots \lambda_m) = \mathrm{Tr}\,\exp\{-\lambda_1 F_1 \dots -\lambda_m F_m\}$$

b) Si N(x) est la densité continue de probabilité

$$S_I = -\int W(x)\,\mathrm{Log}\,\frac{W(x)}{m(x)}\,dx$$

où m(x) est une "mesure" qui rend l'entropie invariante pour un changement de paramètre.

En mécanique classique, si nous appliquons ceci aux fonctions de LIOUVILLE pour un système de N particules, $W_N(x_1\; p_1 \;\dots\; x_N\; p_N\; t)\; d^3x_1 \;\dots\; d^3p_N$ est la probabilité qu'à un instant t le système soit dans l'élément $d^3x_1 \;\dots\; d^3p_N$ de l'espace des phases. On prend m = Constante, en conséquence du fait que chaque état quantique discret correspond à un volume h^{3N} de l'espace de phases classiques. Notre mesure de l'information sera alors l'opposé de la fonction de GIBBS H_G

$$-S_I = H_G = \int W_N \,\mathrm{Log}\, W_N \, d\tau$$

3°) MECANIQUE STATISTIQUE DES ETATS HORS D'EQUILIBRE BASEE SUR LA THEORIE DE L'INFORMATION

Ce formalisme d'entropie d'information maximale a le grand avantage, comme nous l'avons déjà vu au chapitre II, de pouvoir traiter des états hors d'équilibre. Rien dans ce qui a été dit ne restreint les quantités connues à être des constantes du mouvement. Il n'y a aucune raison pour ne pas traiter de la même façon des quantités variables dans le temps. Pour prévoir ce qui se passera pour des temps différents de l'instant initial, on doit en principe résoudre les équations du mouvement ou l'équation de LIOUVILLE

$$W_N + [W_N, H] = 0$$

où H est l'hamiltonien.

Ces calculs étant extrêmement compliqués, il faut réduire cette équation à une équation irréversible appelée indifféremment équation de BOLTZMANN, équation de vitesse ou équation "principale".

La fonction de distribution f(x, p, t) de BOLTZMANN et la fonction de LIOUVILLE W_1 pour une particule ne sont pas équivalentes. BOLTZMANN définit f comme donnant le nombre réel de particules dans les différentes cellules de l'espace des phases à six dimensions.

Si R est l'ensemble des points de phase comprenant une cellule, le nombre de particules dans R est :

$$n_R = \int_R f(x, p, t)\, d^3x\, d^3p$$

f est donc une variable aléatoire et non pas une distribution de probabilité. f et W_1 sont reliés par :

$$< f > = \mathcal{N} \; W_1$$

où la valeur moyenne dénote une moyenne sur la fonction de LIOUVILLE W_N. $W_N \, d\tau$ est la probabilité pour que le système individuel soit dans la région $d\tau$ de l'espace des phases. W_1 ne contient pas toute l'information sur la distribution de particules dans l'espace des phases. Elle ne détermine que la valeur probable du nombre d'occupation des particules :

$$< n_R > = \mathcal{N} \int_R W_1 \, (x, p, t) \, d^3x \, d^3p$$

Cette intégrale représente la probabilité qu'une particule spécifiée soit dans la cellule de phase R. Ce n'est pas la même chose que la fraction de particule dans cette cellule, mais cela représente la valeur probable de cette fraction.

La fonction de LIOUVILLE ne définit rien de précis sur le nombre de particules dans R si la variance n'est pas faible. Si on décrit l'équilibre thermique par $W_N \sim \exp(-\beta H)$ la variance peut nous donner la densité du fluide, nous dire s'il n'y a qu'une phase ou si deux phases coexistent.

De la même façon, en déduisant les lois de l'hydrodynamique de l'équation de LIOUVILLE, on peut prédire la densité de moments. Si la variance est faible, le flux est laminaire. Si elle est grande, en incorporant toute l'information sur les conditions expérimentales imposées, la théorie peut dire que le flux n'est pas reproductible expérimentalement : il est donc turbulent.

Entropie et probabilité

JAYNES compare alors l'entropie de BOLTZMANN

$$H_B = \int f \operatorname{Log} f \, d^3x \, d^3p$$

et l'entropie donnée par GIBBS

$$H_G = \frac{1}{N} \int W_N \operatorname{Log} W_N \, d\nu$$

- L'entropie de GIBBS a une relation simple et universelle avec l'entropie ; pour toute probabilité correspondant aux paramètres mesurés de la thermodynamique, nous avons

$S \geqslant - k\, H_G$, avec égalité si et seulement si H_G est calculé à partir de l'ensemble canonique ou du grand ensemble approprié.

- L'entropie de BOLTZMANN est liée à l'entropie dans le seul cas du gaz idéal de BOLTZMANN (c'est-à-dire ni de BOSE, ni de FERMI). En général, $H_B \leqslant H_G$ et l'entropie peut être soit supérieure soit inférieure à $- k\, H_B$.

- La constance de l'entropie de GIBBS, loin de s'opposer à l'accroissement d'entropie, est la seule propriété dynamique nécessaire pour démontrer cette augmentation.

- L'entropie de GIBBS fournit une définition généralisée de l'entropie pour les cas hors d'équilibre, de sorte que l'énoncé usuel du second principe reste valable. Il donne, cependant, une nouvelle règle disant quels états hors d'équilibre sont accessibles à partir d'autres avec un processus adiabatique.

On a finalement $\quad -k\, H_B \geqslant -k\, H_G \leqslant S_e$

S_e étant l'entropie expérimentale.

Systèmes "fermés"

KATZ a traité du problème des états hors d'équilibre dans le cas de systèmes "fermés", c'est-à-dire dont l'hamiltonien est connu.

Il considère un système classique dont les variables sont les positions et les moments q_i et p_i. Le courant au point x est une fonction définie des variables.

$$J(t) = \frac{1}{m} \sum_i p_i(t)\, \delta^3 [q_i(t) - x]$$

Le courant initial est la même fonction des variables initiales qui à leur tour peuvent s'exprimer en fonction des variables au temps t et du temps :

$$p_i(0) = P_i [t, q_j(t), p_j(t)]$$

$$q_i(0) = Q_i [t, q_j(t), p_j(t)]$$

de sorte que le courant initial peut être exprimé au temps t comme une fonction différente de $q_i(t)$ et $p_i(t)$

$$J(0) = \frac{1}{m} \sum_i P_i [t, q_i(t), p_i(t)]\, \delta^3 \left\{ Q_i [t, q_j(t), p_j(t)] - x \right\}$$

Si J(t) est connu au temps t = 0, alors au temps t on connaît J(0). Cette information peut être prise en considération comme toute autre information au temps t. De cette façon, on peut traiter de situations hors d'équilibre et étudier l'approche de l'équilibre. Il faut noter qu'ici les équations du mouvement sont utilisées pour transférer l'information d'un temps à un autre. L'information ainsi transférée est considérée comme la même et non comme une nouvelle donnée. Les équations du mouvement ne créent pas d'information dont la seule source reste la mesure.

C O N C L U S I O N

La théorie de l'information a pris naissance dans l'étude des télécommunications et la plupart de ses applications actuelles appartiennent à ce domaine. Nous avons montré que la théorie peut également être utile en science pure et plus particulièrement en physique.

La ressemblance entre l'information et l'entropie a été décelée par SHANNON. L'information et l'entropie physique sont de même nature. L'entropie donne une mesure du manque d'information détaillée concernant un système physique. L'information représente un terme négatif dans l'entropie d'un système pour BRILLOUIN qui établit ainsi le principe de néguentropie de l'information. Lors d'une mesure, l'accroissement d'entropie du laboratoire contenant le système observé et les instruments de mesure est toujours supérieur au gain d'information. Ce résultat conduit à une extension du principe de Carnot qu'on peut appliquer au démon de MAXWELL. BONSACK discute les arguments en distinguant entre l'entropie-variabilité de la physique et l'entropie-spécificité de l'information.

En pratique, les quantités d'information correspondent à une valeur extrêmement faible de la néguentropie.

La théorie de l'information, en accord complet avec le principe d'incertitude, nous apprend de plus qu'une expérience rigoureuse est absolument irréalisable car elle donnerait une quantité d'information infinie et exigerait une dépense infinie en néguentropie.

La mécanique statistique permet de faire des prévisions sur la base d'une information incomplète sur le système considéré. Une théorie de l'information sert à définir la "meilleure prévision" basée sur une information donnée. L'obligation de trouver la distribution canonique ou grand canonique correspondant à un problème considéré devient celle de rendre maximale l'information manquante de façon consistante avec l'information donnée. JAYNES a donné en ce sens une formulation très nette de la mécanique statistique des états en équilibre.

La théorie de l'information ainsi faite ignore complètement la valeur de l'information manipulée, transmise ou transformée. Une machine à calculer traduit, code et décode, mais n'ajoute aucune information nouvelle et ne fait que la reproduire dans un langage différent, probablement avec des pertes. Elle ajoute cependant à la valeur pratique de l'information. Alors que la mesure de l'information est une quantité absolue, sa valeur est relative pour un utilisateur donné. Cette valeur relative est au plus égale à l'information absolue. Il semble qu'une étude en ce sens reste à faire.

B I B L I O G R A P H I E

(1) G. RAISBECK - Information Theory The MIT Press - Cambridge U.S.A. (1964)

(2) N. ABRAMSON - Information Theory and Coding - Mac Graw Hill (1963)

(3) C.E. SHANNON et W. WEAVER - The Mathematical Theory of Communication - U. of Illinois - Urbana (1949)

(4) F. BONSACK - Information, Thermodynamique - Vie et pensée Gauthier-Villars - Paris (1961)

(5) A. FEINSTEIN - Foundation of Information Theory - Mac Graw Hill New York (1958)

(6) L. BRILLOUIN - La Science et la Théorie de l'Information Masson (1959)

(7) V. SMOLUCHOVSKI - Physik Z. 13 1069 (1912)

(8) L. SZILARD - Z. Physik 53, 840 (1929)

(9) J.C. SLATER - Introduction to Chemical Physics - Mac Graw Hill New York (1939)

(10) P. DEMERS - Can J. research 22, 27 (1944)
23, 47 (1945)

(11) P. RODD - Am. J. Phys. 32, 333 (1964)

(12) A.M. YAGLOM et I.M. YAGLOM - Probabilité et Information Dunod (Paris) 1959

(13) E.T. JAYNES - Statistical Physics - Brandeis Lectures vol. 3 - Benjamin (1962)

(14) A. KATZ - Nuovo cimento 33, 1544 (1964)
33, 1552 (1964)