

Chapitre 1

Introduction à la Physique Statistique

1.1 Introduction

La physique statistique est née de la tentative de décrire le comportement des gaz en partant d'une description des mouvements des molécules individuelles. Elle a donc eu pour objectif de retrouver les notions de la thermodynamique classique telles que pression, énergie, entropie, à partir d'une description des comportements individuels des molécules. Il s'agit donc avant tout d'un passage d'une description microscopique à une description macroscopique ou, en d'autres termes, d'un changement d'échelle. Cette approche a été étendue à bien d'autres systèmes physiques de sorte que la description des gaz est loin d'être aujourd'hui le meilleur exemple d'application de la physique statistique. Ainsi, il n'est pas possible de comprendre le comportement d'un noyau, d'une étoile, d'un système magnétique, d'un liquide sans faire appel aux notions de physique statistique tant il est vrai que les propriétés microscopiques déterminent le comportement macroscopique. Au-delà du champ de la physique, cette façon d'aborder les problèmes macroscopiques en partant du microscopique a été appliquée à de nombreux autres domaines. La théorie de l'information fondée par Shannon a emprunté à la physique statistique le concept de l'entropie statistique. Les problèmes rencontrés en économie sont de même nature en ce sens que les évolutions économiques à l'échelle d'un pays résultent de l'addition des comportements d'un très grand nombre d'acteurs. Se pose alors la question du passage des comportements individuels à des évolutions collectives. Un autre exemple est celui de la modélisation du trafic routier.

Examinons les motivations qui conduisent à une description statistique d'un système physique. Un gaz parfait est entièrement décrit du point de vue microscopique par la donnée de la position et de l'impulsion de tous les atomes le composant. Cette description complète du système microscopique est impraticable car elle contient un trop grand nombre d'informations que l'on ne peut pas traiter. De surcroît, la description microscopique est souvent illusoire car il est fréquemment impossible de définir complètement avec une précision suffisante l'état d'un système. On voit donc que l'approche statistique est bien souvent la seule voie possible. A ce stade, la méthode statistique apparaît comme une technique de moyenne permettant une simplification. Mais on ne peut pas réduire l'approche statistique à une simple technique de moyenne qui permettrait de remplacer par exemple la connaissance détaillée de l'énergie cinétique de chaque particule par la valeur moyenne de l'énergie cinétique.

L'un des aspects les plus intéressants des approches statistiques est leur aptitude à faire émerger de nouveaux concepts intrinsèquement macroscopiques. La thermodynamique nous apprend qu'une description macroscopique complète du système s'appuie sur des grandeurs telles que la température ou l'entropie. Ces grandeurs ne sont pas présentes dans la description des états microscopiques du système. On ne sait pas définir la température ou l'entropie d'un atome unique. La théorie statistique doit être capable de l'introduire naturellement. De façon générale, une théorie statistique d'un système physique doit être capable de dégager systématiquement les quantités pertinentes pour une description correcte du système physique à l'échelle macroscopique.

Un deuxième exemple de phénomène essentiellement macroscopique est le phénomène de transition de phase. La question de l'organisation de moments magnétiques d'un grand nombre d'atomes ou encore la transition solide-liquide ou bien encore la transition écoulement fluide/ bouchin du trafic routier sont des exemples de phénomènes intrinsèquement macroscopiques. La physique statistique est ainsi la discipline qui s'occupe des systèmes complexes.

Nous allons introduire dans ce chapitre les concepts qui sont indispensables à la description statistique d'un système. Nous allons définir ce qu'est l'état d'un système, puis ce qu'est la probabilité associée à un état. Enfin, nous énoncerons le principe fondamental de la physique statistique. L'une des caractéristiques les plus remarquables de cette approche est que ce principe étant posé, le reste du cours va consister à en déduire des conséquences. Ce faisant, nous serons notamment amenés à reconstruire la thermodynamique classique. Nous pourrions retrouver de nombreuses lois macroscopiques, le plus souvent en utilisant des modèles élémentaires pour décrire le comportement microscopique. Il est frappant de voir à quel point un modèle très simplifié du comportement microscopique (pourvu que les paramètres pertinents aient été bien identifiés) permet de rendre compte avec une grande précision des comportements macroscopiques.

Il faut souligner une autre caractéristique de la physique statistique. Au-delà de sa capacité à faire émerger les grandeurs spécifiques au monde macroscopique et à calculer leur valeur moyenne, la physique statistique permet de calculer les fluctuations et les fonctions de corrélation des différentes grandeurs. Cet aspect -qui ne peut se traiter dans le cadre de la thermodynamique classique- joue un rôle fondamental, notamment dans l'étude des phénomènes de transport au voisinage de l'équilibre. Il existe en effet une relation générale entre les fluctuations d'un système et les phénomènes de transport lorsque le système est hors équilibre.

1.2 État d'un système. Moyenne d'ensemble

1.2.1 État d'un système en mécanique quantique et en physique classique. Espace des phases

La notion d'état a été introduite dans le cours de physique quantique. On sait qu'un état stationnaire d'un système physique est complètement caractérisé par la donnée des nombres quantiques. Par exemple, un système constitué d'un atome dans une boîte est décrit par la donnée de trois nombres quantiques. S'il s'agit d'une molécule diatomique, il faut ajouter les nombres quantiques donnant le moment cinétique et sa projection suivant un axe ainsi que le nombre quantique donnant l'état de la vibration.

La description classique du même système requiert la donnée de la position \mathbf{r} et de l'impulsion \mathbf{p} de l'atome. Un degré de liberté de rotation sera décrit le couple de variable posi-

tion angulaire θ , et moment cinétique J_θ . À partir de là, il est possible de prédire l'évolution du système. Dans le cas d'un gaz comportant N atomes, il faut donner $3N$ coordonnées d'espace et $3N$ coordonnées d'impulsion pour caractériser le système. En introduisant un espace à $6N$ coordonnées, appelé *espace des phases*, on peut associer un point à chaque état.

En résumé :

- Un état quantique est caractérisé par la donnée des nombres quantiques de l'ensemble complet d'observables qui commutent (ECOC) qui décrit le système.

- Un état classique est représenté par un point dans l'espace des phases.

1.2.2 Moyenne statistique et moyenne quantique

À l'évidence, la physique statistique va fréquemment utiliser la notion de moyenne. Il est essentiel de préciser très clairement sa définition. Tout d'abord et afin d'éviter toute confusion, il faut rappeler ce qu'est la moyenne en physique quantique.

Lorsqu'un système se trouve dans un état propre d'une observable, la prédiction du résultat d'une mesure de cette observable est certaine. Si l'état $|r\rangle$ du système est décrit par une superposition linéaire d'états propres ϕ du type

$$|r\rangle = \sum_i c_i \phi_i$$

alors le résultat n'est pas certain mais on sait calculer la probabilité d'obtenir un résultat donné. Si l'on considère maintenant une observable qui ne commute pas avec l'hamiltonien, le résultat de la mesure de la grandeur correspondante effectué sur un système qui est dans un état propre ϕ_i du hamiltonien n'est pas certain. Cependant, la probabilité d'obtenir un résultat peut être calculée également en projetant l'état ϕ_i sur les états propres de l'observable. Il y a donc en physique quantique une notion de probabilité d'un résultat de mesure qui est intrinsèquement quantique. Toutefois, dans ce qui précède, nous avons supposé que le système est dans un état connu. Si cet état est noté $|r\rangle$, alors, la valeur moyenne d'une observable O est donnée par $\langle r|O|r\rangle$.

La moyenne en physique statistique est de toute autre nature. Elle a pour but de calculer des probabilités de résultats de mesure lorsque l'on ne sait pas dans quel état se trouve le système. Le problème clé de la physique statistique est donc de calculer la probabilité que le système soit dans un état donné. Une fois l'état connu, les grandeurs physiques associées s'en déduisent.

Appelons P_r la probabilité que le système physique soit dans l'état $|r\rangle$. Nous avons vu qu'alors, la valeur moyenne de l'observable est donné par $\langle r|O|r\rangle$. La valeur moyenne \bar{O} de la mesure de l'observable O lorsque différents états peuvent être occupés est alors donnée par l'expression :

$$\bar{O} = \sum_r P_r \langle r|O|r\rangle.$$

Cette moyenne notée avec une barre sur la quantité moyennée est la *moyenne statistique*. À ce stade, il y a lieu de faire trois remarques importantes :

i) Si l'on dispose de la loi de probabilité, il est possible de calculer les valeurs moyennes. Dans le cas d'un système d'atomes, on pourra par exemple calculer l'énergie moyenne du système. Dans le cas d'un ensemble de moments magnétiques individuels, on pourra calculer l'aimantation résultante, etc.

ii) Non seulement on peut calculer la valeur moyenne de l'observable mais on peut aussi calculer les *fluctuations autour de cette valeur moyenne*. Il y a là quelque chose qui va bien au-delà de la notion d'état d'équilibre rencontrée en thermodynamique classique. Cela constitue l'un des apports majeurs de la physique statistique.

Remarque : Dans le cas d'une description classique, on peut définir la valeur moyenne d'une grandeur physique O de la même façon. C'est par exemple l'énergie cinétique du système qui est pour un état donné la somme des énergies cinétiques des atomes dans cet état. L'énergie cinétique moyenne du système est alors la moyenne sur tous les états possibles du système.

1.2.3 La notion d'ensemble statistique

La notion d'ensemble statistique permet par exemple d'introduire une probabilité associée à un état. Considérons un système constitué d'une enceinte de volume V contenant N atomes. Ce système est caractérisé macroscopiquement par la donnée du nombre N de molécules, de sa température et de son volume. D'un point de vue microscopique et classique, un état est caractérisé par la donnée de la position et de l'impulsion de tous les atomes. Afin de définir la loi de probabilité des états, on peut imaginer que l'on est capable de préparer un nombre très grand de réalisations du système qui sont *macroscopiquement identiques mais microscopiquement différentes*. Dans le cas d'un gaz dans une enceinte, les différentes réalisations du système ont toutes le même nombre d'atomes, le même volume et la même température. En revanche, les positions et les impulsions des atomes diffèrent. Imaginons que l'on puisse faire une mesure des positions et des impulsions de tous les atomes de chaque réalisation du système. Il est alors possible de construire un histogramme et d'en déduire une loi de probabilité. Cette collection imaginaire de réalisations du système est appelé *ensemble représentatif ou ensemble statistique*. La moyenne définie à l'aide de la densité de probabilité définie sur cet ensemble est appelée *moyenne d'ensemble* ou moyenne statistique. Il faut noter que le caractère *statistique* de la physique statistique tient à l'introduction de cette moyenne d'ensemble.

Il n'est nullement nécessaire de travailler sur un système comportant un grand nombre de particules pour pouvoir faire de la physique statistique. *Le caractère statistique procède de la méconnaissance de l'état dans lequel se trouve le système* et pas du grand nombre de particules. Il se traduit par l'introduction de la densité de probabilité définie sur l'ensemble des états du système étudié. Cette approche peut être mise en oeuvre si l'on travaille sur un système ne comportant que peu de particules. Nous verrons plus loin quel est l'impact d'un faible nombre de particules : les fluctuations des grandeurs physiques sont alors beaucoup plus grandes. C'est justement l'un des cas où une description statistique devient essentielle.

1.2.4 Système à l'équilibre

La notion d'équilibre correspond à une notion relativement intuitive. C'est une notion que l'on emprunte à la thermodynamique classique. Afin de mieux cerner cette notion, nous

allons partir d'un exemple simple. Lorsque l'on ouvre une bouteille de champagne rafraîchie, du gaz s'échappe de la bouteille immédiatement. Ensuite des bulles se forment et du gaz dissous dans le liquide s'échappe. Au bout de quelques heures, la bouteille aura atteint la température de la pièce environnante et la quantité de gaz dissoute aura diminuée. Au bout d'une journée ou deux, la quantité de gaz dissous cessera de décroître. On dit que l'on a atteint l'équilibre. La bouteille a échangé de la chaleur et de la matière avec son environnement jusqu'à atteindre une situation stationnaire. On dit alors que le système est à l'équilibre. On parle parfois d'équilibre dynamique pour souligner qu'au niveau microscopique, à chaque instant des molécules sont dissoutes tandis que d'autres s'échappent du liquide, mais de sorte que le nombre moyen de molécules de CO₂ dissoutes reste constant.

Sur cet exemple, nous observons deux caractéristiques de l'équilibre. D'une part, un état à l'équilibre est *stationnaire*, d'autre part, tous les *flux sont nuls* (de masse, d'énergie, etc). La stationnarité signifie que les variables macroscopiques qui définissent l'état du système n'évoluent pas dans le temps. Toutefois, la stationnarité ne suffit pas à définir l'état d'équilibre. Une autre caractéristique essentielle des états d'équilibre est l'absence de flux. Ainsi, si l'on plaçait la bouteille de champagne fermée sur une plaque chauffante, un flux de chaleur traverserait la bouteille et s'échapperait par les parois vers l'air environnant. Ceci ne serait pas un état d'équilibre bien que l'on puisse atteindre un état stationnaire. Notons qu'à ce flux de chaleur serait associé un gradient de température. On voit que l'on ne pourrait pas définir une température du système, mais un champ de température. Un état d'équilibre est un état pour lequel les variables intensives sont homogènes dans le système.

Une définition possible est : un système à l'équilibre est un système stationnaire pour lequel tous les flux sont nuls. Une définition alternative est : un système à l'équilibre est un système stationnaire caractérisé par des variables intensives homogènes.

1.2.5 Ergodicité

Afin de définir la densité de probabilité, on pourrait imaginer de réaliser des mesures de l'état du système toutes les secondes par exemple plutôt que de considérer un grand nombre de systèmes. Si le système évolue dans le temps, on peut penser qu'au bout d'un temps assez long, il est passé dans un grand nombre de configurations et que cela permet de réaliser une mesure de la densité de probabilité. Utiliser cette densité de probabilité revient à faire une moyenne temporelle. Se pose alors la question de la relation entre la moyenne temporelle et la moyenne d'ensemble. Lorsque la moyenne temporelle est égale à la moyenne d'ensemble, on dit que le système est ergodique. Tous les systèmes ne sont pas ergodiques.

La question de l'ergodicité est souvent mise en avant dans les présentations de la physique statistique car elle permet de répondre à la question de la validité d'une approche statistique lorsque l'on étudie un système particulier. Supposons que l'on s'intéresse à un ensemble de N atomes dans un volume V . Pourquoi des mesures effectuées sur un seul système seraient prédites par des prédictions effectuées sur un ensemble de réalisations de ce système ?

Une première réponse est de dire que le système évolue dans le temps et explore une grande partie des états accessibles de sorte que les mesures sont moyennées temporellement. Il en ressort que si le système est ergodique, la moyenne d'ensemble fournit la valeur moyenne temporelle. Bien que cet argument soit parfaitement raisonnable, il se trouve que les temps requis pour qu'un système explore effectivement une part significative des états accessibles est généralement hors de portée.

Il faut donc en conclure que la moyenne d'ensemble fournit une prédiction en accord

avec les mesures, cela est liée au fait que les grandeurs mesurées sont peu dépendantes des détails des réalisations microscopiques. Cet argument peut être justifié par un calcul direct des fluctuations des grandeurs physiques. Nous effectuerons ce calcul et verrons qu'effectivement, lorsque les systèmes contiennent un grand nombre de particules, les fluctuations autour des valeurs moyennes deviennent négligeables.

1.3 Comment compter les états ? L'approximation classique

Lorsque l'on est amené à effectuer concrètement un calcul, il faut commencer par identifier les variables qui permettent de caractériser le système (tous les nombres quantiques nécessaires par exemple). Dans un deuxième temps, il faut recenser tous les états possibles. La troisième étape consiste à attribuer une probabilité à chacun d'entre eux. Cela étant fait, il devient possible de calculer les valeurs moyennes de toutes les quantités physiques associées au système. Avant de passer au problème central qui est la détermination de la loi de probabilité, il est utile de se doter du moyen de compter les états. Lorsque l'on traite d'un système quantique, la donnée des nombres quantiques permet aisément de repérer et de compter les états. Dans le cas d'un système classique, le problème est plus délicat : il existe une infinité de points dans un espace des phases. On admettra ici un résultat qui est valable lorsque l'écart d'énergie entre deux états voisins est très petit devant $k_B T$ où k_B est la constante de Boltzmann. Cela signifie que le spectre d'énergie du système peut être considéré comme continu. On parle alors d'*approximation classique*. Dans ces conditions, on montre que le nombre d'états est égal au volume $d\Gamma$ de l'espace des phases divisé par $h^{N_{dl}}$ où N_{dl} est le nombre de degrés de liberté du système. Ainsi, pour un gaz contenant N atomes, il y a $3N$ degrés de liberté du système, c'est-à-dire qu'il y a $3N$ coordonnées d'espace auxquelles sont associées $3N$ coordonnées d'impulsion. Le nombre d'états $d\mathcal{N}$ est alors donné par :

$$d\mathcal{N} \approx \frac{d\Gamma}{h^{3N}}$$

où $3N$ est le nombre de degrés de liberté du système. On vérifiera en exercice qu'à la limite où l'énergie typique d'une particule est grande devant l'écart d'énergie entre deux niveaux consécutifs, cette relation est bien satisfaite. Cette approximation est aussi appelée *approximation des grands nombres quantiques ou approximation des grands volumes* suivant les auteurs. Lorsque l'approximation classique est valable, on représente les états par des points dans l'espace des phases. On est alors amené à définir une densité de probabilité sur l'espace des phases. Concrètement, l'opération de somme sur tous les états qui s'introduit naturellement dès lors que l'on veut calculer une moyenne comme on l'a vu plus haut est remplacée par une intégrale sur l'espace des phases :

$$\sum_r \longrightarrow \int_{\Gamma} \frac{d\Gamma}{h^{N_{dl}}} \quad \text{dans le cadre de l'approximation classique.}$$

Exploitions ce résultat dans le cas d'un système de N atomes. On obtient

$$\sum_r \longrightarrow \prod_{i=1}^N \left[\frac{\int dx_i dy_i dz_i \int dp_{x_i} dp_{y_i} dp_{z_i}}{h^3} \right]$$

Il est possible d'effectuer la somme sur les états r en regroupant tous les états de même énergie. Si le système contient plusieurs niveaux d'énergie E_n et de dégénérescence g_n , la

somme peut se mettre sous la forme

$$\sum_r \rightarrow \sum_n g_n$$

où r repère les états et n les niveaux d'énergie. Lorsque l'approximation classique est valable, la somme sur les niveaux d'énergie peut être remplacée par une intégrale. On note alors $g(E)dE$ le nombre d'états d'énergie comprise entre E et $E + dE$. La fonction $g(E)$ est appelé densité d'états.

$$\sum_r \rightarrow \sum_n g_n \xrightarrow{\text{App. class.}} \int g(E)dE$$

1.4 Le principe fondamental de la physique statistique

1.4.1 La loi de probabilité pour un système isolé. Ensemble microcanonique

Enoncé du principe fondamental

Nous abordons maintenant le problème central de la physique statistique. Comment associer à un état physique d'un système une probabilité ? Nous considérons pour débiter un *système qui est isolé* de sorte qu'il ne subit aucune influence. De ce fait, l'énergie du système est nécessairement fixée. Seuls les états du système qui ont cette énergie peuvent être réalisés physiquement. On parle alors d'*états accessibles* au système. Cette contrainte est la seule que l'on puisse imposer a priori. Mis à part cette contrainte, il n'existe aucune autre raison d'attribuer une probabilité plus ou moins importante à un état. Le faire reviendrait à privilégier un état et cela traduirait une certaine information sur le système. En l'absence de toute information supplémentaire sur le système, le seul choix raisonnable est de considérer que *tous les états accessibles sont équiprobables*. Cette hypothèse est érigée en postulat et constitue le postulat de base de la physique statistique :

Pour un système isolé à l'équilibre, tous les états accessibles sont équiprobables.

Ce choix est justifié a posteriori par le fait qu'il permet de construire une théorie en accord avec l'expérience. Lorsque l'on considère des systèmes isolés, les ensembles représentatifs associés sont appelés *ensembles microcanoniques*.

Discussion

Le but de ce paragraphe est de montrer que le principe fondamental joue un rôle analogue à celui du second principe de la thermodynamique. Le second principe est le principe qui permet de choisir quel est l'état d'équilibre d'un système parmi tous les états possibles lorsque l'on supprime une contrainte et que l'on laisse le système évoluer.

Au premier abord, le principe fondamental ne permet pas de choisir puisqu'il postule que tous les états sont équiprobables. Afin de clarifier cette question, il est essentiel de souligner que le principe fondamental de la physique statistique stipule que tous les états *microscopiques* sont équiprobables tandis que le second principe permet de choisir un état *macroscopique*. Essayons donc de répondre à la question de la thermodynamique (quel est l'état macroscopique observé ?) en utilisant le principe fondamental de la physique statistique (tous les états microscopiques sont équiprobables). Il suffit de compter combien

d'états microscopiques correspondent à chaque état macroscopique. C'est une élection à la majorité. Appelons Ω le nombre d'états microscopiques correspondants à un état macroscopique. L'état d'équilibre est celui pour lequel Ω est le plus grand. Comme il s'agit de nombres démesurément grands, il est plus adapté de manipuler $\ln \Omega$, c'est-à-dire l'entropie de Boltzmann.

Prenons un exemple simple. Le système considéré est une bouteille contenant N atomes dans un volume $2V$. La bouteille peut être divisée en deux compartiments de volumes égaux $V_1 = V$ et $V_2 = V$. Le principe fondamental de la physique statistique énonce qu'un état tel que les N atomes se trouvent dans V_1 a la même probabilité qu'un état tel que $N/2$ atomes sont dans V_1 et $N/2$ atomes dans V_2 . Comme on l'a indiqué plus haut, la différence entre les deux situations tient au nombre total d'états du type $(N_1 = N, N_2 = 0)$ et du type $(N_1 = N/2, N_2 = N/2)$. Une conséquence du postulat d'équiprobabilité est que la probabilité pour un atome d'être dans le volume V_1 est $1/2$. Il s'en suit que la probabilité de l'état macroscopique $(N_1 = N, N_2 = 0)$ est donnée par $1/2^N$ tandis que la probabilité d'un état macroscopique du type $(N_1 = N/2, N_2 = N/2)$ est donnée par $C_N^{N/2}/2^N$. Elle est beaucoup plus grande du fait du grand nombre de combinaisons possibles. o

1.4.2 Entropie en physique statistique

Définition de l'entropie statistique

Il est possible de donner une définition de l'entropie dans le cadre d'une théorie statistique. Sa forme a été introduite par L. von Boltzmann dans le cadre de ses études des gaz. Elle a depuis été reprise par Shannon qui a fondé la théorie de l'information. Elle a ensuite été appliquée dans l'étude de la totalité des systèmes stochastiques allant de la physique des systèmes macroscopiques au traitement d'images ou de signaux bruités en passant par l'étude des fluctuations de la bourse. L'expression de l'entropie qui est proposée par Boltzmann est la suivante :

$$S = -k_B \sum_r P_r \ln P_r.$$

où P_r est la probabilité que le système soit dans un état r . Le problème de Boltzmann était double : retrouver la loi de répartition des vitesses formulée par Maxwell d'une part, établir une expression statistique de l'entropie et montrer qu'elle ne peut que croître d'autre part.

Remarque : Dans le cas où il y a Ω états équiprobables, la probabilité est $P_r = 1/\Omega$ et l'on trouve :

$$S = k_B \ln \Omega$$

1.4.3 Une deuxième formulation du principe fondamental : le principe de l'entropie maximale

Enoncé

Il est possible de montrer que la seule loi de probabilité qui rende maximale l'entropie définie plus haut est la loi d'équiprobabilité. Il est possible de choisir une formulation différente du principe fondamental basée sur l'idée d'entropie maximale. On postule la forme de l'entropie statistique donnée ci-dessus. L'énoncé du postulat devient alors :

A l'équilibre, la loi de probabilité est la loi compatible avec les informations (contraintes) sur le système qui rend maximale l'entropie du système.

La loi de probabilité doit alors être cherchée de sorte que ce principe soit satisfait. Soulignons que cet énoncé du principe peut s'appliquer à un système qui n'est pas isolé.

Le point de vue de la théorie de l'information

Dans ce paragraphe, nous donnons un aperçu du point de vue de la théorie de l'information. L'avantage de ce point de vue est de fournir une construction raisonnée de la notion d'entropie statistique.

La question de base de la théorie de l'information est d'introduire une mesure de la quantité d'information d'un évènement. Pour préciser ce que nous entendons par évènement, nous prenons deux exemples. Tout d'abord, nous considérons un lancer de dés répété quatre fois. L'évènement est alors une liste ordonnée telle que $\{1, 6, 4, 2\}$. Un deuxième exemple est un système physique quantique qui peut occuper plusieurs états quantiques notés r . L'évènement est le résultat de la mesure qui fournit l'état r dans lequel se trouve le système. Le but de la discussion est d'illustrer l'idée que l'entropie est une mesure de la *quantité d'information manquante sur le système*. Considérons le cas d'un système physique pour lequel nous savons qu'il est dans l'état r_0 de sorte que $P_{r_0} = 1$ et $P_r = 0$ pour tous les autres états r . Dans ce cas, il n'y a pas d'information manquante. On peut constater que la formule de l'entropie statistique donne 0. Pour toute autre distribution de la loi des P_r , la formule montre que l'entropie est positive. Ce premier exemple montre que l'entropie croît avec l'information manquante. Ici, l'information manquante est le fait de ne pas savoir avec certitude quel est l'état dans lequel se trouve le système. Si le nombre d'états possibles augmente, alors l'entropie augmente.

Revenons maintenant en arrière et examinons le lien entre information et probabilité. Supposons que l'on effectue 600 lancers de dés et que l'on trouve que le résultat est toujours 1. C'est un résultat extrêmement peu probable. C'est si peu probable que si cela arrivait, vous penseriez que les dés sont pipés et vous demanderiez à examiner les dés de près. A contrario, si l'on trouve 100 fois le 1, 99 fois le 2, 101 fois le 3, etc, vous ne serez pas surpris et ne prêterez pas grande attention aux dés. Dans le premier cas, un évènement rare a attiré votre attention. L'évènement rare contient de l'information : le dé est pipé. Le second cas est très probable et contient peu d'information.

Prenons un second exemple. L'évènement est la liste des températures à Paris pendant un an. Si les températures sont dans la fourchette $[0^\circ C - 25^\circ C]$, les évènements sont probables et vous ne vous attendez pas à ce que les télévisions étrangères en fassent état. Si la liste contient $-40^\circ C$, c'est une valeur très peu probable. Il est vraisemblable que l'on en parlera dans la presse.

On voit donc que les évènements rares contiennent de l'information. L'information est ainsi une fonction décroissante de la probabilité. Introduisons une fonction $H(P_r)$ qui mesure la quantité d'information. Nous ajoutons maintenant une hypothèse et considérons le cas d'un évènement A constitué de deux évènements indépendants B et C. Par exemple, quatre jets de dés est le résultat de deux premiers jets de dés suivi de deux autres jets de dés. Comme les jets de dés sont indépendants, on peut écrire : $P(A) = P(B \text{ et } C) = P(B)P(C)$. Il s'en suit que $H[P(A)] = H[P(B)P(C)]$. Nous cherchons une fonction qui mesure l'information et qui soit additive de sorte que $H[P(B)P(C)] = H[P(B)] + H[P(C)]$. Ceci entraîne $H(P) = -K \ln(P)$ où K est une constante arbitraire positive. Le signe négatif provient du fait que la fonction doit être décroissante.

Nous avons donc trouvé que la mesure de la quantité d'information est donnée par $-K \ln P$ si l'on veut satisfaire aux deux contraintes : fonction décroissante de P_r et fonc-

tion additive. Il ne nous reste plus qu'à introduire la valeur moyenne de cette quantité et à l'appeler entropie :

$$S = -K \sum_r P_r \ln P_r.$$