CHAPITRE PREMIER

PRINCIPES FONDAMENTAUX DE LA STATISTIQUE

§ 1. Distribution statistique

La *physique statistique*, ou *statistique* tout court, a pour objet l'étude des lois particulières régissant le comportement et décrivant les propriétés des corps macroscopiques, i.e. des corps composés d'un nombre énorme de particules — atomes ou molécules. La forme générale de ces lois est pour une large part indépendante de la mécanique que l'on utilise pour décrire les mouvements individuels des particules des corps, i.e. de la mécanique classique ou de la mécanique quantique. La justification de ces lois se fonde cependant sur des raisonnements différents suivant qu'on utilise l'une ou l'autre de ces mécaniques. Pour raison de commodité nous commencerons par fonder tous nos raisonnements sur la mécanique classique.

Si le nombre d'équations de mouvement du système mécanique que l'on doit utiliser est égal au nombre des degrés de liberté, l'intégration de ces équations permettrait d'obtenir, en principe, toutes les données relatives au mouvement du système. Mais si le système régi par la mécanique classique comporte un nombre colossal de degrés de liberté, la mise en œuvre des méthodes de la mécanique conduit à écrire et à résoudre un nombre égal d'équations différentielles, ce qui est pratiquement irréalisable. Notons encore que même si on arrivait à effectuer l'intégration de ces équations, il serait absolument impossible de substituer dans la solution générale les conditions initiales relatives aux vitesses et aux coordonnées de toutes les particules.

Il semblerait à première vue que les propriétés du système mécanique deviendraient de plus en plus compliquées à mesure qu'augmenterait le nombre de particules et qu'on ne pourrait alors trouver aucune régularité dans le comportement des corps macroscopiques. Or il n'en est rien et nous verrons par la suite que lorsque le nombre de particules est très grand de nouvelles lois de caractère spécifique apparaissent.

Ce sont des lois dites *statistiques* qui apparaissent précisément lorsque le nombre de particules constituant les corps est très grand; ces lois sont telles qu'elles ne peuvent être ramenées à des lois pu ment mécaniques. La spécificité des lois statistiques réside en ce qu'elles perdent toute signification lorsqu'on passe à des systèmes mécaniques ne présentant qu'un petit nombre de degrés de liberté. Quoique le mouvement des systèmes possédant un nombre colossal de degrés de liberté obéisse aux mêmes lois de mécanique que celles qui régissent les systèmes composés d'un petit nombre de particules, l'existence d'un très grand nombre de degrés de liberté détermine l'apparition de lois qualitativement différentes.

Toute l'importance de la physique statistique parmi d'autres domaines de la physique théorique tient à ce que la Nature nous met constamment en présence des corps macroscopiques dont le comportement ne peut être complètement décrit, pour les raisons invoquées ci-dessus, par des méthodes purement mécaniques; les corps macroscopiques sont régis par des lois statistiques.

Avant de formuler l'objet principal de la statistique classique, il nous faut introduire d'abord la notion d'*espace des phases*, notion que nous aurons à utiliser constamment.

Posons que le système mécanique macroscopique considéré possède s degrés de liberté. Cela signifie que les positions spatiales des différents points du système sont définies par s coordonnées que nous noterons q_i, l'indice i prenant toutes les valeurs 1, 2, . . ., s. A un instant donné l'état du système sera déterminé par les s coordonnées q_i et par les s vitesses $\dot{q}_i$ correspondant à ces coordonnées. En statistique il est d'usage, pour raison de commodité, de caractériser les systèmes par les coordonnées et par les impulsions p_i et non par les vitesses. On peut représenter mathématiquement les différents états d'un système par des points dans un espace dit des phases (l'espace des phases est un concept purement mathématique) ; on porte sur les axes de coordonnées de cet espace les valeurs des coordonnées et des impulsions du système considéré. Chaque système possède donc son propre espace des phases, caractérisé par un nombre de dimensions égal au double du nombre de ses degrés de liberté. Tout point de l'espace des phases, étant défini par des valeurs déterminées des coordonnées q_i et des impulsions p_i du système, caractérise un état déterminé de celui-ci. Comme l'état du système varie dans le temps, le point de l'espace des phases qui représente un état donné du système (« point de phase ») y décrit une certaine ligne que l'on appelle trajectoire de phase.

Considérons un corps macroscopique quelconque ou un système du corps. Posons que le système considéré est fermé, i.e. n'interagit avec aucun corps. Délimitons en pensée, dans ce système, une région très petite devant le système tout entier, mais de dimension macroscopique. Si le système est composé d'un très grand nombre de particules, le nombre de particules contenues dans une petite partie du système peut être grand. Les parties relativement petites mais macroscopiques d'un système sont appelées *sous-systèmes*. Tout sous-système est encore un système mécanique, mais qui n'est pas fermé, étant soumis à toutes sortes d'actions de la part des autres parties du système. Le nombre des degrés de liberté de ces autres parties du système étant extrêmement grand, leurs interactions avec le sous-système considéré présentent un caractère très compliqué et imbriqué. De ce fait la variation dans le temps de l'état du sous-système sera elle aussi très compliquée.

On ne peut parvenir à une solution exacte du problème du comportement du sous-système qu'en résolvant le problème de mécanique correspondant concernant la totalité du système fermé, ce qui revient à dresser et à résoudre toutes les équations différentielles du mouvement pour des conditions initiales données; nous avons déjà signalé que c'est là un travail irréalisable. Or c'est précisément l'allure extrêmement compliquée de la variation de l'état des sous-systèmes, rendant inutilisables les méthodes de la mécanique, qui permet de tourner la difficulté et d'utiliser une approche différente.

Cette approche nouvelle est fondée sur le fait que les actions extérieures exercées par les autres parties du système sur le sous-système considéré étant très compliquées, au cours d'un intervalle de temps suffisamment long, le sous-système se retrouvera maintes fois dans tous les états dans lesquels il peut se trouver. On peut formuler cette proposition de manière plus précise. Désignons par $\Delta p \Delta q$ une petite région du « volume » de l'espace des phases du sous-système considéré, correspondant aux valeurs q_i et p_i de ses coordonnées et de ses impulsions comprises dans les intervalles Δq_i et Δp_i. On peut affirmer qu'au cours d'un intervalle de temps T suffisamment long la trajectoire de phase extrêmement compliquée repassera maintes fois par une telle région de l'espace des phases. Notons Δt la partie de l'intervalle de temps T pendant laquelle le sous-système « se trouvait » dans la région $\Delta p \Delta q$ de l'espace des phases [1]. En faisant croître indéfiniment le temps T, le rapport $\Delta t/T$ tend vers une limite

$$w = \lim_{T\to\infty} \frac{\Delta t}{T}. \tag{1,1}$$

Cette grandeur peut évidemment être considérée comme la probabilité de trouver le sous-système à un instant arbitraire dans la région $\Delta p \Delta q$ de l'espace des phases.

En passant à un élément infiniment petit du volume des phases [2]

$$dq\, dp = dq_1\, dq_2 \ldots dq_s\, dp_1\, dp_2 \ldots dp_s, \tag{1,2}$$

[1] Pour raison de concision on dit généralement que le système « se trouve dans la région $\Delta p\, \Delta q$ de l'espace des phases », en entendant par là que le système se trouve dans différents états représentés par des points de phase de la région considérée.

[2] Dans ce qui suit nous désignerons respectivement par dp et dq les produits des différentielles de toutes les impulsions et de toutes les coordonnées.

on peut introduire la probabilité dw des états représentés par des points contenus dans cet élément de volume, i.e. la probabilité pour que les coordonnées q_i et les impulsions p_i aient des valeurs comprises dans les intervalles infinitésimaux: q_i, $q_i + dq_i$ et p_i, $p_i + dp_i$. Cette probabilité dw se laisse exprimer par

$$dw = \rho\,(p_1, \ldots, p_s, q_1, \ldots, q_s)\, dp\, dq, \tag{1,3}$$

où $\rho\,(p_1, \ldots, p_s, q_1, \ldots, q_s)$ est une fonction de toutes les impulsions et de toutes les coordonnées (nous écrirons $\rho\,(p, q)$ ou même ρ tout court). La fonction ρ qui joue le rôle de « densité » de distribution de la probabilité dans l'espace des phases est appelée *fonction de distribution statistique* (ou fonction de distribution) du corps considéré. La fonction de distribution doit évidemment vérifier la *condition de normalisation*

$$\int \rho\, dp\, dq = 1 \tag{1,4}$$

(l'intégrale doit être étendue à tout l'espace des phases) qui ne fait qu'exprimer le fait que la somme des probabilités d'existence de tous les états est égale à l'unité.

En statistique l'assertion suivante présente une très grande importance. La distribution statistique d'un sous-système donné ne dépend pas de l'état initial d'une autre petite partie du système étudié, car l'influence de cet état initial sera effacée au bout d'un temps suffisant par celles des parties restantes du système, qui sont beaucoup plus étendues. La distribution statistique ne dépend pas non plus de l'état initial du sous-système considéré, puisque celui-ci passe au cours du temps par tous les états imaginables dont chacun peut être pris pour état initial. On peut donc trouver la distribution statistique d'une petite partie du système étudié sans avoir à résoudre le problème de mécanique, compte tenu des conditions initiales.

Le principal objectif de la statistique consiste à déterminer la distribution statistique de n'importe quel sous-système. En parlant de « petites parties » d'un système fermé on doit remarquer que les corps macroscopiques auxquels on a généralement affaire sont eux-mêmes de « petites parties » d'un grand système fermé constitué par les corps et par le milieu extérieur dans lequel ils se trouvent.

Si le problème est résolu et qu'on connaisse la distribution statistique pour le sous-système considéré, on peut calculer les probabilités des différentes valeurs que prennent n'importe quelles grandeurs physiques dépendant de l'état du sous-système (i.e. des valeurs de ses coordonnées q et de ses impulsions p). Nous pouvons aussi calculer la valeur moyenne d'une quantité quelconque $f\,(p, q)$ en multipliant ses valeurs possibles par les probabilités correspondantes et en intégrant ensuite sur tous les états. En dénotant les valeurs moyennes par une lettre surmontée d'une barre, on écrira

$$\overline{f} = \int f(p, q)\, \rho(p, q)\, dp\, dq. \tag{1,5}$$

Cette formule sert à calculer les valeurs moyennes des différentes grandeurs à l'aide de la fonction de distribution statistique [1].

Le calcul de la valeur moyenne à l'aide de la fonction de distribution (*moyenne statistique*) rend inutile l'étude des variations dans le temps de la valeur réelle de la grandeur physique $f(p, q)$ afin de calculer sa valeur moyenne. Compte tenu de la définition même du concept de probabilité exprimé par la formule (1,1), il est évident que la moyenne statistique est absolument équivalente à la moyenne temporelle. Donc, s'il en est ainsi, en suivant les variations dans le temps de la grandeur considérée, on pourrait construire la fonction $f = f(t)$; la moyenne cherchée serait alors donnée par

$$\overline{f} = \lim_{T\to\infty} \frac{1}{T} \int_0^T f(t)\, dt.$$

Il résulte de ces différentes considérations que les conclusions et les prévisions de l'étude statistique du comportement des corps macroscopiques présentent un caractère probabiliste. C'est ce qui distingue la statistique de la mécanique (classique), les résultats de cette dernière étant parfaitement univoques. Il importe cependant de remarquer que le caractère probabiliste des résultats que fournit la statistique classique n'est pas inhérent à la nature des objets à l'étude et tient uniquement à ce qu'on obtient ces résultats en utilisant beaucoup moins de données qu'il n'en faudrait pour une description mécanique exhaustive des objets (on n'a pas besoin de connaître les valeurs initiales de toutes les coordonnées et de toutes les impulsions).

En pratique la mise en œuvre de la statistique dans l'étude des corps macroscopiques ne laisse généralement pas apparaître le caractère probabiliste des résultats auxquels elle conduit. Si on observe pendant un temps suffisamment long un corps macroscopique (se trouvant dans des conditions stationnaires, i.e. des conditions extérieures ne dépendant pas du temps), on constate que toutes les grandeurs physiques caractérisant ce corps ont des valeurs pratiquement constantes (égales à leurs valeurs moyennes) et ne présentent que rarement des écarts notables à ces valeurs ; il s'agit évidemment des grandeurs *macroscopiques* caractérisant le corps tout entier ou ses différentes parties macroscopiques et non pas des particules indi-

[1] Dans ce livre nous dénoterons les valeurs moyennes soit par une lettre surmontée d'une barre $\overline{f}$, soit par des chevrons $\langle f \rangle$ en se guidant par la commodité de l'écriture. La deuxième notation convient mieux pour dénoter les valeurs moyennes d'expressions compliquées.

viduelles qui le composent [1]. Cette circonstance, essentielle pour la statistique, découle des considérations très générales (exposées au paragraphe suivant) et se trouve d'autant mieux vérifiée que le corps étudié est plus compliqué et plus grand. En termes de distribution statistique, on peut dire que si on construit, à l'aide de la fonction $\rho(p, q)$, la fonction de distribution de la quantité $f(p, q)$, cette dernière fonction présentera un pic très aigu pour $f = \bar{f}$ et ce n'est qu'à proximité immédiate de ce pic qu'elle aura une valeur différente de zéro.

Ainsi la statistique, qui permet de calculer les valeurs moyennes des grandeurs caractérisant les corps macroscopiques, permet de faire des prévisions de grande précision, valables pour la majeure partie de n'importe quel intervalle de temps, qui doit être cependant assez grand pour que l'influence de l'état initial du corps se trouve effacée. A ce point de vue la statistique fournit des prévisions déterminées et non pas probabilistes. (C'est pour cela que dans ce qui suit nous omettrons presque toujours de placer une barre au-dessus du symbole des grandeurs macroscopiques lorsqu'il s'agira de leurs valeurs moyennes).

Lorsqu'un système macroscopique fermé se trouve dans un état tel que les valeurs des grandeurs physiques macroscopiques caractérisant n'importe quelle partie macroscopique du système sont égales, avec un degré de précision relativement grand, à leurs valeurs moyennes, on dit que le système se trouve dans un état d'*équilibre statistique* (celui-ci est désigné aussi sous le nom d'*équilibre thermodynamique* ou *thermique*). Il résulte des considérations qui précèdent que, si on observe un système macroscopique fermé pendant un intervalle de temps suffisamment long, le système se trouve, pendant la majeure partie de ce temps, dans un état d'équilibre statistique. Si un système macroscopique fermé ne se trouvait pas, à un instant initial arbitraire, dans un état d'équilibre statistique (par exemple, a été artificiellement mis hors d'équilibre par une action extérieure, puis abandonné à lui-même, i.e. est redevenu un système fermé), il reviendra obligatoirement à un état d'équilibre au bout d'un certain temps. L'intervalle de temps au bout duquel un système revient à l'état d'équilibre statistique est appelé *temps de relaxation*. Lorsqu'il était question plus haut d'intervalles de temps « suffisamment longs », on avait en vue des temps longs devant le temps de relaxation.

---

[1] L'exemple suivant permet de constater la surprenante précision de la méthode statistique. Si on considère un élément de volume d'un gaz contenant, par exemple, 1/100 de molécule-gramme, on constate que l'écart relatif moyen de la valeur de l'énergie de cette masse de gaz ne représente que $\sim 10^{-11}$ par rapport à sa valeur moyenne. La probabilité de déceler (lors d'une seule observation) un écart relatif de l'ordre de 10^{-6} par exemple, est égale à un nombre aussi petit que $\sim 10^{-3 \cdot 10^{15}}$.

La théorie des processus liés au passage à l'état d'équilibre constitue la *cinétique*. Comme la statistique concerne l'étude de systèmes se trouvant dans un état d'équilibre statistique, l'étude de la cinétique n'en fait pas partie.

§ 2. Indépendance statistique

Les sous-systèmes dont il a été question au § 1 ne sont pas des systèmes fermés. Bien au contraire, ils sont constamment soumis aux actions exercées par les autres parties du système considéré. Néanmoins, du fait que ces différentes parties, quoique petites vis-à-vis du système tout entier, sont des corps macroscopiques, on peut admettre qu'ils se comportent approximativement comme des systèmes fermés si le temps d'observation n'est pas très grand. En effet, aux interactions d'un sous-système donné avec les parties du système qui l'entourent, participent surtout les particules se trouvant à proximité de sa surface. Or la quantité relative de ces particules est petite devant le nombre total de particules contenues dans le sous-système et diminue rapidement lorsque les dimensions du sous-système augmentent ; lorsque ce dernier devient suffisamment grand, l'énergie de ses interactions avec l'entourage devient petite devant son énergie interne. On peut donc dire que les sous-systèmes sont *quasi fermés*. Soulignons encore une fois que le caractère quasi fermé des sous-systèmes ne subsiste que pendant des intervalles de temps pas trop longs. Si on considère un intervalle de temps long, l'influence des interactions mutuelles des sous-systèmes finira toujours par se manifester, aussi faibles que puissent être les interactions. Ce sont précisément ces interactions relativement faibles qui déterminent, en fin de compte, l'instauration de l'équilibre statistique.

Le fait qu'on puisse admettre que les interactions mutuelles des différents sous-systèmes sont faibles conduit à admettre que ces derniers sont statistiquement indépendants. Dire que les sous-systèmes sont *statistiquement indépendants* signifie que l'état d'un sous-système donné n'exerce aucune influence sur les probabilités d'apparition des différents états dans les autres sous-systèmes.

Soient, par exemple, deux sous-systèmes et soient $dp^{(1)}\, dq^{(1)}$ et $dp^{(2)}\, dq^{(2)}$ des éléments de volume de leurs espaces des phases. Si on considère les deux sous-systèmes comme un sous-système composé, au point de vue mathématique l'indépendance statistique des sous-systèmes signifie que la probabilité pour que le sous-système composé se trouve dans un élément de volume $dp^{(12)}\, dq^{(12)} = dp^{(1)}\, dq^{(1)} \cdot dp^{(2)}\, dq^{(2)}$ de son espace des phases est égale au produit des probabilités pour que les sous-systèmes initiaux se trouvent respectivement dans les éléments de volume $dp^{(1)}\, dq^{(1)}$ et $dp^{(2)}\, dq^{(2)}$, chacune de ces dernières probabilités ne dépendant que des coordonnées et des impulsions du sous-système donné. On peut donc écrire

$$\rho_{12}dp^{(12)}\,dq^{(12)} = \rho_1\,dp^{(1)}dq^{(1)}\cdot\rho_2\,dp^{(2)}\,dq^{(2)},$$

ou

$$\rho_{12} = \rho_1\rho_2, \tag{2,1}$$

où ρ_{12} est la distribution statistique du sous-système composé, et ρ_1 et ρ_2 sont les fonctions de distribution des sous-systèmes initiaux. On peut écrire une relation analogue pour un ensemble de plusieurs sous-systèmes [1].

La proposition inverse est évidemment valable: si la distribution des probabilités pour un système composé se laisse décomposer en un produit de facteurs qui ne dépendent chacun que de grandeurs caractérisant une des parties du système, cela signifie que ces parties sont statistiquement indépendantes et que chacun des facteurs est proportionnel à la probabilité des états de la partie correspondante.

Si f_1 et f_2 désignent deux grandeurs physiques relatives à deux sous-systèmes différents, il résulte de la relation (2,1) et de la définition (1,5) des valeurs moyennes que la valeur moyenne du produit f_1f_2 est égale au produit des valeurs moyennes de f_1 et f_2 prises isolément:

$$\overline{f_1f_2} = \overline{f_1}\cdot\overline{f_2}. \tag{2,2}$$

Considérons une grandeur quelconque f caractérisant un certain corps macroscopique ou une de ses parties. Au cours du temps la valeur de cette grandeur varie en oscillant autour de sa valeur moyenne. Introduisons une quantité caractérisant en moyenne la largeur de l'intervalle de variation de la grandeur f. On ne peut utiliser pour cela la valeur moyenne de la différence $\Delta f = f - \overline{f}$, puisque f présente des écarts de part et d'autre de la valeur moyenne, de sorte que la valeur moyenne de la différence $f - \overline{f}$, étant tantôt positive et tantôt négative, sera égale à zéro, quel que soit le nombre de fois que f se sera écarté de façon notable de sa valeur moyenne. Aussi utilisera-t-on, pour caractériser l'intervalle de variation de la grandeur f, la valeur moyenne du carré de la différence $f - \overline{f}$. Comme la quantité $(\Delta f)^2$ est toujours positive, sa valeur moyenne ne tend vers zéro que si elle-même tend vers zéro, autrement dit elle ne sera petite que si la probabilité d'apparition d'écarts importants de f par rapport à $\overline{f}$ est petite. La quantité $\langle(\Delta f)^2\rangle^{1/2}$ est appelée *fluctuation quadratique moyenne* de la grandeur f. En développant le carré $(f - \overline{f})^2$ on obtient

$$\langle(\Delta f)^2\rangle = \overline{f^2} - \overline{f}^2, \tag{2,3}$$

[1] A condition, bien entendu, que l'ensemble de plusieurs sous-systèmes constitue encore une petite partie du système fermé.