

Chapitre IV

Mécanique
statistique

§ 16 MESURES SUR UNE VARIÉTÉ

VARIÉTÉS COMPOSÉES

Nous allons donner, dans ce chapitre, une plus grande extension à la notion de *variété* ; ceci, par les conventions suivantes :

- (16.1) [Nous noterons R^0 l'ensemble composé du seul nombre 0 ;
Les deux parties de R^0 (R^0 et la partie vide) seront dites *ouvertes* ;
Une application A de R^n à R^0 sera dite *différentiable* si elle est définie sur un ouvert ; si $x \in \text{def}(A)$, on définira $D(A)(x)$ comme l'application (unique) de R^n sur R^0 ;
Toute application A de R^0 à R^n sera dite différentiable ; si $x \in \text{def}(A)$, on notera $D(A)(x)$ l'application nulle de R^0 à R^n .

- (16.2) [Soit V un ensemble ; F une application.
On dira que F est un *système de coordonnées* de V s'il existe un entier n ($n = 0, 1, 2, \dots$) tel que :
 F est une application de R^n à V ;
 F est régulier ;
 $\text{def}(F)$ est un ouvert de R^n .
Il est clair que l'entier n est complètement défini par F ; on l'appelle *dimension* du système de coordonnées F .

- (16.3) Avec ces conventions, les définitions suivantes du chapitre I peuvent encore s'appliquer, avec un sens élargi :

- *systèmes de coordonnées cohérents* (1.5) ;
- *atlas* (1.6) ;
- *cartes* (1.7) ;

on appellera encore *variété* tout ensemble V muni d'un atlas — la structure de variété étant définie par l'ensemble de toutes les cartes (qui est encore un atlas) — on pourra aussi étendre les notions suivantes :

- *ouvert* d'une variété (1.13),
- application *continue* (resp. *différentiable*) de variété à variété (1.15) (resp. (1.18)),
- ouverts *connexes*, *composantes* (1.48), (1.49).

Mais il faut modifier la notion de *dimension* ; on voit en effet que :

(16.4) Si V est une variété, x un point de V , toutes les cartes F telles que $x \in \text{val}(F)$ ont même dimension (16.2) n ; ce nombre s'appelle *dimension* de V au point x .

(16.5) On dit qu'une variété V est *pure* si sa dimension n est la même en tous les points ; n s'appelle alors « dimension de V ».

(16.6) — Les variétés qui ont été étudiées jusqu'ici sont donc les variétés *pures* de dimension ≥ 1 .

(16.7) — Si V est une variété *pure* de dimension 0, toute partie de V est ouverte ⁽¹⁾.

(16.8) — Une variété qui n'est pas pure sera dite *composée* ; ce qui s'explique par le théorème suivant :

(16.9) Soit V une variété ; pour tout entier $n \geq 0$, désignons par V_n l'ensemble (éventuellement vide) des points x de V où la dimension est égale à n . Alors :

V_n est un ouvert de V , et possède (au sens (1.39)) une structure de variété pure de dimension n .

— On vérifie, en effet, que l'ensemble des cartes de V dont la dimension est n est un atlas de V_n , et lui confère donc une structure de variété pure de dimension n .

(16.10) On voit que les V_n forment une *partition* de V en *ouverts disjoints* ; il en résulte immédiatement que les composantes connexes de V (1.49) sont contenues chacune dans l'un des V_n ; une variété composée a donc au moins deux composantes.

— Réciproquement, si on s'est donné des variétés pures V_j , de dimensions respectives n_j , deux à deux sans point commun, on donne à leur réunion V une structure de variété en appelant carte de V toute carte de l'une des V_j ; V sera une variété composée si les n_j ne sont pas tous égaux.

(16.11) Ces résultats ramènent donc l'étude des variétés au sens généralisé (16.3) à celle des variétés pures (16.6), (16.7) ; la nouvelle définition — que nous adopterons désormais — n'apporte donc pas de résultats mathématiques nouveaux, mais des facilités de langage pour l'étude qui va suivre.

⁽¹⁾ On exprime ce fait en disant que V est un espace topologique *discret*.

ENSEMBLES COMPACTS

DÉFINITION, THÉORÈME

Soit V une variété ; Φ une partie de V .

- (16.12) — On dit que Φ est *fermé* si le *complémentaire* de Φ dans V est *ouvert*.
— V et la partie vide de V sont fermés.
— La réunion de deux parties fermées de V est fermée.
— Toute intersection de parties fermées de V est fermée.

La vérification est immédiate (voir (1.14)).

DÉFINITION

(16.13) On appelle *fermeture* ⁽¹⁾ d'une partie Φ_0 d'une variété V l'intersection des parties fermées de V qui contiennent Φ_0 .

(16.14) Il résulte immédiatement de (16.12) que la fermeture de Φ_0 est *le plus petit ensemble fermé contenant* Φ_0 .

— Dans le cas $V = R^n$, la fermeture Φ d'un ensemble Φ_0 peut se caractériser comme suit

(16.15) $[x \in \Phi] \Leftrightarrow$ [toute boule de centre x rencontre Φ_0].

DÉFINITION

Soit V une variété *séparée* ; K une partie de V .

— On appelle *recouvrement* de K une famille d'ouverts Ω_j ($j \in I$) ⁽²⁾ dont la réunion contient K :

(16.16)
$$\forall j \in I : \Omega_j \subset V, \Omega_j \text{ ouvert} \quad \left| \quad K \subset \bigcup_{j \in I} \Omega_j \right.$$

— On dit que K est *compact* si tout recouvrement de K contient un recouvrement fini.

(16.17) Il est clair que toute partie *finie* de V est compacte ; que la réunion de deux parties compactes de V est compacte.

(16.18) Une variété V (considérée comme partie d'elle-même) peut être *compacte* (exemple : une sphère S_n) ou *non* (exemple : un espace vectoriel).

⁽¹⁾ Synonyme de fermeture : *adhérence*.

⁽²⁾ Rappelons que l'ensemble d'indices I est quelconque, donc pas nécessairement fini ou dénombrable. Il est clair que toute partie de V possède un recouvrement : l'ensemble des ouverts, indexés par eux-mêmes.

On montre les théorèmes suivants :

(16.19) [Si V est une variété séparée, K une partie compacte de V , K est fermée.

(16.20) [Soit K une partie de R^n ; alors :
 $[K \text{ compact}] \Leftrightarrow [K \text{ fermé et borné}]^{(1)}$ (Borel-Lebesgue).

(16.21) [Soient V_1, \dots, V_n des variétés séparées; si on a
 $K_j = \text{partie compacte de } V_j \quad \forall j \in [1, 2, \dots, n]$
 le produit direct $K = K_1 \times K_2 \times \dots \times K_n$ est une partie compacte de la variété produit $V = V_1 \times \dots \times V_n$.

(16.22) [Soient V et V' deux variétés séparées, K une partie compacte de V , A une application continue de K dans V' (Définition (1.15)).
 Alors l'ensemble de valeurs de A est une partie compacte de V' .
 Si de plus A est injectif, A^{-1} est continue.

En appliquant ce théorème au cas $V' = R$, on voit facilement (grâce à (16.20)) que toute fonction réelle continue sur un compact K est bornée supérieurement et inférieurement

(16.23) $\exists a, b \in R \quad [a \leq f(x) \leq b \quad \forall x \in K]$

et qu'elle atteint effectivement ses bornes :

(16.24) $\exists x_0, x_1 \in K; \quad [f(x_0) \leq f(x) \leq f(x_1) \quad \forall x \in K].$

(16.25) [— Une partie K d'une variété séparée V est dite *relativement compacte* si la fermeture de K est compacte $^{(2)}$;
 Toute partie d'un ensemble relativement compact est relativement compact; toute variété séparée V possède un recouvrement Ω_j tel que les Ω_j soient *relativement compacts* $^{(3)}$.

DÉFINITION, THÉORÈME

(16.26) [Soient V et V' deux variétés séparées.
 Nous dirons qu'une application A de V à V' est *étalée* si l'image réciproque par A de tout ouvert relativement compact de V' est un ouvert relativement compact de V .

Si A est étalée de V à V' , $\text{def}(A)$ est un ouvert de V ; A est continue.
 Si A est étalée de V à V' , B étalée de V' à V'' , $B.A$ est étalée de V à V'' .

Vérification immédiate.

$^{(1)}$ On dit qu'une partie de R^n est *bornée* si elle est contenue dans une boule.

$^{(2)}$ Exemple : une partie de R^n est relativement compacte si elle est bornée (voir (16.20)).

$^{(3)}$ On énonce ce théorème en disant que V est *localement compacte*; diverses notions que nous allons étudier sur les variétés séparées peuvent s'étendre à tous les *espaces topologiques localement compacts* (voir Bourbaki, *Intégration*; Naimark : *Normed Rings*).

ESPACES DE RIESZ

DÉFINITION

On dit que E est un *espace de Riesz* si :

- 1) E est un espace vectoriel réel.
- 2) On a défini dans E une *relation d'ordre* \geq $^{(1)}$:

$$(16.27) \quad \diamond \quad \begin{cases} [X \geq Y, Y \geq Z] \Rightarrow [X \geq Z] \\ [X \geq Y, Y \geq X] \Leftrightarrow [X = Y] \end{cases}$$

- 3) On a

$$\begin{cases} [X \geq Y] \Leftrightarrow [X - Y \geq 0] \\ [X \geq 0, s \geq 0 \text{ (s réel)}] \Rightarrow [sX \geq 0] \text{ }^{(2)}. \end{cases}$$

- 4) Tout couple X, Y d'éléments de E possède une *borne supérieure* $\text{sup}(X, Y)$, élément de E tel que

$$\clubsuit \quad \begin{cases} [Z \geq X] \\ [Z \geq Y] \end{cases} \Leftrightarrow [[Z \geq \text{sup}(X, Y)]]$$

(16.28) — On vérifie facilement que la borne supérieure $\text{sup}(X, Y)$ est *entièrement définie* par la relation \clubsuit .

— Dans tout espace de Riesz, on vérifie immédiatement les relations

$$(16.29) \quad \begin{cases} [X \geq Y] \Leftrightarrow [-Y \geq -X] \\ [X \geq 0, Y \geq 0] \Rightarrow [X + Y \geq 0]. \end{cases}$$

Il en résulte que tout couple X, Y possède une *borne inférieure* $\text{inf}(X, Y)$, caractérisée par $^{(3)}$

$$(16.30) \quad \begin{cases} [Z \leq X] \\ [Z \leq Y] \end{cases} \Leftrightarrow [Z \leq \text{inf}(X, Y)]$$

et que l'on obtient par la formule

$$(16.31) \quad \boxed{\text{inf}(X, Y) = - \text{sup}(-X, -Y)}$$

$^{(1)}$ Une relation s'appelle *relation d'ordre* si elle vérifie \diamond ; on n'exige pas que tout couple X, Y vérifie l'une des relations $X \geq Y, Y \geq X$.

$^{(2)}$ On dit que X est *positif* si $X \geq 0$.

$^{(3)}$ On écrit bien entendu $Z \leq X$ pour $X \geq Z$.

Si X est un élément d'un espace de Riesz, on pose :

$$(16.32) \quad \boxed{X^+ = \sup(X, 0) \quad X^- = \sup(-X, 0) \quad |X| = \sup(X, -X)} \quad (1)$$

on établit aisément les formules :

$$(16.33) \quad \boxed{X \equiv X^+ - X^- \quad |X| \equiv X^+ + X^-}$$

$$(16.34) \quad \boxed{|sX| = |s| |X|} \quad \forall s \in \mathbb{R}$$

$$(16.35) \quad \boxed{|X+Y| \leq |X| + |Y|}$$

$$(16.36) \quad \boxed{||X| - |Y|| \leq |X - Y|}$$

$$(16.37) \quad \boxed{\sup(X, Y) = \frac{1}{2} [X + Y + |X - Y|]}$$

DÉFINITION, THÉORÈME

Soit E un espace de Riesz, E' un sous-espace vectoriel de E .

(16.38) 1) Les conditions suivantes sont équivalentes :

a) $[X, Y \in E'] \Rightarrow [\sup(X, Y) \in E']$.

b) $[X \in E'] \Rightarrow [X^+ \in E']$.

c) $[X \in E'] \Rightarrow [|X| \in E']$.

2) Lorsqu'elles sont réalisées, nous dirons que E' est un *sous-espace de Riesz* de E ; E' est alors un espace de Riesz pour les opérations induites.

Vérification immédiate, à l'aide des formules (16.32), (16.37).

DÉFINITION, THÉORÈME

(16.39) Soit E un espace de Riesz; un élément Φ du dual E^* de E (2) sera dit *positif* si

$$\diamond \quad [X \geq 0] \Rightarrow [\Phi(X) \geq 0]$$

(1) Ces formules s'appliquent en particulier à l'espace de Riesz $E = \mathbb{R}$; mais il faut se souvenir, dans le cas général, que $|X|$ (« valeur absolue de X ») désigne un élément de E , non de \mathbb{R} .

(2) E^* est par définition l'ensemble des applications linéaires de E dans \mathbb{R} , appelées aussi *formes linéaires* de E .

(16.39) on appellera *mesure* de E les éléments de E^* qui sont *différence de deux éléments positifs*.

L'ensemble E' des mesures de E est un *sous-espace vectoriel* de E^* ; E' est lui-même un *espace de Riesz*, tel que

$$\heartsuit \quad \left[\begin{array}{l} \Phi \in E^* \\ \Phi \text{ positif} \end{array} \right] \Leftrightarrow \left[\begin{array}{l} \Phi \in E' \\ \Phi \geq 0 \end{array} \right]$$

Le lecteur trouvera la démonstration dans N. Bourbaki (*Intégration*), ainsi que celle de la formule

$$(16.40) \quad \boxed{X \geq 0 \Rightarrow |\Phi|(X) = \sup_{|Y| \leq X} \Phi(Y)} \quad \forall \Phi \in E'$$

THÉORÈME

$$(16.41) \quad \boxed{|\Phi| \leq \Phi', \quad |X| \leq X'} \Rightarrow \boxed{|\Phi(X)| \leq \Phi'(X')}$$

On a d'une part

$$\Phi'(X') - |\Phi|(|X|) = [\Phi' - |\Phi|](X') + |\Phi|(X' - |X|) \geq 0;$$

d'autre part :

$$|\Phi(X)| = |[\Phi^+ - \Phi^-](X^+ - X^-)| = |[\Phi^+(X^+) + \Phi^-(X^-)]$$

$$- [\Phi^+(X^-) + \Phi^-(X^+)]| \leq [\Phi^+ + \Phi^-](X^+ + X^-) = |\Phi|(|X|).$$

C.Q.F.D.

MESURES

DÉFINITIONS

(16.42) Soit V une variété séparée. Nous désignerons par $\mathcal{C}(V)$ l'ensemble des applications *continues* de V dans \mathbb{R} (« fonctions continues sur V »). Si $f \in \mathcal{C}(V)$, on appelle *support* de f la *fermeture* de l'ensemble des points x tels que $f(x) \neq 0$.

Nous appellerons $\mathcal{K}(V)$ l'ensemble des éléments de $\mathcal{C}(V)$ dont le support est *compact*.

THÉORÈME

(16.43) $\mathcal{C}(V)$ est un espace vectoriel; si l'on pose $[\forall f, g \in \mathcal{C}(V)]$

$$[f \geq g] \Leftrightarrow [f(x) \geq g(x) \quad \forall x \in V]$$

$\mathcal{C}(V)$ devient un *espace de Riesz*.

La vérification est immédiate, et conduit aux formules suivantes :

$$(16.44) \quad f^+(x) = \begin{cases} f(x) & \text{si } f(x) \geq 0 \\ 0 & \text{si } f(x) \leq 0 \end{cases} \quad f^-(x) = \begin{cases} -f(x) & \text{si } f(x) \leq 0 \\ 0 & \text{si } f(x) \geq 0 \end{cases}$$

$$(16.45) \quad |f|(x) = |f(x)| \quad \forall x$$

$$(16.46) \quad \sup(f, g)(x) = \sup(f(x), g(x)) = \frac{1}{2}[f(x) + g(x) + |f(x) - g(x)|] \quad \forall x$$

— Comme le support de $|f|$ est égal à celui de f , on vérifie immédiatement (grâce à (16.38)) que

$$(16.47) \quad \mathcal{K}(V) \text{ est un sous-espace de Riesz de } \mathcal{C}(V).$$

On démontre le lemme suivant (Urysohn) :

$$(16.48) \quad \text{Soit } K \text{ un compact d'une variété séparée } V. \text{ Il existe } f \in \mathcal{K}(V) \text{ tel que} \\ [0 \leq f(x) \leq 1 \quad \forall x \in V]; [f(x) = 1 \quad \forall x \in K]$$

et le résultat plus complet suivant ⁽¹⁾ :

Soit Ω_j un recouvrement fini d'un compact K ; il existe des f_j différentiables sur V tels que

$$(16.49) \quad \begin{array}{l} f_j \geq 0 \quad \forall j \\ \text{support}(f_j) \subset \Omega_j \quad \forall j \\ \sum_j f_j(x) \leq 1 \\ \left[\sum_j f_j(x) = 1 \quad \forall x \in K \right] \end{array}$$

DÉFINITION

$$(16.50) \quad \text{On appelle mesure d'une variété } V \text{ (supposée séparée) toute mesure de l'espace de Riesz } \mathcal{K}(V) \text{ (Définition (16.39)).}$$

En d'autres termes, une mesure μ sur V est une forme linéaire sur l'espace vectoriel des fonctions continues à support compact de V , qui est soit positive, soit différence de deux mesures positives.

On démontre le théorème ⁽²⁾.

⁽¹⁾ Voir Bourbaki, *Intégration*, chap. III; L. Schwartz, *Théorie des distributions* (Chap. I).
⁽²⁾ Voir Bourbaki, *Intégration*, Chap. III.

Soit V une variété séparée; μ une forme linéaire sur $\mathcal{K}(V)$. Les propositions suivantes sont équivalentes :

$$(16.51) \quad \begin{array}{l} 1) \mu \text{ est une mesure.} \\ 2) \text{ Pour tout compact } K, \text{ il existe un nombre } M \text{ tel que} \end{array}$$

$$\diamond \quad |\mu(f)| \leq M \|f\| \quad \text{si } f \in \mathcal{K}(V), \text{ support}(f) \subset K \quad (1)$$

qui fournit donc une nouvelle définition des mesures de V .

$$(16.52) \quad \text{— Les mesures de } V \text{ forment un espace de Riesz (Théorème (16.39)); à toute mesure } \mu \text{ on peut donc associer les mesures positives } \mu^+, \mu^-, |\mu|; \text{ on peut appliquer les formules (16.40), (16.41); soit :}$$

$$(16.53) \quad [f \geq 0] \Rightarrow [|\mu|(f) = \sup_{|g| \leq f} \mu(g)] \quad f, g \in \mathcal{K}(V)$$

$$(16.54) \quad [|\mu| \leq \mu', |f| \leq f'] \Rightarrow |\mu(f)| \leq \mu'(f')$$

— Si $f, g \in \mathcal{C}(V)$, nous désignerons par $f \times g$ leur produit ordinaire

$$(16.55) \quad [f \times g](x) = f(x)g(x) \quad \forall x \in V.$$

Il est clair que $f \times g \in \mathcal{C}(V)$; cette opération \times donne à $\mathcal{C}(V)$ une structure d'algèbre commutative ⁽²⁾.

— Soit μ une mesure de V , f une fonction continue sur V ; nous poserons

$$(16.56) \quad [\mu \times f](g) = \mu(f \times g) \quad \forall g \in \mathcal{K}(V).$$

$$(16.57) \quad \text{On vérifie aisément que } \mu \times f \text{ est encore une mesure, appelée produit de la mesure } \mu \text{ par la fonction continue } f; \mu \times f \text{ est positive si } \mu \text{ et } f \text{ sont positifs; l'application } (\mu, f) \mapsto \mu \times f \text{ est bilinéaire et associative (à droite) on dit que l'algèbre } \mathcal{C}(V) \text{ opère sur l'espace des mesures de } V.$$

Théorème (notations (16.56)).

$$(16.58) \quad |\mu \times f| = |\mu| \times |f|$$

Utilisons la formule (16.53); on est ramené à montrer, pour tout $g \geq 0$ dans $\mathcal{K}(V)$, que les nombres

$$X = \sup_{k \in I} \mu(k), \quad X' = \sup_{k' \in I'} \mu(k')$$

⁽¹⁾ La notation $\|f\|$ est définie ci-dessous en (16.110).
⁽²⁾ C'est-à-dire : l'opération \times est bilinéaire, associative (et commutative).

sont égaux, les ensembles I et I' étant donnés par

$$[k \in I] \Leftrightarrow [\exists h \in \mathcal{X}(V), |h| \leq g, k = f \times h]$$

$$[k' \in I'] \Leftrightarrow [k' \in \mathcal{X}(V), |k'| \leq |f|g].$$

— Il est clair que $I \subset I'$, donc que $X \leq X'$.

— Soit d'autre part K le support de g , M le nombre associé à μ et K par la formule (16.51 \diamond)

$$\heartsuit [l \in \mathcal{X}(V), \text{support}(l) \subset K, |l(x)| \leq \lambda \quad \forall x] \Rightarrow [|\mu(l)| \leq M\lambda].$$

Si ε est un nombre positif, désignons par φ la fonction continue dont le graphe, donné par la figure (16.I), est composé d'un segment de droite et d'un morceau d'hyperbole; pour tout $k' \in I'$, posons

$$h(x) = k'(x) \varphi(f(x)), \quad k = f \times h, \quad l = k' - k.$$

On vérifie facilement que $h \in \mathcal{X}(V)$, $|h| \leq g$, donc que $k \in I$; par suite $|\mu(k)| \leq X$; d'autre part que

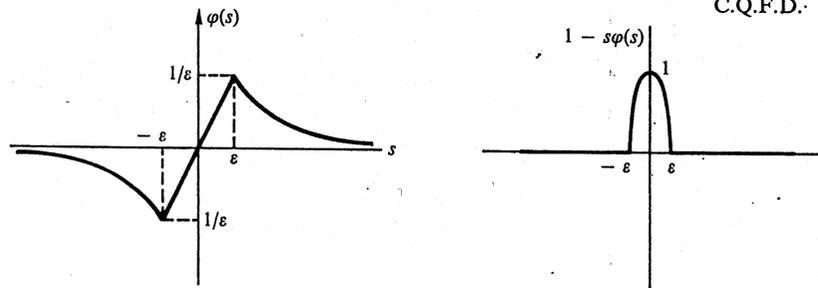
$$[l(x) \neq 0] \Rightarrow [1 - f(x) \varphi(f(x)) \neq 0] \Rightarrow [|f(x)| \leq \varepsilon] \Rightarrow$$

$$\Rightarrow |l(x)| \leq |k'(x)| \leq |f(x)|g(x) < \varepsilon \|g\| \text{ (notation (16.110 } \diamond \text{))};$$

comme $l \in \mathcal{X}(V)$, et $\text{support}(l) \subset K$, il résulte de \heartsuit que $|\mu(l)| \leq \varepsilon M \|g\|$; d'où $\mu(k') = \mu(k) + \mu(l) \leq X + \varepsilon M g$; $X' \leq X + \varepsilon M \|g\|$; ε étant arbitraire, on a donc $X' = X$.

C.Q.F.D.

Fig. 16.I.



PRODUIT TENSORIEL DE MESURES

THÉORÈME

Soient V et V' deux variétés séparées; μ et μ' des mesures de V et V' respectivement.

(16.59) Il existe une mesure de $V \times V'$, appelée *produit (tensoriel) de μ et μ'* , notée $\mu \otimes \mu'$, définie par

$$\diamond [\mu \otimes \mu'](f) = \mu \left(x \mapsto \mu' \left(x' \mapsto f \left(\begin{matrix} x \\ x' \end{matrix} \right) \right) \right) \quad \forall f \in \mathcal{X}(V \times V').$$

— Ce théorème contient implicitement les résultats suivants : $V \times V'$ est une variété séparée; si $f \in \mathcal{X}(V \times V')$, et si $x \in V$, l'application $x' \mapsto f \left(\begin{matrix} x \\ x' \end{matrix} \right)$ appartient à $\mathcal{X}(V')$; de même l'application

$$x \mapsto \mu' \left(x' \mapsto f \left(\begin{matrix} x \\ x' \end{matrix} \right) \right)$$

appartient à $\mathcal{X}(V)$. Le lecteur trouvera l'ensemble des démonstrations dans Bourbaki.

THÉORÈME

$$(16.60) \quad [\mu \geq 0, \mu' \geq 0] \Rightarrow [\mu \otimes \mu' \geq 0].$$

Trivial.

Un simple calcul donne le théorème suivant :

$$(16.61) \quad \text{L'opération } \otimes \text{ sur les mesures est bi-linéaire et associative.}$$

La vérification de l'associativité conduit d'ailleurs à la formule

$$(16.62) \quad [\mu \otimes \mu' \otimes \mu''](f) = \mu \left(x \mapsto \mu' \left(x' \mapsto \mu'' \left(x'' \mapsto f \left(\begin{matrix} x \\ x' \\ x'' \end{matrix} \right) \right) \right) \right)$$

— Si $g \in \mathcal{C}(V)$, $g' \in \mathcal{C}(V')$, nous poserons

$$(16.63) \quad [g \otimes g'] \left(\begin{matrix} x \\ x' \end{matrix} \right) \equiv g(x) g'(x')$$

il est immédiat que

$$(16.64) \quad [\mu \otimes \mu'](g \otimes g') = \mu(g) \mu'(g') \quad \forall g, g' \in \mathcal{X}(V)$$

pour toutes mesures μ et μ' de V et V' ; réciproquement, on démontre (voir Bourbaki) le théorème :

$$(16.65) \quad \text{La seule mesure } \nu \text{ de } V \times V' \text{ vérifiant}$$

$$\nu(g \otimes g') = \mu(g) \mu'(g') \quad \forall g \in \mathcal{X}(V), \quad \forall g' \in \mathcal{X}(V')$$

est égale au produit tensoriel $\mu \otimes \mu'$.

On en déduit aisément la formule

$$(16.66) \quad [\mu \otimes \mu'] \left(\begin{pmatrix} x \\ x' \end{pmatrix} \mapsto y \right) = [\mu' \otimes \mu] \left(\begin{pmatrix} x' \\ x \end{pmatrix} \mapsto y \right)$$

valable si μ et μ' sont des mesures et si $\left[\begin{pmatrix} x \\ x' \end{pmatrix} \mapsto y \right] \in \mathcal{H}(V \times V')$.

On notera aussi les formules :

$$(16.67) \quad [\mu \times f] \otimes [\mu' \times f'] = [\mu \otimes \mu'] \times [f \otimes f'] \quad \forall f, f' \in \mathcal{C}(V)$$

$$(16.68) \quad |\mu \otimes \mu'| = |\mu| \otimes |\mu'|$$

$$(16.69) \quad \text{support}(\mu \otimes \mu') = \text{support}(\mu) \times \text{support}(\mu') \quad (\text{voir 16.102})$$

EXEMPLES DE MESURES

(16.70) Soit V une variété de dimension 0.

Comme toute partie de V est ouverte (16.7), on vérifie facilement que V est séparée; que toute application de V dans R est continue; que tout compact de V est fini; donc que $\mathcal{H}(V)$ est l'ensemble des applications de V dans R nulles en dehors d'un ensemble fini. On peut donc, pour tout $g \in \mathcal{H}(V)$, définir un nombre $I(g)$ par la formule

$$\diamond \quad I(g) = \sum_{x \in V} g(x).$$

On vérifie immédiatement que I est une *mesure positive* de V ; que toute mesure de V est de la forme

$$\heartsuit \quad I \times f \quad f \in \mathcal{C}(V). \quad \blacksquare$$

(16.71) — Soit λ une fonction *non décroissante* sur R ; on peut montrer, si $f \in \mathcal{H}(R)$, que la somme

$$\sum_j f(x_j) [\lambda(x_{j+1}) - \lambda(x_j)],$$

où les x_j sont des nombres croissant avec leur numéro j , qui encadrent le support de f , a une limite S lorsque le plus grand des $x_{j+1} - x_j$ tend vers 0;

cette limite, appelée *intégrale de Stieltjes*, se note

$$(16.72) \quad \int_{-\infty}^{+\infty} f(x) d\lambda(x)$$

on vérifie facilement que $f \mapsto S$ est une mesure positive sur R ; on peut montrer réciproquement que toute mesure positive sur R est obtenue par ce procédé, à l'aide d'une fonction non décroissante λ ; on obtiendra donc toute mesure de R au moyen d'une intégrale de Stieltjes comportant une fonction λ qui est différence de deux fonctions croissantes (« fonction à variation bornée »). ■

En particulier, on définit la *mesure de Lebesgue* de R , I , par la formule

$$(16.73) \quad I(f) = \int_{-\infty}^{+\infty} f(x) dx \quad \forall f \in \mathcal{H}(R)$$

c'est évidemment une mesure positive. ■

— I étant la mesure de Lebesgue de R , le produit tensoriel I_n de n mesures égales à I est une mesure de R^n , encore appelée *mesure de Lebesgue*, et définie par

$$(16.74) \quad I_n(f) = \int_{-\infty}^{+\infty} d^1x \int_{-\infty}^{+\infty} d^2x \dots \int_{-\infty}^{+\infty} d^n x f \begin{pmatrix} 1x \\ \dots \\ nx \end{pmatrix} \quad \forall f \in \mathcal{H}(R^n).$$

Le théorème (16.66) permet de changer arbitrairement l'ordre des intégrations; on emploie souvent la notation suivante pour l'intégrale multiple $I_n(f)$

$$(16.75) \quad \int f(x) d^1x d^2x \dots d^nx \quad \blacksquare$$

— Soit V une variété séparée, a un point de V . On appelle *mesure de Dirac* au point a la mesure δ_a définie par

$$(16.76) \quad \delta_a(f) = f(a) \quad \forall f \in \mathcal{H}(V). \quad \blacksquare$$

— Soit A un homéomorphisme ⁽¹⁾ d'une variété séparée V sur une variété V^* ; on vérifie immédiatement le théorème suivant :

Si μ est une mesure de V , il existe une mesure $A^+(\mu)$ de V^* , dite *image* de μ par A , définie par

(16.77)

$$A^+(\mu)(f^*) = \mu(f^* \circ A) \quad \forall f \in \mathcal{K}(V^*).$$

Nous étendrons cette notion à un cas plus général (ci-dessous (16.149)). ■

Exemple : L'image par $\begin{pmatrix} x \\ x' \end{pmatrix} \mapsto \begin{pmatrix} x' \\ x \end{pmatrix}$ de la mesure $\mu \otimes \mu'$ est égale à $\mu' \otimes \mu$ (notations (16.64)); c'est une façon de formuler le théorème (16.66).

MESURES COMPLÈTEMENT CONTINUES

DÉFINITION

Soit E un espace vectoriel de dimension n ; on appelle *densité* de E toute fonction φ , définie sur l'ensemble des bases de E , et vérifiant

(16.78)

$$\varphi(S.M) = \varphi(S) | \det(M) |$$

pour toute base S et toute matrice $M \in GL(R^n)$.

On vérifie facilement que les densités forment un espace vectoriel de dimension 1, qui est même un espace de Riesz : φ est positive si $\varphi(S) \geq 0 \quad \forall S$.

(16.79)

Soit V une variété pure, séparée, de dimension n ($n \geq 1$); soit $x \mapsto \varphi$ une *champ de densités* de V (si $x \in V$, φ est une densité de l'espace vectoriel D_x tangent à V en x); supposons que ce champ soit *continu* (c'est-à-dire que la correspondance $\xi \mapsto \varphi(D(F)(\xi))$ soit continue pour toute carte $F[\xi \mapsto x]$ d'un atlas; cette propriété ne dépend pas du choix de l'atlas); nous allons associer à ce champ une mesure μ de V .

Supposons d'abord que f soit une fonction réelle sur V dont le support, compact, est contenu dans l'ensemble de valeurs d'une carte $F[\xi \mapsto x]$; posons ⁽²⁾

(16.80)

$$\mu(f) = \int f(x) \varphi \left(\frac{\partial x}{\partial \xi} \right) d^1 \xi d^2 \xi \dots d^n \xi ;$$

⁽¹⁾ C'est-à-dire une bijection continue ainsi que son inverse; exemple : un difféomorphisme.

⁽²⁾ On calcule l'intégrale dans R^n en prolongeant $f(x)$ par 0 en dehors de l'ensemble de définition de la carte (voir (16.100)).

la « formule de changement de variable dans les intégrales multiples », exprime que cette intégrale *ne dépend pas du choix de la carte F* , ce qui justifie la notation $\mu(f)$.

Soit maintenant A un atlas de V , f une fonction continue à support compact de V . On peut démontrer ⁽¹⁾ que l'on peut mettre f sous forme d'une somme finie

(16.81)

$$f = \sum_{j=1}^N f_j.$$

chaque terme f_j étant une fonction continue, dont le support, compact, est contenu dans l'ensemble de valeur de l'une des cartes de l'atlas; on peut vérifier que la somme $\sum_j \mu(f_j)$ ⁽²⁾ ne dépend ni du choix de l'atlas, ni de la décomposition (16.81); on pourra donc poser

(16.82)

$$\mu(f) = \sum_j \mu(f_j)$$

définissant ainsi une fonction μ sur $\mathcal{K}(V)$; il est immédiat que μ est une mesure de V , positive si φ est positive ⁽³⁾; réciproquement, si μ est positive, on peut montrer que $\varphi \geq 0 \quad \forall x$ ⁽⁴⁾; en résumé :

(16.83)

Soit $x \mapsto \varphi$ un *champ continu de densités* sur une variété pure séparée V ; il lui correspond une *mesure* μ de V , *caractérisée* par la formule (16.80), valable chaque fois que le support de f , compact, est contenu dans l'ensemble de valeurs d'une carte $\xi \mapsto x$; nous noterons

◇

$$\mu(f) = \int_V f(x) \varphi(dx).$$

On a :

♡

$$[\mu \geq 0] \Leftrightarrow [\varphi \geq 0 \quad \forall x].$$

(16.84)

— Une mesure μ sera dite *complètement continue* ⁽⁵⁾ si elle est définie ainsi à l'aide d'un champ continu $x \mapsto \varphi$ de densités; il est clair que la correspondance $[x \mapsto \varphi] \mapsto \mu$ est *linéaire*; il résulte évidemment de (16.83♡) qu'elle est *régulière*.

⁽¹⁾ Appliquer (16.16) et (16.49).

⁽²⁾ Dont chaque terme est défini par la formule (16.80).

⁽³⁾ Voir (16.49) et (16.81).

⁽⁴⁾ S'il existe un x_0 où $\varphi < 0$, on peut trouver un ouvert Ω tel que $x_0 \in \Omega$, $[x \in \Omega] \Rightarrow [\varphi < 0]$ (à cause de la continuité de $x \mapsto \varphi$); on peut trouver un $f \in \mathcal{K}(V)$, ayant son support contenu dans Ω et dans l'ensemble de valeurs d'une carte, tel que $f \geq 0$ et $f \neq 0$; on vérifie (à l'aide de (16.80)) que $\mu(f) < 0$, d'où contradiction.

⁽⁵⁾ On étend parfois ce terme à tous les éléments de $L^1(V)$ (ci-dessous (16.129)).

Exemples de mesures complètement continues

(16.85) La mesure de Lebesgue de R^n (16.75) est complètement continue, parce qu'on peut la définir à l'aide de la « densité de Lebesgue » φ :

$$\varphi(S) \equiv |\det(S)|$$

(une base S de R^n est simplement une matrice régulière). ■

Soit φ un champ continu de densités sur une variété V , μ la mesure associée :

$$\mu(f) = \int_V f(x) \varphi(dx)$$

Si $A[x \mapsto x^*]$ est un difféomorphisme de V sur une variété V^* , on vérifie sans peine que l'image $A^+(\mu)$ (16.77) est donnée par la formule

$$\diamond \quad A^+(\mu)(f^*) = \int_V f^*(x^*) \varphi^*(dx^*) \quad \forall f^* \in \mathcal{X}(V^*)$$

(16.86) $[x^* \mapsto \varphi^*]$ étant le champ de densités de V^* défini par

$$\heartsuit \quad \varphi^*(S^*) \equiv \varphi\left(\frac{\partial x}{\partial x^*} S^*\right)$$

appelé *image par A de la densité φ* ; ce champ est lui aussi continu ; $A^+(\mu)$ est donc aussi une mesure complètement continue. ■

(16.87) Soit $x = \begin{pmatrix} x^1 \\ \dots \\ x^n \end{pmatrix}$ un point variable d'un produit direct $V = V_1 \times \dots \times V_n$

de variétés séparées V_j ; supposons donné sur chaque V_j un champ continu ${}^j x \mapsto \varphi_j$ de densités, auxquels sont associées les mesures μ_j :

$$\diamond \quad \mu_j(f_j) = \int_{V_j} f_j({}^j x) \varphi_j(d{}^j x) \quad (\text{ne pas sommer !})$$

Si $S^{(j)}$ est une base de l'espace vectoriel tangent à V_j en ${}^j x$, on vérifie facilement que

$$\heartsuit \quad S = \begin{pmatrix} S^{(1)} & 0 & \dots & 0 \\ 0 & S^{(2)} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & S^{(n)} \end{pmatrix}$$

est une base de l'espace vectoriel tangent à V en x (voir (1.23)), qu'il existe un *champ continu de densités* $[x \mapsto \varphi]$ du produit direct caractérisé par

$$\clubsuit \quad \varphi(S) \equiv \varphi_1(S^{(1)}) \times \varphi_2(S^{(2)}) \times \dots \times \varphi_n(S^{(n)})$$

quels que soient les $S^{(j)}$; nous noterons

$$(16.88) \quad \varphi = \varphi_1 \otimes \varphi_2 \cdots \otimes \varphi_n$$

ou, dans la notation (16.83) \diamond

$$(16.89) \quad \varphi(dx) \equiv \varphi_1(d{}^1 x) \dots \varphi_n(d{}^n x)$$

(16.90) On vérifie — à l'aide de (16.80) et de (16.64) — que la mesure associée au champ $x \mapsto \varphi_1 \otimes \dots \otimes \varphi_n$ est le *produit tensoriel des mesures* associées à $\varphi_1, \dots, \varphi_n$; un *produit tensoriel de mesures complètement continues est donc une mesure complètement continue*. On a donc

$$(16.91) \quad \int_{V_1 \times V_2 \times \dots \times V_n} f(x) [\varphi_1 \otimes \dots \otimes \varphi_n](dx) \\ = \int_{V_1} \varphi_1(d{}^1 x) \int_{V_2} \varphi_2(d{}^2 x) \cdots \int_{V_n} \varphi_n(d{}^n x) f(x)$$

le théorème (16.66) permet de changer l'ordre des intégrations.

(16.92) — Soit G un groupe de Lie, \mathcal{G} son algèbre de Lie.

Choisissons une densité φ_0 de l'espace vectoriel \mathcal{G} ; on vérifie facilement qu'il existe sur G un seul champ de densités $a \mapsto \varphi_L$ qui soit *invariant* ⁽¹⁾ par les translations à gauche a_G de G (Définition (6.6)) et qui prenne la valeur φ_0 pour $a = e$ (élément neutre) : ce champ est donné par

$$\diamond \quad \varphi_L(S) \equiv \varphi_0([D(a_G)(e)]^{-1} \cdot S)$$

on constate sur \diamond que ce champ est continu ; il définit donc une mesure μ_L sur G ,

$$\heartsuit \quad \mu_L(f) = \int_G f(x) \varphi_L(dx)$$

qui est *invariante par toute translation à gauche* ⁽²⁾ ; μ_L s'appelle *mesure de Haar* du groupe G ; elle est définie à un facteur près par le choix de φ_0 (voir (16.78)) ; on démontre que ces mesures de Haar sont les seules qui soient *invariantes par translation à gauche*.

⁽¹⁾ C'est-à-dire *image de lui-même*, au sens (16.86), par ces translations.

⁽²⁾ C'est-à-dire que l'image de μ_L par une translation à gauche, au sens (16.77), est encore égale à μ_L .

— On peut aussi définir, de façon analogue, une densité φ_R invariante par translation à droite (6.21), coïncidant avec φ_0 au point e ; il en résulte une mesure de Haar μ_R invariante par translation à droite; on vérifie aisément que

$$\mu_R = \mu_L \times \lambda$$

λ étant la fonction définie sur G par

$$\lambda(a) \equiv |\det(\underline{a}_g)|$$

$a \mapsto \underline{a}_g$ étant la représentation adjointe de G (6.24).

La variable $\lambda(a)$ s'appelle le *module* de G ; lorsque $\lambda(a) \equiv 1$, on dit que G est *unimodulaire* (c'est évidemment le cas pour un groupe abélien; on montre aussi que les groupes *compacts* ou *semi-simples* sont unimodulaires); alors la même mesure de Haar est invariante à gauche et à droite. ■

(16.93) La mesure de Lebesgue (16.75) est invariante à droite et à gauche sur le groupe additif \mathbb{R}^n ; sur le groupe de Lie des matrices réelles

$$a \equiv \begin{bmatrix} x & y \\ 0 & 1/x \end{bmatrix}$$

le module est égal à x^2 . ■

LEMME

(16.94) Soit E un espace vectoriel euclidien (resp. symplectique); il existe une densité positive φ de E , appelée *densité euclidienne* (resp. *symplectique*), caractérisée par

$$\diamond \quad \varphi(S) = 1 \text{ pour toute base orthonormale (resp. canonique) de } E.$$

Il suffit en effet de choisir une base orthonormale (resp. canonique) S_0 ; \diamond et (16.78) donnent pour toute base S

$$(16.95) \quad \varphi(S) = |\det(S_0^{-1} \cdot S)|;$$

réciroquement, cette formule (16.95), prise comme définition de φ , fournit bien une densité positive; si S est aussi une base orthonormale (resp. symplectique) nous savons que $|\det(S_0^{-1} \cdot S)| = 1$ ((6.63) (resp. (8.30))); d'où \diamond .

— Soit V une variété symplectique séparée; en chaque point x de V , le théorème précédent permet de définir la densité symplectique φ ; on vérifie (par exemple en prenant un atlas canonique) que le champ $x \mapsto \varphi$ est continu; on définit donc une mesure μ par

$$(16.96) \quad \mu(f) \equiv \int_V f(x) \varphi(dx)$$

μ s'appelle la *mesure de Liouville* de V ; c'est une mesure positive.

— Si le support de f est contenu dans l'ensemble de valeurs d'une *carte canonique* $[\xi \mapsto x]$, il résulte immédiatement de (16.80) et (16.94) la formule

$$(16.97) \quad \mu(f) = \int \cdots \int f(x) dp_1 \cdots dp_n d[q^1] \cdots d[q^n]$$

avec la notation habituelle (10.4) pour les coordonnées canoniques.

— A l'aide de cette formule, et de (16.87) on établit le théorème suivant :

(16.98) Si on appelle μ_j les mesures de Liouville de variétés symplectiques V_j , la mesure de Liouville de la variété symplectique produit (9.7) est le produit tensoriel (16.45) des μ_j .

Soit V une variété symplectique séparée, A un élément de $\text{can}(V)$ (notation (10.15)); en comparant (16.94) et (16.86) on vérifie que l'image par A de la densité symplectique φ de V est égale à φ ; d'où, par application de (16.86), le *théorème de Liouville* :

(16.99) La mesure de Liouville d'une variété symplectique séparée V est invariante par tout symplectomorphisme de V .

SUPPORT D'UNE MESURE

(16.100) Soit Ω un ouvert d'une variété séparée V ; on sait que Ω possède aussi une structure de variété (séparée, de même dimension); si $f \in \mathcal{K}(\Omega)$, \hat{f} défini sur V par

$$\diamond \quad f(x) = \begin{cases} f(x) & \text{si } x \in \Omega \\ 0 & \text{si } x \notin \Omega \end{cases}$$

est un prolongement de f ; on vérifie que $\hat{f} \in \mathcal{K}(V)$.

Si donc μ est une mesure de V , on peut définir une application μ_Ω par $\mu_\Omega(f) \equiv \mu(\hat{f})$; on vérifie que μ_Ω est une mesure de Ω , positive si μ est positive; μ_Ω s'appelle mesure *induite* par μ sur l'ouvert Ω .

THÉORÈME

(16.101) — Ω étant donné, la correspondance $[\mu \mapsto \mu_\Omega]$ est linéaire et croissante (entre les espaces de Riesz des mesures de V et Ω); on a $|\mu_\Omega| \equiv |\mu|_\Omega$.

(16.102) — La réunion des ouverts Ω_j vérifiant $[\mu_{\Omega_j} = 0]$ est un ouvert (éventuellement vide) Ω ; on vérifie que $\mu_\Omega = 0$: Ω est donc le plus grand ouvert tel que $\mu_\Omega = 0$; son complémentaire — qui est fermé — s'appelle *support* de la mesure μ .

(16.103) **Exemples :** Le support de la mesure de Dirac δ_a est réduit au point a ; le support de la mesure associée à un champ continu de densités $x \mapsto \varphi$ est la fermeture de l'ensemble des x tels que $\varphi \neq 0$ (cf. (16.42)).

— On démontre ⁽¹⁾ que

(16.104) Toute fonction $f \in \mathcal{X}(V)$ qui est nulle sur le support d'une mesure μ vérifie $\mu(f) = 0$.

MESURES BORNÉES

DÉFINITION

Soit E un espace vectoriel (réel ou complexe). On dit que E est un *espace vectoriel normé* si on a associé à chaque $X \in E$ un nombre réel $\|X\|$, appelé *norme* de X , de façon que

(16.105)
$$\diamond \begin{cases} \|X\| > 0 & \text{si } X \neq 0, \\ \|sX\| = |s| \times \|X\| & \text{si } X \in E, \ s \text{ scalaire } ^{(2)}, \\ \|X + Y\| \leq \|X\| + \|Y\| & \forall X, Y \in E. \end{cases}$$

Exemples : R (ou C) est évidemment un espace vectoriel normé; de même R^n (1.2).

DÉFINITION, THÉORÈME

On appelle *espace de Banach*, ou espace vectoriel normé *complet* tout espace vectoriel normé où les séries *absolument convergentes* sont *convergentes* ⁽³⁾ :

(16.106)
$$\boxed{X_j \in E \ \forall j = 1, 2, 3, \dots; \exists a \in R; \forall n; \sum_{j=1}^n \|X_j\| \leq a} \Rightarrow$$

$$\boxed{\exists X \in E; \forall \varepsilon > 0; \exists N; \forall n \geq N; \left\| X - \sum_{j=1}^n X_j \right\| \leq \varepsilon}$$

⁽¹⁾ (Voir Bourbaki, loc. cit.).

⁽²⁾ $|s|$ désigne la valeur absolue ou le module du nombre réel ou complexe s — on voit que

$$[\|X\| = 0] \Leftrightarrow [X = 0].$$

⁽³⁾ On écrit en abrégé $\sum_{j=1}^{\infty} \|X_j\| < \infty \Rightarrow \sum_{j=1}^{\infty} X_j = X$; notons que $\left\| \sum_{j=1}^{\infty} X_j \right\| \leq \sum_{j=1}^{\infty} \|X_j\|$.

Exemple : On démontre que tout espace vectoriel normé de dimension finie est un espace de Banach.

DÉFINITION, THÉORÈME

Soient E et E' deux espaces vectoriels normés, $L(E, E')$ l'ensemble des applications linéaires de E dans E' .

Soit $A \in L(E, E')$; on dit que A est *borné* si le nombre

$$\diamond \quad \|A\| = \sup_{\|X\| \leq 1} \|A(X)\|$$

(16.107) est fini; on a alors

$$\heartsuit \quad \|A(X)\| \leq \|A\| \times \|X\| \quad \forall X \in E.$$

— L'ensemble des éléments bornés de $L(E, E')$ est un *espace vectoriel normé* : c'est un espace de Banach si E' est un espace de Banach.

La vérification est élémentaire.

Exemple

(16.108) Soit E un espace vectoriel normé; on appelle *dual topologique* de E l'ensemble E' des éléments bornés du dual $E^*(= L(E, R)$ ou $L(E, C)$); E' est donc un espace de Banach; on démontre le théorème (Hahn-Banach)

(16.109)
$$\boxed{\|X\| = \sup_{C \in E', \|C\| \leq 1} |C(X)|} \quad \forall X \in E. \blacksquare$$

THÉORÈME

Soit V une variété séparée, $\mathcal{X}(V)$ l'espace vectoriel des fonctions continues sur V à support compact (16.42).

Si l'on pose

(16.110)
$$\diamond \quad \|f\| = \sup_{x \in V} |f(x)| \quad \forall f \in \mathcal{X}(V)$$

$\mathcal{X}(V)$ devient un espace vectoriel normé.

Le fait que \diamond définisse un nombre $\|f\|$ fini résulte du théorème (16.24); la vérification des axiomes (16.105 \diamond) est alors triviale.

Il résulte immédiatement de (16.51) que

(16.111) *La dual topologique de l'espace vectoriel normé $\mathcal{X}(V)$ est l'ensemble des mesures bornées.*

$[\mathcal{X}(V)]'$ est donc un espace de Banach (16.107).

THÉORÈME

(16.112) [Toute mesure μ à support compact est bornée.

Soit K le support de μ , f l'une des fonctions dont l'existence est assurée par le théorème (16.48); on vérifie que $\mu = \mu \times f$; μ est donc bornée en vertu du théorème (16.133) ci-dessous.

(16.113) — En particulier, sur une variété compacte, toutes les mesures sont bornées; par contre la mesure de Lebesgue de R^n n'est pas bornée.

THÉORÈME

(16.114) a) Si μ et μ' sont des mesures bornées de V et V' , $\mu \otimes \mu'$ est une mesure bornée de $V \times V'$, et l'on a

$$\|\mu \otimes \mu'\| = \|\mu\| \times \|\mu'\|$$

b) $[\mu \otimes \mu' \neq 0, \mu \otimes \mu' \text{ bornée}] \Rightarrow [\mu \text{ bornée}, \mu' \text{ bornée}]$.

La démonstration résulte simplement de (16.59) et de (16.64).

THÉORÈME

(16.115) Si μ est une mesure

$$[\mu \text{ bornée}] \Leftrightarrow [|\mu| \text{ bornée}] \Rightarrow [\|\mu\| = \||\mu|\|].$$

(16.54) donne $|\mu(f)| \leq |\mu|(|f|)$; d'où $\|\mu\| \leq \||\mu|\|$; (16.53) donne $|\mu|(|f|) \leq |\mu|(|f|) = \sup_{|g| \leq |f|} \mu(g)$, d'où $\||\mu|\| \leq \|\mu\|$.

THÉORÈME

(16.116) Si μ et ν sont des mesures bornées,

$$[\mu \geq 0, \nu \geq 0] \Rightarrow \|\mu + \nu\| = \|\mu\| + \|\nu\|.$$

On sait que $\|\mu + \nu\| \leq \|\mu\| + \|\nu\|$ (16.105 \diamond); soit d'autre part $\varepsilon > 0$; $\exists f, g \in \mathcal{X}(V)$, $\|f\| \leq 1$, $\|g\| \leq 1$; $\mu(f) \geq \|\mu\| - \varepsilon$, $\nu(g) \geq \|\nu\| - \varepsilon$; si $h = \sup(|f|, |g|)$, on a $\|h\| \leq 1$, $[\mu + \nu](h) \geq \|\mu\| + \|\nu\| - 2\varepsilon$; d'où $\|\mu + \nu\| \geq \|\mu\| + \|\nu\| - 2\varepsilon$.

C.Q.F.D.

Grâce à (16.33), on en déduit pour toute μ bornée

$$\|\mu\| = \|\mu^+\| + \|\mu^-\|.$$

En collationnant (16.38), (16.107) et (16.115), on voit que

(16.118) Dans l'espace de Riesz des mesures d'une variété V , les mesures bornées forment un sous-espace de Riesz.

— Si μ et ν sont des mesures positives bornées, il résulte immédiatement de (16.116) que le nombre $\|\mu\| - \|\nu\|$ ne dépend que de la différence $\mu - \nu$; il en résulte que la masse d'une mesure bornée, définie par la formule

$$\text{masse}(\mu) = \|\mu^+\| - \|\mu^-\|$$

vérifie

$$[\mu \geq 0, \nu \geq 0, \mu \text{ et } \nu \text{ bornées}] \Rightarrow \|\mu\| - \|\nu\| = \text{masse}(\mu - \nu)$$

On déduit aisément de cette formule et de (16.116) que l'application « masse » est linéaire :

$$\begin{cases} \text{masse}(\mu + \nu) \equiv \text{masse}(\mu) + \text{masse}(\nu) \\ \text{masse}(s\mu) \equiv s \text{masse}(\mu) \quad s \in R. \end{cases}$$

Notons aussi les formules

$$|\text{masse}(\mu)| \leq \|\mu\|$$

$$[\text{masse}(\mu) = \|\mu\|] \Leftrightarrow [\mu \geq 0]$$

THÉORÈME

(16.123) Si μ et μ' sont des mesures bornées de V et V' , on a

$$\text{masse}(\mu \otimes \mu') = \text{masse}(\mu) \times \text{masse}(\mu').$$

Il suffit de remarquer que $\mu \otimes \mu'$ est la différence des mesures $\mu^+ \otimes \nu^+ + \mu^- \otimes \nu^-$, $\mu^+ \otimes \nu^- + \mu^- \otimes \nu^+$ qui sont positives (16.60) et bornées (16.114), et d'appliquer (16.114), (16.116).

DÉFINITION

Soit Ω un ouvert de V , μ une mesure de V .

(16.124) On dit que Ω est μ -mesurable si la mesure μ_Ω (Définition (16.100)) est bornée; alors le nombre masse (μ_Ω) s'appelle μ -mesure de Ω ; on le note généralement $\mu(\Omega)$.

On vérifie facilement que Ω est μ -mesurable dès que μ est bornée ou que Ω est relativement compact; que

$$[\Omega \cap \Omega' = \emptyset] \Rightarrow \mu(\Omega \cup \Omega') = \mu(\Omega) + \mu(\Omega');$$

dans le cas $\mu \geq 0$, on a

$$(16.126) \quad \begin{cases} \mu(\Omega) \geq 0 \\ [\Omega \subset \Omega'] \Rightarrow \mu(\Omega) \leq \mu(\Omega'). \end{cases}$$

Exemples

(16.127) Dans R^n , une *boule* B de rayon r ($r > 0$) :

$$x \in B \Leftrightarrow \|x\| < r$$

est un ouvert relativement compact; μ étant la mesure de Lebesgue de R^n , le nombre $\mu(B)$ s'appelle souvent *volume* de B ; on montre qu'il vaut ⁽¹⁾

$$\mu(B) = \frac{\pi^{n/2}}{[n/2]!} r^n.$$

— La *sphère* S de rayon r

$$x \in S \Leftrightarrow \|x\| = r$$

possède une structure de variété compacte plongée dans R^n ; la densité euclidienne (16.94) définit une mesure μ sur S ; le nombre $\mu(S) = \|\mu\|$ s'appelle *aire* de la sphère; le calcul montre que

$$\|\mu\| = \frac{n\pi^{n/2} r^{n-1}}{[n/2]!}.$$

(16.128) — La notion de mesure des ensembles peut s'étendre à une classe beaucoup plus vaste que celle des ensembles ouverts (voir Bourbaki); mais nous n'utiliserons pas cette extension. ■

(16.129) Soit V une variété pure séparée; nous appellerons $L^1(V)$ l'ensemble des mesures μ de V qui sont bornées et qui vérifient

$$\forall \varepsilon > 0; \exists \mu_0 \text{ mesure complètement continue de } V; \|\mu - \mu_0\| < \varepsilon.$$

Le lecteur pourra vérifier, à titre d'exercice, les théorèmes suivants :

(16.130) $L^1(V)$ est un espace de Banach.

(16.131) $L^1(V)$ est un sous-espace de Riesz ⁽²⁾ de l'espace des mesures (bornées) de V .

(16.132) $L^1(V) \otimes L^1(V') \subset L^1(V \otimes V')$ ⁽³⁾.

⁽¹⁾ Nous posons $x! = \Gamma(1+x) = \int_0^\infty e^{-t} t^x dt$; pour n impair, $[\frac{n}{2}]!$ se calcule par récurrence, à partir de la valeur $[-1/2]! = \sqrt{\pi}$, par la formule $x! = x[x-1]!$

⁽²⁾ Etablir d'abord l'identité $\|\mu - \mu_0\| \leq \|\mu - \mu_0\|$, conséquence de (16.36) et de (16.54).

⁽³⁾ Abus de notations signifiant : $\mu \in L^1(V), \mu' \in L^1(V') \Rightarrow \mu \otimes \mu' \in L^1(V \times V')$; utiliser (16.90) et (16.114).

FONCTIONS INTÉGRABLES

LEMME

(16.133) Si μ est une mesure de V , si $f \in \mathcal{K}(V)$, $\mu \times f$ (notation (16.56)) est une mesure bornée, et l'on a

$$\diamond \quad \text{masse}(\mu \times f) = \mu(f).$$

Il est facile de vérifier, si μ et f sont positives, que $\mu \times f$ est positive, et que ⁽¹⁾

$$(16.134) \quad \|\mu \times f\| = \mu(f).$$

On passe à \diamond en y remplaçant μ et f par des différences d'éléments positifs, et en appliquant les propriétés de linéarité.

COROLLAIRE, DÉFINITION

(16.135) Soit μ une mesure de V , f une fonction continue sur V .

On dit que f est μ -intégrable si la mesure $\mu \times f$ est bornée.

L'application linéaire μ se prolonge linéairement à l'espace vectoriel des fonctions continues μ -intégrables en posant :

$$\mu(f) = \text{masse}(\mu \times f).$$

Bien entendu, dans le cas d'une mesure définie par une densité φ (16.83), nous emploierons encore la notation

$$(16.136) \quad \int_V f(x) \varphi(dx)$$

pour désigner le nombre $\mu(f)$, même si le support de f n'est pas compact : on dit dans ce cas que l'intégrale est *convergente*.

Exemples

(16.137) Soit V une variété de dimension 0; I la mesure définie en (16.70). On sait que toute fonction f définie sur V est continue; on dit que f est *sommable* si f est I -intégrable; alors le nombre $I(f)$ se note

$$\sum_{x \in V} f(x)$$

mais il s'agit en général d'une « somme infinie ».

⁽¹⁾ Utiliser (16.48).

Dans le cas particulier où V est l'ensemble des entiers positifs, on peut montrer que f est sommable si la série

$$f(1) + f(2) + \dots + f(n) + \dots$$

est absolument convergente; alors sa somme est égale à $I(f)$. ■

(16.138) Soit f une fonction continue sur un intervalle ouvert borné $a < x < b$ de R ; si f est intégrable pour la densité de Lebesgue, son intégrale se note $\int_a^b f(x) dx$; on peut montrer que cette intégrale est la limite de $F(\beta) - F(\alpha)$, (F étant une primitive de f), lorsque β tend vers b et α vers a ; ainsi

$$\int_{-1}^{+1} \frac{dx}{\sqrt{1-x^2}} = \pi;$$

si f n'est pas positive, cette limite de $F(\beta) - F(\alpha)$ peut exister sans que f soit intégrable. ■

THÉORÈME

Si des mesures μ, μ' et des fonctions continues f, f' vérifient

(16.139)

$ \mu \leq \mu'$	$ f \leq f'$	f' est μ' -intégrable
-------------------	---------------	-----------------------------

f est μ -intégrable, et l'on a

$$|\mu(f)| \leq \mu'(f').$$

Si $g \in \mathcal{H}(V)$, $|g| \leq 1$, on a $|f \times g| \leq f'$; d'où grâce à (16.54),

$$|[\mu \times f](g)| = |\mu(f \times g)| \leq \mu'(f');$$

on voit que $\mu \times f$ est borné, et que $\|\mu \times f\| \leq \mu'(f')$; il suffit d'appliquer l'inégalité (16.122) à $\mu \times f$.

THÉORÈME

(16.140)

Si f est μ -intégrable, $|f|$ est $|\mu|$ -intégrable.

En effet $[f \text{ est } \mu\text{-intégrable}] \Leftrightarrow [\mu \times f \text{ bornée}] \Rightarrow$
 $[|\mu \times f| \text{ bornée}]$ (16.115) $\Leftrightarrow [|\mu| \times |f| \text{ bornée}]$ (16.58). ■

COROLLAIRE

(16.141)

$[\mu \geq 0, f \in \mathcal{C}(V); f \geq 0; f \text{ est } \mu\text{-intégrable}] \Rightarrow [\mu(f) \geq 0]$. ■

(16.142)

En combinant les deux théorèmes (16.139), (16.140) on constate que les 4 hypothèses $[f \mu\text{-intégrable}]$, $[|f| \mu\text{-intégrable}]$, $[f | \mu| \text{-intégrable}]$, $[|f| | \mu| \text{-intégrable}]$, sont équivalentes; elles entraînent

$$\diamond \quad \mu(f) = \mu(f^+) - \mu(f^-). \quad \blacksquare$$

(16.143)

Soient μ et μ' deux mesures des variétés V et V' . Soient $f \in \mathcal{C}(V)$, $f' \in \mathcal{C}(V')$; alors :

$[f \text{ est } \mu\text{-intégrable}, f' \text{ est } \mu'\text{-intégrable}] \Rightarrow$

$$\left[\begin{array}{l} f \otimes f' \text{ est } [\mu \otimes \mu']\text{-intégrable} \\ [\mu \otimes \mu'](f \otimes f') = \mu(f) \times \mu'(f') \end{array} \right]$$

Simple corollaire de (16.67) et (16.114).

(16.144)

— Soit F une application continue de V dans un espace vectoriel E de dimension finie ⁽¹⁾.

On dit que F est μ -intégrable, si, pour tout élément C du dual de E , l'application $x \mapsto C(F(x))$ est μ -intégrable; dans ce cas, il existe un élément de E , que nous noterons $\mu(F)$ — ou $\int_V F(x) \varphi(dx)$ dans le cas (16.83) — défini par

(16.145)

$$C(\mu(F)) = \mu(C \circ F) \quad \forall C \in E^*.$$

Pour établir ce résultat, il suffit de choisir une base S de E ; si l'on pose

(16.146)

$$F(x) \equiv S \begin{pmatrix} {}^1f(x) \\ \dots \\ {}^nf(x) \end{pmatrix}$$

on vérifie immédiatement que

(16.147)

$$\left\{ \begin{array}{l} [F \text{ est } \mu\text{-intégrable}] \Leftrightarrow [{}^j f \text{ est } \mu\text{-intégrable} \quad \forall j = 1, 2, \dots, n] \\ \Rightarrow \left[\mu(F) = S \begin{pmatrix} \mu({}^1f) \\ \dots \\ \mu({}^nf) \end{pmatrix} \right]. \quad \blacksquare \end{array} \right.$$

⁽¹⁾ Le lecteur trouvera dans Bourbaki une extension de cette notion au cas où E est un espace de Banach.

THÉORÈME

Soient V une variété séparée ; μ une mesure de V ; f et g deux fonctions continues ⁽¹⁾ sur V telles que

(16.148)

$$\begin{cases} f \text{ est } \mu\text{-intégrable,} \\ f(x) = g(x) \quad \forall x \in \text{support de } \mu. \end{cases}$$

Alors g est μ -intégrable, et l'on a

$$\mu(g) = \mu(f).$$

La fonction continue $h = g - f$ est nulle sur le support de μ ; donc aussi, $\forall k \in \mathcal{X}(V)$, la fonction à support compact $h \times k$; on a donc (16.104) $\mu(h \times k) = 0$; d'où, k étant arbitraire, $\mu \times h = 0$; d'où $\mu \times g = \mu \times f$; d'où évidemment $\|\mu \times g\| = \|\mu \times f\| < \infty$, $\mu(g) = \text{masse}(\mu \times g) = \text{masse}(\mu \times f) = \mu(f)$. ■

IMAGES D'UNE MESURE

DÉFINITION

(16.149)

Soient V et V^* deux variétés séparées ; A une application continue d'un ouvert de V dans V^* .

a) Nous dirons qu'une mesure μ de V est A -transportable ⁽²⁾ si

$$\heartsuit \quad [f^* \in \mathcal{X}(V^*)] \Rightarrow [f^* \circ A \text{ est } \mu_{\text{def}(A)}\text{-intégrable}].$$

b) Dans l'espace de Riesz des mesures de V , celles qui sont A -transportables forment un sous-espace de Riesz.

c) Si μ est A -transportable, il existe une mesure de V , appelée image de μ par A , notée $A^+(\mu)$, définie par

$$\diamond \quad A^+(\mu)(f^*) = \mu_{\text{def}(A)}(f^* \circ A) \quad [\forall f^* \in \mathcal{X}(V^*)].$$

d) L'application A^+ ainsi définie est linéaire et croissante ⁽³⁾.

— La vérification de (b) et de la linéarité de A^+ est élémentaire, compte tenu de (16.101) et de (16.57) ; si $\mu \geq 0$, on voit que $A^+(\mu)$ est une forme linéaire positive sur $\mathcal{X}(V^*)$, donc une mesure ; pour toute mesure μ A -transportable, on a $A^+(\mu) = A^+(\mu^+) - A^+(\mu^-)$; $A^+(\mu)$ est donc une mesure. La croissance de A^+ est alors immédiate.

C.Q.F.D.

⁽¹⁾ Le résultat est établi en supposant f et g à valeurs réelles ; il s'étend immédiatement au cas des valeurs dans un espace vectoriel de dimension finie.

⁽²⁾ Synonyme : [μ est A -transportable] \Leftrightarrow [A est μ -propre].

⁽³⁾ De l'espace de Riesz des mesures A -transportables de V dans celui des mesures de V^* .

THÉORÈME (hypothèses de (16.149)) :

Supposons que μ soit A -transportable et que $\mu \geq 0$.
Alors la formule (16.149) \diamond

(16.150)

$$\diamond \quad A^+(\mu)(f^*) = \mu_{\text{def}(A)}(f^* \circ A)$$

reste vraie pour toute $f^* \in \mathcal{X}(V^*)$ telle que l'un des deux membres existe ⁽¹⁾.

a) Supposons d'abord $f^* \geq 0$; compte tenu de la définition (16.135) et de (16.122), la proposition à démontrer s'écrit $X = Y$, avec

$$X = \sup_h \mu_{\text{def}(A)}([f^* \circ A] \times h), \quad Y = \sup_{g^*, h} \mu_{\text{def}(A)}([f^* \times g^*] \circ A \times h)$$

les indices g^* , h appartenant aux ensembles respectifs suivants :

$$[g^* \in \mathcal{X}(V^*), \quad 0 \leq g^* \leq 1]; \quad [h \in \mathcal{X}(\text{def}(A)); \quad 0 \leq h \leq 1].$$

Il est évident que $Y \leq X$, puisque $[f^* \times g^*] \circ A \times h \leq [f^* \circ A] \times h$; d'autre part, $\forall h$, désignons par K le support de h (qui est un compact de $\text{def}(A)$) ; son image K^* par A , qui est continu, est un compact de V^* (Théorème (16.22)) ; d'après le lemme d'Urysohn, il existe un g^* vérifiant $g^* \in \mathcal{X}(V^*)$, $0 \leq g^* \leq 1$, tel que $[x^* \in K^*] \Rightarrow [g^*(x^*) = 1]$; si $x \in \text{def}(A)$, on a

$$[h(x) \neq 0] \Rightarrow [x \in K] \Rightarrow [A(x) \in K^*] \Rightarrow [g^*(A(x)) = 1]; \text{ d'où } h = [g^* \circ A] \times h,$$

et évidemment $[f^* \circ A] \times h = [f^* \times g^*] \circ A \times h$; d'où $X \leq Y$.

b) Supprimons maintenant l'hypothèse $f^* \geq 0$; on remarque que

$$[f^*]^+ \circ A = [f^* \circ A]^+, \quad [f^*]^- \circ A = [f^* \circ A]^- \quad (\text{voir (16.44)});$$

il en résulte que les deux membres de \diamond , s'ils existent, sont respectivement égaux à (Théorème (16.142) \diamond)

$$A^+(\mu)(f^{*+}) - A^+(\mu)(f^{*-}) \\ \mu_{\text{def}(A)}(f^{*+} \circ A) - \mu_{\text{def}(A)}(f^{*-} \circ A)$$

donc égaux puisque le résultat établi en a) est valable pour les fonctions positives f^{*+} , f^{*-} . ■

THÉORÈME

(16.151)

Si $A^+(\mu)$ existe et si $B^+(A^+(\mu))$ existe, on a

$$[B \circ A]^+(\mu) = B^+(A^+(\mu)).$$

Dans le cas $\mu \geq 0$, il suffit de développer $[B \circ A]^+(\mu)(f)$ et $B^+(A^+(\mu))(f)$ par la formule (16.150 \diamond) ; on passe au cas général en considérant μ comme différence des deux mesures positives μ^+ et μ^- , auxquelles la formule est applicable, et en appliquant (16.149).

⁽¹⁾ C'est-à-dire, donc, si f^* est $A^+(\mu)$ -intégrable.

THÉORÈME (hypothèses de (16.149))

(16.152) Si μ est A -transportable, le support de $A^+(\mu)$ est contenu dans la fermeture de l'image par A du support de μ .

Il suffit de vérifier, si Ω est un ouvert de V^* qui ne rencontre pas cette image, que $[A^+(\mu)]_\Omega = 0$; or $[A^+(\mu)]_\Omega = [1_\Omega]^+(A^+(\mu)) = C^+(\mu)$, en posant $C = 1_\Omega \cdot A$ (Théorème précédent); l'ensemble de définition de C est un ouvert de V qui ne rencontre pas le support de μ ; on a donc $\mu_{\text{def}(C)} = 0$ (16.102); d'où, grâce à la définition (16.149), $C^+(\mu) = 0$.

THÉORÈME (hypothèses de (16.149))

(16.153) Toute mesure bornée μ de V est A -transportable; on a alors

$$\|A^+(\mu)\| \leq \|\mu\|.$$

Si $f^* \in \mathcal{X}(V^*)$, $g \in \mathcal{X}(\text{def}(A))$, on vérifie que

$$\|[\widehat{f^* \circ A}] \times g\| \leq \|f^*\| \times \|g\|$$

le $\widehat{}$ désignant le prolongement par 0 dans le complémentaire de $\text{def}(A)$ (16.100 \diamond); d'où $\|\mu_{\text{def}(A)} \times [\widehat{f^* \circ A}]\| \leq \|\mu\| \times \|f^*\|$; ce qui conduit immédiatement au résultat.

(16.154) — On voit que A^+ définit une application *linéaire, bornée, croissante* de l'espace de Banach-Riesz des mesures bornées de V dans l'espace correspondant de V^* .

THÉORÈME

(16.155) Si A est un homéomorphisme de V à V^* , et μ une mesure bornée de V

$$\|A^+(\mu)\| = \|\mu\|.$$

Soit $\mu^* = A^+(\mu)$; le théorème (16.151) montre que $\mu = [A^{-1}]^+(\mu^*)$; on a donc $\|\mu^*\| \leq \|\mu\|$, $\|\mu\| \leq \|\mu^*\|$.

THÉORÈME (hypothèses de (16.149))

(16.156) Si $\text{def}(A) = V$, et si μ est bornée, on a

$$\text{masse}(A^+(\mu)) = \text{masse}(\mu).$$

a) Supposons d'abord $\mu \geq 0$; si on pose $f^*(x^*) = 1[\forall x^* \in V^*]$, notons que $f = f^* \circ A$ vérifie $f(x) = 1[\forall x \in V]$; comme la masse d'une mesure bornée est égale à la valeur de cette mesure appliquée à la constante 1 (Définition (16.135)), on a $\text{masse}(\mu) = \mu(f)$, $\text{masse}(A^+(\mu)) = A^+(\mu)(f^*)$; ces deux nombres sont égaux d'après (16.150).

b) Dans le cas général, il suffit d'appliquer le résultat (a) aux mesures positives μ^+ et μ^- , et d'utiliser la linéarité des applications A^+ et « masse ».

C.Q.F.D.

THÉORÈME

(16.157) Soient V et V^* deux variétés séparées; A une application *étalée* (1) de V à V^* .

Alors toute mesure μ de V est A -transportable.

a) Supposons $\mu \geq 0$; posons $\Omega = \text{def}(A)$.

Si $f^* \in \mathcal{X}(V^*)$, l'ensemble des $x^* \in V^*$ tels que $f^*(x^*) \neq 0$ est un ouvert (image réciproque par f^* d'un ouvert de R) relativement compact (car contenu dans le support compact de f^*); son image réciproque par A est donc un ouvert relativement compact Ω' (puisque A est étalée); notons que

$$x \in \Omega' \Leftrightarrow f(x) \neq 0,$$

en posant

$$f = f^* \cdot A.$$

Soit $\bar{\Omega}'$ la fermeture (compacte) de Ω' ; grâce au lemme d'Urysohn (16.48),

$$\exists h \in \mathcal{X}(V) : h \geq 0, \quad [x \in \bar{\Omega}'] \Rightarrow [h(x) = 1].$$

Si $x \in \Omega$ on a, soit $f(x) = 0$, soit $x \in \Omega'$; d'où $f(x) = f(x) h(x)$.

Soit $g : g \in \mathcal{X}(\Omega)$, $\|g\| \leq 1$; $\forall x \in \Omega$, on a

$$|[f \times g](x)| = |f(x) h(x) g(x)| = |f^*(A(x)) h(x) g(x)| \leq \|f^*\| h(x);$$

cette inégalité vaut aussi pour le prolongement $\widehat{f \times g}$ nul en dehors de Ω (16.100), puisque $h \geq 0$; on a donc $|\widehat{f \times g}| \leq \|f^*\| h$, d'où

$$|[\mu_\Omega \times f](g)| = |\mu_\Omega(f \times g)| = |\mu(\widehat{f \times g})| \leq \mu(\widehat{f \times g}) \leq \|f^*\| \mu(h);$$

ce qui montre que $\mu_\Omega \times f$ est *borné*, c'est-à-dire que f est μ_Ω intégrable; on a donc $\mu_{\text{def}(A)}(f^* \cdot A) = \mu_\Omega(f) = \text{masse}(\mu_\Omega \times f)$ (Définition (16.135)); la formule (16.149 \diamond) définit donc un opérateur $A^+(\mu)$ sur $\mathcal{X}(V^*)$, visiblement linéaire; on constate si $f^* \geq 0$, que $A^+(\mu)(f^*) \geq 0$; $A^+(\mu)$ est donc une mesure positive de V .

b) On passe au cas général en écrivant $\mu = \mu^+ - \mu^-$, en appliquant (a) aux mesures positives μ^+ et μ^- , et en utilisant (16.149).

(16.158) — On voit donc, si A est étalée de V à V^* , que A^+ définit une application linéaire croissante de l'espace de Riesz des mesures de V dans celui de V^* ; cette application prolonge celle, définie sur les mesures *bornées*, qui a été étudiée en (16.154).

(16.159) En particulier, si A est un homéomorphisme de V à V^* , A est étalée, $A^+(\mu)$ existe pour toute mesure μ de V ; on retrouve d'ailleurs la définition déjà donnée en (16.77).

(1) Définition (16.26); on rappelle que $\text{def}(A)$ est ouvert, et que A est continue.

Exemples

Convolution sur un groupe de Lie

(16.160) Soit G un groupe de Lie, μ et μ' deux mesures bornées sur G . On appelle *convoluté* de μ et μ' la mesure bornée, image de $\mu \otimes \mu'$ par l'application $\begin{pmatrix} x \\ y \end{pmatrix} \mapsto x \times y$; on la note (1) $\mu \times \mu'$; on vérifie facilement que la *convolution* \times est une opération interne, bilinéaire, et associative (2) (3) sur l'espace B des mesures bornées de G ; on dit que B est une *algèbre*.

(16.161) On sait d'autre part que B est un espace de Banach (16.111); la convolution vérifie (4)

$$\|\mu \times \nu\| \leq \|\mu\| \times \|\nu\|;$$

on dit que B est une *algèbre de Banach*.

Si $\mu \in B$, on désigne par μ^* l'image par $[x \mapsto x^{-1}]$ de μ ; il est clair que $\mu^* \in B$.

Il est immédiat (voir (16.154)) que

$$(16.162) \quad [\mu + \mu']^* \equiv \mu^* + \mu'^*$$

$$(16.163) \quad [s\mu]^* \equiv s[\mu^*] \quad \forall s \in \mathbb{R};$$

Puisque $x \mapsto x^{-1}$ est un difféomorphisme de G à G , on a (5)

$$(16.164) \quad \|\mu^*\| \equiv \|\mu\|;$$

d'autre part les identités $[x^{-1}]^{-1} \equiv x$, $[x \times y]^{-1} \equiv y^{-1} \times x^{-1}$, valables dans G , entraînent les formules

$$(16.165) \quad [\mu^*]^* \equiv \mu,$$

$$(16.166) \quad [\mu \times \nu]^* \equiv \nu^* \times \mu^*;$$

on résume les formules (16.162) à (16.166) en disant que l'algèbre de Banach B est *involutive*. On peut montrer que $L^1(G)$ est une sous-algèbre de Banach de B , évidemment involutive; on l'appelle généralement *algèbre du groupe* G .

(1) Dans le cas d'un groupe additif, on emploie généralement la notation $\mu * \mu'$.

(2) La vérification montre que $\mu \times \mu' \times \mu''$ est l'image par $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto x \times y \times z$ de

$$\mu \otimes \mu' \otimes \mu''.$$

(3) Si G est abélien, la convolution est commutative sur G .

(4) Appliquer (16.114) et (16.153).

(5) Théorème (16.155).

— Notons aussi la formule

$$(16.167) \quad \text{masse}(\mu \times \nu) = \text{masse}(\mu) \text{ masse}(\nu),$$

valable sur l'algèbre $B(1)$ et évidemment sur sa sous-algèbre $L^1(G)$. ■

Photométrie

(16.168) La photométrie classique a pour objet de déterminer le *flux lumineux* qui traverse une surface donnée (dans un sens ou dans l'autre) lorsqu'elle est exposée à diverses sources lumineuses; ce flux est un nombre, qui se mesure par exemple en *lumens*.

La théorie doit évidemment indiquer comment le flux dépend de la surface; il se trouve qu'on peut formuler cette dépendance de la façon suivante :

(16.169) — Un *état photométrique* est une *mesure positive* μ sur la variété V des rayons lumineux (2) .

— Si Ω désigne l'ensemble des rayons lumineux qui traversent une surface donnée dans un sens donné (3) , le flux lumineux correspondant est égal à $\mu(\Omega)$ (4) .

Sans développer complètement les bases scientifiques de l'éclairage, indiquons quelques utilisations de ces axiomes :

(16.170) — Soit S une source lumineuse bornée. L'ensemble Ω des rayons lumineux issus de cette source, qui traversent une sphère entourant S ne dépend pas du choix de la sphère; le nombre $\mu(\Omega)$ est le *flux* de la source; il se mesure donc en lumens.

(16.171) — Supposons que la mesure μ soit complètement continue; la densité correspondante est alors proportionnelle à la *densité symplectique* de V (5) ; le facteur de proportionnalité sera noté $\lambda(\xi)/\pi$ (ξ désignant un rayon lumineux variable); le nombre $\lambda(\xi)$ s'appelle la *brillance* (ou *luminance*); il se mesure en lumen \otimes mètre $^{-2}$, appelés *blondels* pour la circonstance.

(16.172) — Un état photométrique est *uniforme* si la luminance est constante; on considère souvent des états *induits sur un ouvert par un état uniforme* (16.100); exemples : le ciel vu à travers une lucarne; le rayonnement calorifique à l'intérieur d'un four isotherme (voir ci-dessous (17.151)).

(1) Utiliser (16.123) et (16.156).

(2) Définition (15.87); la dimension de V est 4.

(3) On rappelle qu'un rayon lumineux est une droite orientée (15.87).

(4) Définition (16.124); $\mu(\Omega)$ est égal à $\|\mu\|$ puisque μ est positive (16.122), donc ≥ 0 .

(5) Définition (16.94); nous choisissons pour V la structure symplectique définie par la formule (15.90).

(16.173) — Considérons un écran dont la surface est une variété de dimension 2 plongée dans l'espace R^3 ; si on désigne par r le point où un rayon lumineux ξ perce cet écran (dans le bon sens !) on peut montrer que l'application $\xi \mapsto r$ est étalée; l'image de l'état photométrique par cette application constitue justement l'image reçue sur l'écran (au sens photographique du terme); si celle-ci est complètement continue, la densité correspondante peut s'écrire $E(r) \varphi$, φ étant la densité euclidienne du plan tangent ⁽¹⁾; le nombre $E(r)$ s'appelle *éclairage* de l'écran au point r ; il se mesure en lumen \otimes mètre⁻² (appelés cette fois-ci *lux*); si l'état est complètement continu, l'éclairage est donné par la formule

$$(16.174) \quad E(r) = \frac{1}{\pi} \int_H \lambda(\xi) \cos I \varphi'(du)$$

H étant l'hémisphère parcouru par le vecteur unitaire u (notations (15.86)) pour l'ensemble des rayons ξ qui arrivent sur l'écran en r ; φ' désigne la densité euclidienne de l'espace vectoriel tangent à la sphère ⁽²⁾; I est l'angle d'incidence ($\cos I = |\langle n, u \rangle|$, n étant un vecteur unitaire normal à l'écran).

(16.175) — Si l'état photométrique est uniforme, on en déduit que l'éclairage est constant et égal à la *luminance* (grâce au facteur $1/\pi$ mis dans la définition (16.171) de la luminance).

(16.176) — Un écran éclairé se comporte en général comme une source; s'il est *parfaitement diffusant*, la brillance d'un rayon issu du point r de l'écran est égale à l'éclairage $E(r)$; une sphère parfaitement diffusante mise dans un éclairage uniforme est donc invisible.

(16.177) — Si l'on observe des sources lumineuses d'état photométrique μ , à travers un instrument d'optique, et si l'on peut négliger l'absorption et les réflexions partielles dans l'instrument, on admet que l'énergie d'un faisceau lumineux (ouvert de V) qui traverse l'instrument est conservée; il en résulte que l'état photométrique des rayons qui ont traversé l'appareil est égal à $S^+(\mu)$ ⁽³⁾, S désignant le symplectomorphisme local de diffusion (voir (15.87) à (15.90)); comme S conserve la densité symplectique (voir (16.99)), on en déduit que la luminance est conservée pour les rayons qui traversent l'instrument ⁽⁴⁾.

⁽¹⁾ Définition (16.94); $\varphi(dr)$ s'appelle « élément d'aire ».

⁽²⁾ $\varphi'(du)$ s'appelle « élément d'angle solide ».

⁽³⁾ Notation (16.149); dans les exemples usuels, S est étalé.

⁽⁴⁾ Mais pas l'éclairage de la rétine, si l'ouvert Ω des rayons traversant l'instrument ne contient pas l'ouvert Ω' des rayons qui traversent la pupille.

VARIABLES ALÉATOIRES

DÉFINITION

(16.178) [Soit V une variété séparée; on appelle *loi de probabilité* de V toute mesure de V positive, bornée, de masse égale à 1.

Soit μ une loi de probabilité sur V , Ω un ouvert de V ; on appelle *probabilité* de l'ensemble Ω le nombre

$$(16.179) \quad p = \mu(\Omega) \quad (1)$$

on vérifie facilement que

$$(16.180) \quad 0 \leq p \leq 1$$

les formules (16.125), (16.126) fournissent quelques propriétés usuelles de la notion de probabilité; mais l'établissement d'une correspondance complète entre les axiomes du calcul des probabilités (Kolmogorov) et la définition (16.179) nécessite une extension de la notion de mesure à des ensembles non ouverts, extension à laquelle nous avons renoncé (16.128); nous nous contenterons d'exploiter directement la définition (16.178).

Exemples

(16.181) Soit a un point d'une variété V ; la mesure de Dirac δ_a est évidemment une loi de probabilité sur V . ■

(16.182) — Soit μ une loi de probabilité définie sur une variété V de dimension 0; si $j \in V$, l'ensemble réduit à j est un ouvert (16.7); notons p_j sa probabilité; on vérifie immédiatement que [notation (16.137)]

$$\diamond \quad \sum_{j \in V} p_j = 1$$

et que, pour tout $g \in \mathcal{K}(V)$

$$\heartsuit \quad \mu(g) = \sum_{j \in V} p_j g(j)$$

la loi de probabilité est donc entièrement définie par les « probabilités élémentaires » p_j .

⁽¹⁾ Définition (16.124); tout ouvert est μ -mesurable, puisque μ est bornée.

DÉFINITION

- (16.183) Soit x une variable à valeurs dans une variété V .
On dit que x est une *variable aléatoire* si on a associé à x une loi de probabilité sur V (on l'appelle « loi de probabilité de x »).

THÉORÈME

- (16.184) Soient V et V^* deux variétés séparées; A une application continue de V dans V^* ; μ une loi de probabilités sur V .
Alors l'image $A^+(\mu)$ est une loi de probabilité sur V .
On sait que $A^+(\mu)$ est une mesure (16.153) positive (16.149) de masse 1 (16.156).

Usage

- (16.185) — Si x est une variable aléatoire ayant μ comme loi de probabilité, on considère encore la variable $A(x)$ comme une variable aléatoire — ayant $A^+(\mu)$ comme loi de probabilité; le théorème (16.151) assure la cohérence de cette règle ⁽¹⁾.

Exemple

- (16.186) Soit $X \equiv \begin{pmatrix} x \\ x' \end{pmatrix}$ une variable aléatoire prenant ses valeurs dans le produit direct $V \times V'$ de deux variétés; puisque les applications $X \mapsto x$, $X \mapsto x'$ sont continues, x et x' sont deux variables aléatoires.

THÉORÈME

- (16.187) Si μ et μ' sont des lois de probabilité sur des variétés V et V' , le produit tensoriel $\mu \otimes \mu'$ est une loi de probabilité sur $V \times V'$.

Vérification immédiate : appliquer (16.60) et (16.123).

- (16.188) — Revenons aux notations (16.186); on dit que x et x' sont des variables aléatoires *indépendantes* si la loi de probabilité de $\begin{pmatrix} x \\ x' \end{pmatrix}$ est le produit tensoriel de deux lois de probabilités μ et μ' ; on vérifie alors ⁽²⁾ que μ et μ' sont les lois de probabilités respectives de x et x' .

⁽¹⁾ Si x est une variable aléatoire, s'il existe des applications continues $x \mapsto y$, $y \mapsto z$, la définition de z comme variable aléatoire au moyen de l'application $x \mapsto z$, ou par l'intermédiaire de y , est la même.

⁽²⁾ Excellent exercice !

- (16.189) — Soit en particulier G un groupe de Lie; si la variable $\begin{pmatrix} x \\ x' \end{pmatrix}$ parcourt G^2 , la variable aléatoire $x \times x'$ parcourt G ; si x et x' sont indépendantes, il est immédiat que la loi de probabilité de $x \times x'$ est la convolutive (16.160) de celles de x et x' ; ceci s'applique en particulier au cas où G est un espace vectoriel, considéré comme groupe additif (« addition des variables aléatoires »).

VALEURS MOYENNES

DÉFINITION

- (16.190) Soit x une variable aléatoire à valeurs dans un espace vectoriel E (de dimension finie); soit μ sa loi de probabilité.
On appelle *valeur moyenne* de x l'élément de E

$$\text{moyenne } x = \mu(1_E)$$

défini si 1_E est μ -intégrable ⁽¹⁾.

THÉORÈME

- (16.191) Soit x une variable aléatoire à valeurs dans une variété V , de loi de probabilité μ ; soit F une application continue de V dans un espace vectoriel E (de dimension finie).
On a alors

$$\text{moyenne } F(x) = \mu(F)$$

si F est μ -intégrable.

$$\text{En effet } F^+(\mu)(1_E) = \mu(1_E \circ F) = \mu(F).$$

Exemple

- (16.192) Soit x une variable aléatoire à valeurs dans un espace vectoriel E . On appelle *moment d'ordre p* de x le tenseur (s'il existe)

$$\text{moyenne } x \otimes x \cdots \otimes x$$

le produit tensoriel comportant p termes ⁽²⁾.

⁽¹⁾ Définition (16.144); 1_E est l'application $[x \mapsto x]$, pas $[x \mapsto 1]$.

⁽²⁾ Il est commode de considérer le nombre 1 comme moment d'ordre 0; la moyenne de x comme moment d'ordre 1. Pour la définition du produit tensoriel $x \otimes x \cdots \otimes x$, voir par exemple J.-M. Souriau, *Calcul linéaire* (P. U. F.), § 13.

DÉFINITION, THÉORÈME

Soit E un espace vectoriel de dimension finie ; Q une partie de E .

— On dit que Q est *convexe* si

$$[x \in Q, y \in Q, 0 \leq s \leq 1] \Rightarrow [sx + [1 - s]y \in Q].$$

- (16.193) \diamond — Toute intersection de parties convexes de E est convexe.
— Soit H une partie de E ; soit Q l'intersection des demi-espaces fermés contenant H :

$$\heartsuit \quad [x \in Q] \Leftrightarrow [\forall \Gamma \in E^*, \Gamma(x) \leq \sup_{y \in H} \Gamma(y)]$$

Q est le plus petit ensemble convexe et fermé contenant H .

La vérification de \diamond est triviale, ainsi que le fait que l'ensemble Q , défini par \heartsuit , contient H , est fermé et convexe. Pour montrer que c'est le plus petit, le lecteur se référera à Bourbaki (*Espaces vectoriels topologiques*, Chap. II), qui traite le cas d'un espace de Banach.

COROLLAIRE (notations de (16.190)).

- (16.194) Si le support de x ⁽¹⁾ est contenu dans un ensemble fermé convexe Q ,
moyenne $x \in Q$ (si moyenne x existe).

Soit H le support de μ ; si $\Gamma \in E^*$, soit $a = \sup_{y \in H} \Gamma(y)$; posons $f(x) = \inf(\Gamma(x), a)$; on a visiblement $f \leq a$ et $[x \in H] \Rightarrow [f(x) = \Gamma(x)]$; en utilisant (16.145), (16.148) on constate donc que

$$\Gamma(\text{moyenne } x) = \Gamma(\mu(1_E)) = \mu(\Gamma \circ 1_E) = \mu(\Gamma) = \mu(f) \leq \mu(a) = a ;$$

moyenne x appartient donc à l'ensemble Q défini par (16.193 \heartsuit), qui est le plus petit fermé convexe contenant H .

C,Q,F,D,

COROLLAIRE (hypothèses de (16.191))

- (16.195) Si Q est un fermé convexe de E , si $\text{val}(F) \subset Q$, et si moyenne $F(x)$ existe :

$$\text{moyenne } F(x) \in Q.$$

Il suffit d'appliquer (16.152).

⁽¹⁾ Abus de langage désignant le support de la loi de probabilité de x .

Exemple

(16.196) Soit x une variable aléatoire. Posons $F(x) \equiv \begin{pmatrix} x \\ x \otimes x \end{pmatrix}$; on vérifie que

l'ensemble Q des $\begin{pmatrix} {}^1M \\ {}^2M \end{pmatrix}$, où 1M et 2M sont des tenseurs contravariants de E , respectivement de degré 1 et 2, tels que ⁽¹⁾

$$1 + 2 {}^1M(C) + {}^2M(C)(C) \geq 0 \quad \forall C \in E^*$$

est un ensemble fermé convexe contenant $\text{val}(F)$; on a donc

$$\text{moyenne } F(x) \in Q,$$

ce qui s'écrit

$$1 + 2 {}^1M(C) + {}^2M(C)(C) \geq 0 \quad \forall C \in E^*$$

1M et 2M désignant les *moments* d'ordre 1 et 2 de x ; on peut aussi écrire cette propriété sous la forme

$$\Phi(C)(C) \geq 0 \quad \forall C \in E^*,$$

Φ étant le tenseur ${}^2M - {}^1M \otimes {}^1M$; Φ s'appelle tenseur *variance* de x .

ENTROPIE. LOI DE GIBBS

DÉFINITION

a) Soit V une variété séparée ; soit $x \mapsto \varphi$ un champ continu de densités, tel que $\varphi > 0 \quad \forall x$ ⁽²⁾ ; soit λ la mesure correspondante (16.83).

b) Si μ est une loi de probabilité complètement continue de V , il existe une fonction $f \in \mathcal{C}(V)$ telle que

$$f \geq 0 ; \quad \mu = \lambda \times f$$

(16.197) on appelle λ -entropie de la loi de probabilité μ le nombre

$$s = \begin{cases} \lambda(H \circ f) & \text{si } H \circ f \text{ est } \lambda\text{-intégrable} \\ -\infty & \text{si } H \circ f \text{ n'est pas intégrable} \end{cases}$$

H étant la fonction

$$H(\rho) = \begin{cases} 0 & \text{si } \rho = 0 \\ -\rho \log \rho & \text{si } \rho > 0. \end{cases}$$

⁽¹⁾ Pour les notations, voir la référence de la note ⁽²⁾ (p. 267).

⁽²⁾ Nous dirons que $x \mapsto \varphi$ est *strictement positif*.

On peut donc — avec quelques abus de notation — écrire

(16.198)

$$s = \int_V -\rho \log \rho \varphi(dx) \quad [\rho \equiv f(x)].$$

— Il est utile de connaître le graphe de H ; il est donné par la figure (16. II).

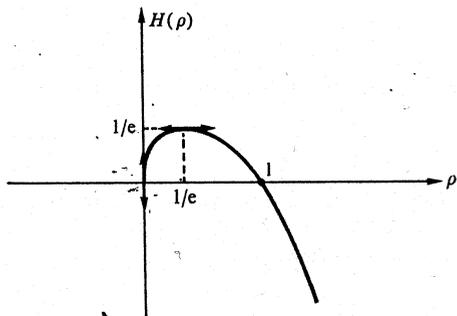


Fig. 16. II.

DÉFINITION, THÉORÈME

(16.199)

Soit V une variété séparée; supposons donnés :

- Un champ de densités $x \mapsto \varphi$, continu et strictement positif, définissant une mesure λ de V ;
- Une application continue Ψ de V dans un espace vectoriel E (de dimension finie).

(16.200)

1) Nous appellerons *loi de Gibbs* (généralisée) toute loi de probabilité ζ telle que

$$\diamond \quad \begin{cases} \exists z \in \mathbb{R}, \exists Z \in E^*; \zeta = \lambda \times f; f(x) \equiv e^{-[z + Z(\Psi(x))]} \\ \Psi \text{ est } \zeta\text{-intégrable.} \end{cases}$$

Alors la λ -entropie de cette loi existe, et vaut

♡

$$s = z + Z(M)$$

en posant

♣

$$M = \zeta(\Psi)$$

2) Toute *autre* loi de probabilité complètement continue ζ' , vérifiant aussi

$$\zeta'(\Psi) = M$$

a une entropie strictement inférieure à celle de la loi de Gibbs.

Puisque la courbe représentative de la fonction H (Fig. 16. II) a sa concavité tournée vers le bas (la dérivée seconde est $H''(\rho) = -1/\rho$) elle est située en dessous de sa tangente en un point ρ_0 ; ce qui s'écrit

$$[\rho \geq 0, \rho_0 > 0, \rho \neq \rho_0] \Rightarrow [H(\rho) < H(\rho_0) + [\rho - \rho_0] H'(\rho_0)];$$

en remplaçant H' par sa valeur, cette inégalité devient

$$H(\rho) < \rho_0 - \rho - \rho \log(\rho_0);$$

il en résulte, si deux lois de probabilités complètement continues ont comme densités respectives $\rho \varphi, \rho_0 \varphi$, que l'entropie s de la première vérifie ⁽¹⁾

$$s \leq \int_V -\rho \log(\rho_0) \varphi(dx).$$

Si la seconde loi est une loi de Gibbs :

$$\rho_0 \equiv e^{-[z + Z(\Psi(x))]}$$

on a donc $s \leq \int_V \rho [z + Z(\Psi(x))] \varphi(dx) = z + Z(M)$,

l'égalité ne pouvant avoir lieu que si $\rho \equiv \rho_0$ (Théorème (16.84)).

C.Q.F.D.

— Soit ζ une loi de Gibbs, définie à partir de z et Z (notations (16.200)); en écrivant que masse (ζ) = 1 et que Ψ est ζ -intégrable, on constate que les deux intégrales :

$$I_0 = \int_V e^{-Z(\Psi(x))} \varphi(dx)$$

$$I_1 = \int_V \Psi(x) e^{-Z(\Psi(x))} \varphi(dx)$$

(16.201)

(16.202)

⁽¹⁾ On utilise le fait que les deux lois ont la même masse 1, donc que

$$\int_V [\rho_0 - \rho] \varphi(dx) = 0.$$

sont convergentes, et que

(16.203)

$$z = \log(I_0)$$

(16.204)

$$M = \frac{I_1}{I_0}$$

(16.205)

Réciproquement, si on se donne Z dans E^* , si les intégrales I_0 et I_1 sont convergentes, on voit que I_0 est positive, qu'on peut donc définir un nombre z par (16.203); alors la densité

$$e^{-[z + Z(\Psi(x))]} \varphi$$

est bien celle d'une loi de Gibbs.

Exemple : (loi normale) :

(16.206)

Prenons le cas $V = R^n$, λ = mesure de Lebesgue; $\Psi(x) \equiv \begin{pmatrix} x \\ x \otimes x \end{pmatrix}$;

un élément Z du dual de E peut se définir par la formule

$$Z(\Psi(x)) \equiv \bar{a} \cdot x + \frac{1}{2} \bar{x} \cdot H \cdot x$$

[$\bar{a} \in R^n$; H = matrice symétrique]. On vérifie que la convergence de l'intégrale I_0 a lieu si la matrice H est positive ⁽¹⁾; dans ce cas la loi de Gibbs s'appelle *loi normale de Gauss*; on calcule facilement I_0 en faisant le changement de variable $x^* = H^{1/2} x + H^{-1/2} \bar{a}$ ⁽²⁾; il vient

(16.207)

$$z = \frac{1}{2} [\bar{a} \cdot H^{-1} \cdot \bar{a} - \log(\det(H)) + n \log(2\pi)]$$

alors la convergence de I_1 a lieu également; on peut donc calculer M , qui est défini par les moments du premier et du second ordre de la loi (16.196); le calcul montre que le moment du premier ordre est égal à $-H^{-1} \cdot \bar{a}$ et que les composantes du tenseur *variance* (16.196) sont égales aux éléments de la matrice H^{-1} ; le moment du second ordre s'en déduit immédiatement.

La formule (16.200) donne l'entropie :

(16.208)

$$s = \frac{n}{2} \log(2\pi e) - \frac{1}{2} \log(\det(H)) ;$$

⁽¹⁾ Voir *Calcul linéaire*, tome II.

⁽²⁾ C'est-à-dire en recherchant l'image de la loi par l'application $x \mapsto x^*$.

nous pouvons donc associer, à toute loi de probabilité de R^n ayant une fluctuation régulière, une loi normale ayant les mêmes moments d'ordre 1 et 2; la loi donnée a une entropie inférieure à la valeur (16.208), à moins qu'elle ne soit déjà normale.

THÉORÈME

(16.209)

La correspondance $Z \mapsto M$ définie par (16.204) est *injective*, à moins que l'ensemble de valeurs de F soit contenu dans un sous-espace affine de dimension inférieure à celle de E .

Supposons en effet que deux valeurs différentes Z, Z' conduisent au même M ; le théorème (16.200) montre que chacune de ces deux lois donne l'entropie maximum compatible avec cette valeur de M ; les entropies sont donc égales, et par conséquent les densités sont les mêmes; on a donc $z + Z(\Psi(x)) \equiv z' + Z'(\Psi(x))$, ce qui montre bien que $\text{val}(\Psi)$ est contenu dans le sous-espace des y tels que

$$[Z - Z'](y) = z' - z.$$

THÉORÈME (hypothèses (16.200))

a) Soit Ω un ouvert de E^* ; supposons qu'il existe une fonction $f \in \mathcal{C}(V)$, telle que

(16.210)

$$e^{-Z(\Psi(x))} \leq f(x) \quad \forall x \in V, \quad \forall Z \in \Omega$$

et que les intégrales

$$I_0' = \int_V f(x) \varphi(dx)$$

$$I_1' = \int_V \Psi(x) f(x) \varphi(dx)$$

convergent ⁽¹⁾. Alors $[Z \mapsto I_0]$ est une fois différentiable dans Ω ; on peut différentier I_0 sous le signe \int :

$$\diamond \quad \delta I_0 \equiv \int_V -[\delta Z](\Psi(x)) e^{-Z(\Psi(x))} \varphi(dx).$$

b) Si de plus l'intégrale

$$I_2' = \int_V [\Psi(x) \otimes \Psi(x)] f(x) \varphi(dx)$$

⁽¹⁾ Nous dirons dans ce cas que les intégrales I_0, I_1 convergent uniformément dans Ω .

(16.210) converge, on peut aussi différentier I_1 sous le signe \int :

$$\heartsuit \quad \delta I_1 \equiv \int_V -\Psi(x) [\delta Z] (\Psi(x)) e^{-Z(\Psi(x))} \varphi(dx).$$

Nous admettrons ce résultat. ■

— Plaçons-nous dans ces hypothèses (16.210a); en comparant la formule \heartsuit avec (16.203, 204), on trouve

$$(16.211) \quad \delta z \equiv - [\delta Z] (M)$$

ce que l'on peut encore écrire (en identifiant E avec son bidual $[E^*]^*$:

$$(16.212) \quad M = - \frac{\partial z}{\partial Z}$$

■ Cette formule peut se vérifier sur l'exemple (16.207) ⁽¹⁾. ■

— Supposons les conditions (16.210b) vérifiées; alors M est fonction une fois différentiable de Z ; on tire facilement de (16.210 \heartsuit) la formule

$$(16.213) \quad [\delta Z] (\delta M) = - \frac{1}{I_0} \int_V \left\{ [\delta Z] \left(\Psi(x) - \frac{I_1}{I_0} \right) \right\}^2 e^{-Z(\Psi(x))} \varphi(dx)$$

d'où découle l'inégalité

$$(16.214) \quad [\delta Z] (\delta M) \leq 0$$

et même, si l'ensemble de valeurs de F n'est pas contenu dans un sous-espace affine de dimension inférieure à celle de E

$$(16.215) \quad [\delta Z \neq 0] \Rightarrow [[\delta Z] (\delta M) < 0]$$

Il en résulte évidemment que $\partial M / \partial Z$ est régulier; comme $[Z \mapsto M]$ est injectif, on voit que $[Z \mapsto M]$ est un difféomorphisme ⁽²⁾ (de l'ouvert de E^*

⁽¹⁾ L'ensemble des Z tels que I_0 converge est alors un ouvert, qui peut être recouvert par des ouverts Ω_j où les conditions (16.210a et b) sont vraies.

⁽²⁾ En un sens restreint puisque nous ne sommes pas assurés de l'existence des dérivées d'ordre ≥ 2 ; mais les théorèmes permettant l'inversion s'appliquent encore.

où les conditions d'existence et de différentiabilité sont vérifiées sur un ouvert de E); alors $[M \mapsto z]$ et $[M \mapsto s]$ sont différentiables (une fois au moins); en joignant (16.200 \heartsuit) et (16.211), on a immédiatement

$$(16.216) \quad \delta s \equiv Z(\delta M)$$

ce qui peut s'écrire

$$(16.217) \quad Z \equiv \frac{\partial s}{\partial M}$$

ENSEMBLE DE GIBBS D'UN GROUPE DYNAMIQUE

HYPOTHÈSES

Soit U une variété symplectique, G un groupe dynamique de U ⁽¹⁾; nous supposons que :

- 1) U est une variété séparée;
- 2) G est un groupe de Lie connexe ⁽²⁾;
- 3) G possède un moment μ :

(16.218) il existe donc une application $\Psi[x \mapsto \mu]$ de U dans le dual \mathcal{G}^* de l'algèbre de Lie \mathcal{G} de G ⁽³⁾ telle que

$$\sigma(Z_v(x)) \equiv - \nabla[\mu, Z] \quad \forall Z \in \mathcal{G} \text{ (Z constant)}$$

σ étant la forme de Lagrange de U .

Nous savons que U possède une mesure λ — la mesure de Liouville — définie à partir de la densité symplectique \mathcal{G} (voir (16.96)).

Cette mesure est absolument continue et strictement positive : elle vérifie les hypothèses (16.197); toute loi de probabilité complètement continue de U possède donc un Liouville-entropie s (nous l'appellerons simplement entropie). Nous pouvons aussi construire les lois de Gibbs formées à partir de l'application Ψ ; ce qui conduit à l'énoncé suivant (hypothèses (16.218)) :

⁽¹⁾ Définition (11.1).

⁽²⁾ On rappelle que tout groupe dynamique est un groupe de Lie, nécessairement séparé (6.12).

⁽³⁾ Rappelons que les éléments de \mathcal{G}^* s'appellent *torseurs* de G . Des conditions suffisantes d'existence d'un moment ont été données en (11.8).

1) Soit Ω le plus grand ouvert de \mathcal{G} tel que $\forall Z \in \Omega$, les intégrales

(16.219)

◇

$I_0 = \int_U e^{-\mu Z} \varphi(dx)$
$I_1 = \int_U \mu e^{-\mu Z} \varphi(dx)$
$I_2 = \int_U \mu \otimes \mu e^{-\mu Z} \varphi(dx)$

convergent, et que les applications $Z \mapsto I_0, Z \mapsto I_1$ soient une fois différentiables sous le signe \int ⁽¹⁾.

Si $Z \in \Omega$, nous poserons

♡

$z \equiv \log(I_0)$
$M \equiv \frac{I_1}{I_0}$

♠

2) Si $Z \in \Omega$, il existe une loi de probabilité ζ , définie par la densité

♣

$e^{-[z+\mu Z]} \varphi$	(loi de Gibbs)
--------------------------	----------------

elle vérifie ⁽²⁾

(

moyenne $\mu = M$

son entropie est

σ

$s = z + M \cdot Z$

parmi toutes les lois de probabilité complètement continues vérifiant (, la loi ♣ est celle qui a la plus grande entropie.

⁽¹⁾ Nous avons donné en (16.210) des conditions suffisantes; il suffit qu'elles soient vérifiées sur les éléments d'un recouvrement de Ω .

⁽²⁾ L'écriture correcte est $\zeta(\Psi) = M$; on suppose implicitement que l'on a fait de x une variable aléatoire, admettant la loi de probabilité ζ ; alors la formule (résulte de (16.191).

(16.219)

3) Les applications $Z \mapsto z, Z \mapsto s, Z \mapsto M$ sont une fois différentiables, et l'on a

♀

$\frac{\partial z}{\partial Z} \equiv -M$	$\delta M \cdot \delta Z \leq 0 \quad \forall \delta Z$
---	---

Ces résultats sont une simple transcription des résultats (16.200-214), à un changement de notation près: puisque l'ensemble de valeurs de Ψ est le dual \mathcal{G}^* d'un espace vectoriel \mathcal{G} , nous identifions tout élément Z du bidual \mathcal{G}^{**} avec un élément de \mathcal{G} , en écrivant μZ au lieu de $Z(\mu)$.

THÉORÈME (suite de (16.219))

— Supposons de plus que G opère *effectivement* sur U , c'est-à-dire que

$$[a \in G, a_U = 1_U] \Rightarrow [a = \text{élément neutre de } G].$$

Alors l'application $Z \mapsto M$ est injective; son ensemble de valeurs est un ouvert Ω^* de \mathcal{G}^* ; les applications

(16.220)

$$M \mapsto z, \quad M \mapsto s, \quad M \mapsto Z$$

sont une fois différentiables, et l'on a

♂

$\frac{\partial s}{\partial M} \equiv Z$ ⁽¹⁾	$\delta M \cdot \delta Z < 0 \quad (\forall \delta Z \neq 0)$
---	---

Grâce aux théorèmes (16.209), (16.215), il suffit de montrer que l'ensemble de valeurs de Ψ n'est pas contenu dans un espace affine de dimension inférieure à celle de \mathcal{G}^* , c'est-à-dire qu'il n'existe pas $\delta Z \in \mathcal{G}, c \in R$ tels que

$$\mu \delta Z \equiv c \quad (\delta Z \neq 0, \delta Z \text{ et } c \text{ constants}).$$

en utilisant la définition du moment (rappelée en (16.218)) on voit qu'on aurait $\sigma(Z_U(x)) \equiv 0$; soit, puisque σ est régulier (9.1), $Z_U(x) = 0$; il résulte immédiatement de (6.12) que pour tout α réel, l'élément de $G: a \equiv \exp(\alpha Z_U)$ (e) vérifierait $a_U \equiv 1_U$; d'où $a \equiv e$ grâce à (16.220); ce qui est impossible, puisque, selon (3.1),

$$\frac{da}{d\alpha} \equiv Z_G(a) \neq 0.$$

(16.221)

— On sait que le moment μ n'est défini par le groupe dynamique qu'à une constante additive près (Théorème (11.8)); si on remplace μ par $\mu - \mu_0$, on constate que la définition (16.219, 1)) de l'ouvert Ω ne change pas; si $Z \in \Omega$, que les variables z, M, s subissent la substitution:

◇

$$z \rightarrow z + \mu_0 Z; \quad M \rightarrow M - \mu_0; \quad s \rightarrow s$$

⁽¹⁾ Abus de notations signifiant $\delta s \equiv \delta M \cdot Z$.

(16.221) la loi de probabilité Φ associée à Z ne change pas; l'ensemble de ces lois de probabilités ne dépend donc pas du choix du moment μ ; on l'appellera ensemble canonique de Gibbs du groupe dynamique G .

— Soit a un élément du groupe G (Fig. 16. III); puisque \underline{a}_G^{-1} est un symplectomorphisme, l'image par \underline{a}_G^{-1} de la mesure de Liouville λ est égale à λ (Théorème de Liouville (16.99); si $Z \in \Omega$, l'intégrale I_0 peut s'écrire $\lambda(x \mapsto e^{-\Psi(x)Z})$; le théorème (16.150) montre donc que l'on a

◇
$$I_0 = \lambda(x \mapsto e^{-\Psi(\underline{a}_G^{-1}(x))Z}) ;$$

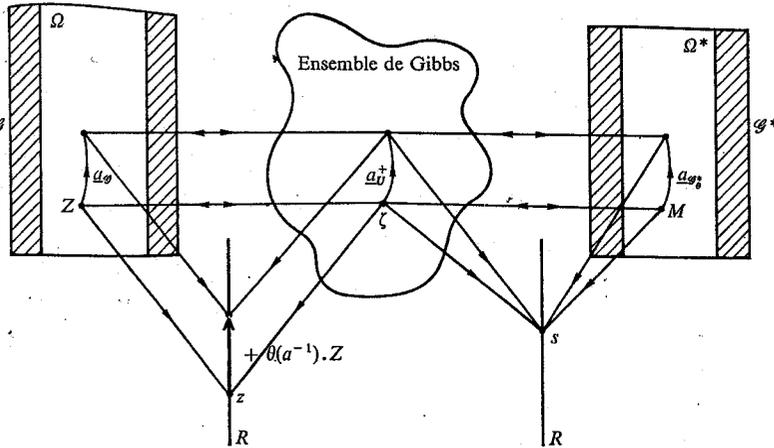
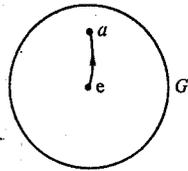


Fig. 16. III.

or nous savons que l'on a

$$\Psi(\underline{a}_G^{-1}(x)) \equiv \underline{a}_G^{-1}(\Psi(x)) + \theta(a^{-1})$$

θ étant un cocycle symplectique de G (11.17 ◇); en portant dans ◇, et en posant

♡
$$Z' = \underline{a}_G(Z)$$

on voit (grâce à (11.15)) que l'intégrale I_0 correspondant à la valeur ♡ de Z est convergente, et que l'on a

♣
$$I_0' = e^{\theta(a^{-1})Z} I_0.$$

Un calcul analogue montre que les intégrales I_1' et I_2' convergent aussi, et que les applications $Z' \mapsto I_0'$, $Z' \mapsto I_1'$ sont dérivables sous le signe \int ; en utilisant les formules ♣, (16.219), (11.17 ♡), (16.86), on en déduit le

THÉORÈME (suite de (16.219); Fig. 16. III)

— L'ensemble Ω est invariant par toute transformation \underline{a}_G ($a \mapsto \underline{a}_G$ étant la représentation adjointe de G); les variables (16.219) subissent alors la substitution

(16.222)
$$\begin{aligned} Z &\rightarrow \underline{a}_G(Z) \\ z &\rightarrow z + \theta(a^{-1})Z \equiv z - \theta(a)\underline{a}_G(Z) \quad (1) \\ s &\rightarrow s \\ M &\rightarrow \underline{a}_G(M) + \theta(-a) = \underline{a}_G(M) \quad (2) \\ \zeta &\rightarrow \underline{a}_G^+(\zeta) \quad (3). \end{aligned}$$

Formulons un lemme purement algébrique :

Soit f un cocycle symplectique d'une algèbre de Lie \mathcal{G} , c'est-à-dire une 2-forme de \mathcal{G} vérifiant :

(16.223)
$$f(Z_1)([Z_2, Z_3]) + f(Z_2)([Z_3, Z_1]) + f(Z_3)([Z_1, Z_2]) \equiv 0.$$

Soit Z un élément du noyau de f ; soit E l'ensemble de valeurs de $\text{ad}(Z)$. Il existe alors un tenseur symétrique g de E défini par

$$g([Z, Z']) (Z'') = f(Z') (Z'') \quad \forall Z' \in \mathcal{G} \quad \forall Z'' \in E.$$

La vérification est élémentaire.

THÉORÈME (suite de (16.222))

Soit f la dérivée de θ en l'élément neutre; posons (notation (6.13))

$$f_Z \equiv f + M \cdot \text{Ad} \quad \forall Z \in \Omega.$$

Alors :

a) f_Z est un cocycle symplectique de \mathcal{G} , qui ne dépend pas du choix du moment μ de G ;

b)
$$f_Z(Z) = 0 \quad \forall Z \in \Omega ;$$

(1) Le cocycle θ , associé au groupe G et au moment μ , a été défini en (11.17).

(2) Notation (11.28).

(3) $\underline{a}_G^+(\zeta)$ est l'image par \underline{a}_G de la loi de probabilité ζ (notation (16.86) ou (16.149)).

(16.224) c) Il existe un tenseur symétrique g , défini sur l'ensemble de valeurs E de $\text{Ad}(Z)$, tel que

$$\diamond \quad g([Z, Z']) (Z'') = f_Z(Z') (Z'') \quad \forall Z' \in \mathcal{G}, \quad \forall Z'' \in E$$

il vérifie

$$\heartsuit \quad g(Z'') (Z'') \leq 0 \quad \forall Z'' \in E.$$

d) Si E n'est pas réduit à 0, et si G opère *effectivement* sur U , g donne à E une structure d'espace euclidien négatif.

— On sait que f est un cocycle symplectique (11.33); donc aussi f_Z , grâce à l'identité de Jacobi.

— Si on change le moment μ , en le remplaçant par $\mu - \mu_0$ (Théorème (11.8)), il résulte des formules (11.18), (11.16), (16.221 \diamond) que f_Z ne change pas; d'où (a).

— Différentions l'identité

$$[Z \mapsto z] (a_{\mathcal{G}}(Z)) \equiv z - \theta(a) a_{\mathcal{G}}(Z)$$

qui est une conséquence immédiate de (16.222), en prenant $a = e$, $\delta a = Z'$, $\delta Z = 0$; compte tenu de (6.11), (6.26), (16.219 \heartsuit), il vient

$$M[Z', Z] = -f(Z') (Z),$$

d'où (b).

L'existence du tenseur g vérifiant \diamond en découle immédiatement (par application du lemme (16.223)).

— Différentions comme ci-dessus l'identité

$$[Z \mapsto M] (a_{\mathcal{G}}(Z)) \equiv a_{\mathcal{G}}(M) + \theta(a)$$

autre conséquence de (16.222); on trouve

$$\frac{\partial M}{\partial Z} (-[Z', Z]) \equiv f(Z') + M \text{ ad}(Z') \equiv f_Z(Z').$$

Posons $dZ = [Z, Z'] = Z''$; la formule (16.219 \heartsuit)

$$dM \cdot dZ \geq 0$$

s'écrit donc $f_Z(Z') (Z'') \geq 0$, soit \heartsuit . De même, si $Z'' \neq 0$, et si G opère *effectivement*, l'inégalité $dM \cdot dZ < 0$ (16.220) montre que $g(Z'') (Z'') < 0$; g est nécessairement régulier.

C.Q.F.D.

§ 17 LES PRINCIPES DE LA MÉCANIQUE STATISTIQUE

ÉTATS STATISTIQUES

(17.1) Soit U l'espace des mouvements d'un système dynamique. En mécanique classique ou relativiste, l'état du système (c'est-à-dire la description complète de son histoire) est caractérisé par un point x de U .

On appelle *état statistique* du système une *loi de probabilité* μ sur la variété U ⁽¹⁾.

Autrement dit : en mécanique statistique, le mouvement x d'un système dynamique est une *variable aléatoire* (Définition (16.183)), dont la loi de probabilité définit l'*état statistique* du système.

(17.2) Quelle interprétation doit-on donner à un état statistique ? Les opinions à ce sujet sont assez diverses ⁽²⁾; nous n'envisagerons cette question qu'après avoir étudié des exemples (ci-dessous (17.52)).

HYPOTHÈSES DE LA THÉORIE CINÉTIQUE DES GAZ

(17.3) Le problème initial de la Mécanique Statistique fut l'étude des gaz; la « théorie cinétique » consiste à décrire un gaz comme un *système*, avec un

⁽¹⁾ On définit souvent (voir par exemple G. Mackey, *Mathematical Foundations of Quantum Mechanics*) un état statistique au moyen d'une loi de probabilité μ_t , définie sur l'espace de phases U_t à l'instant t (définition (12.31)); on postule en plus une condition (formulée au moyen d'une équation aux dérivées partielles dans le cas d'une loi complètement continue) qui revient à dire que l'*image* $P^+(\mu_t)$ de cette loi, par la projection P sur l'espace des mouvements U (Fig. 12.1) est *indépendante* de t ; c'est cette image μ que nous considérons directement.

⁽²⁾ L'état statistique représenterait :

- une moyenne faite sur des expériences nombreuses répétées;
- une simplification pour l'étude des systèmes dont l'espace des mouvements a une dimension très grande;
- une évaluation de l'influence d'un « grand » système sur un « petit » auquel il est faiblement couplé;
- une mesure de notre connaissance partielle de l'état réel d'un système (théorie de l'information);
- une description réaliste de la nature (ce point de vue doit nécessairement être rapproché du point de vue quantique; voir (19.17)).

espace des mouvements de dimension grande, mais finie ; ce système étant décrit, de façon microscopique, au moyen de molécules (elles-mêmes traitées le plus souvent comme des points matériels ou des solides) agissant entre elles et les parois de la boîte supposée contenir le gaz.

(17.4) — On peut traiter ces actions par la *théorie de la diffusion* ; supposons, pour simplifier, que les actions directes sur les molécules (champ électromagnétique, gravitation) soient négligeables, et que les interactions se réduisent à des collisions instantanées ⁽¹⁾ ; les symplectomorphismes de diffusion peuvent alors être indexés par le temps ⁽²⁾. Rappelons le formalisme (15.61) :

Si x est un mouvement du gaz qui ne comporte pas de collision à la date t , on écrit

$$(17.5) \quad x = A_t(x^*)$$

x^* étant le mouvement qu'auraient les molécules si elles étaient toutes libres, avec les positions et les vitesses qu'elles ont effectivement à la date t ; on pose

$$(17.6) \quad S_t^{t'} = A_t^{-1} \cdot A_{t'} \quad \forall t, t' \in R$$

$S_t^{t'}$ est un symplectomorphisme local de U^* à U^* (U^* désignant la variété des mouvements du système des molécules libres) qui vérifie les relations (15.67) :

$$(17.7) \quad \begin{cases} S_t^t = 1_{E_t} \\ S_t^{t'} = [S_{t'}^t]^{-1} \\ S_t^t \circ S_{t'}^{t''} < S_{t'}^{t''} \end{cases}$$

(17.8) E_t est l'ouvert de U^* composé des mouvements libres tels que, à l'instant t , toutes les molécules soient à l'intérieur de la boîte, sans se toucher les unes les autres.

(17.9) — On sait que la variété des mouvements du gaz est entièrement définie par la connaissance des $S_t^{t'}$ (voir (15.68)).

— Le groupe de Galilée opère sur la variété U^* ; par conséquent aussi le sous-groupe à une dimension des translations dans le temps : nous note-

⁽¹⁾ Nous nous plaçons dans l'hypothèse où il existe un *libre parcours* pour les molécules ; l'hypothèse selon laquelle les collisions sont instantanées n'est pas indispensable pour appliquer la théorie de la diffusion.

La théorie cinétique envisage aussi le cas où les actions mutuelles des molécules ne sont pas négligeables à distance ; ayant formulé des hypothèses sur ces interactions, on recherche notamment si on peut trouver des états d'équilibre à n molécules ayant — en un certain sens — une limite lorsque n tend vers l'infini.

⁽²⁾ Voir par exemple le cas du miroir plan (15.110).

rons simplement \underline{e}_U le symplectomorphisme \underline{e}_U , avec (notations (13.3)) :

$$(17.10) \quad a \equiv \begin{bmatrix} 1_{R^3} & 0 & 0 \\ 0 & 1 & e \\ 0 & 0 & 1 \end{bmatrix} \quad e \in R$$

l'application $e \mapsto \underline{e}_U$ permet de considérer le groupe additif R comme groupe dynamique de U ; si x^* est un mouvement des molécules libres, $\underline{e}_U(x^*)$ est le mouvement qui s'en déduit par un retard de e .

(17.11) — Supposons la boîte immobile : on admet alors que les $S_t^{t'}$ commutent avec les \underline{e}_U (c'est-à-dire que les lois de collision ne sont pas modifiées par un retard) ; il en résulte que les translations dans le temps forment un *groupe de symétrie* (voir (15.69) à (15.73)) pour la diffusion ; R opère aussi sur les mouvements du gaz, de sorte que

$$\underline{e}_U \cdot A_t = A_t \cdot \underline{e}_U \quad (\forall t, e \in R)$$

$\underline{e}_U(x)$ est évidemment le mouvement du gaz déduit du mouvement x par un retard e .

— Alors le gaz peut être considéré comme un *système conservatif* — au sens (12.153) ; le moment du groupe des translations dans le temps est simplement le *hamiltonien* h (ou si on préfère l'énergie), au signe près ; ce qui s'écrit :

$$(17.12) \quad \frac{\partial}{\partial t} [L_U(x)] \equiv F(L_U(x))$$

F étant le champ de vecteurs

$$(17.13) \quad F(x) \equiv - \text{grad } h$$

défini par le gradient symplectique de la variable dynamique h .

ÉQUILIBRES D'UN SYSTÈME CONSERVATIF

(17.14) — Considérons un tel système conservatif ; on appelle *équilibre* du système un mouvement x qui est *invariant par le groupe des translations dans le temps* :

$$L_U(x) = x \quad [\forall t \in R]$$

on voit immédiatement que les équilibres sont les points de U où $\text{grad } h$ s'annule.

De même, on appelle *équilibre statistique* du système un *état statistique* μ invariant par le groupe, c'est-à-dire tel que

$$\underline{t}_U^+(\mu) = \mu \quad [\forall t \in \mathbb{R}];$$

ici \underline{t}_U^+ désigne l'image de la mesure μ par le difféomorphisme \underline{t}_U , image qui est encore une loi de probabilité (16.184); cette définition contient la précédente, si l'on identifie un mouvement x avec la loi de probabilité définie par la mesure de Dirac δ_x ⁽¹⁾.

(17.15) Pour rechercher des équilibres, on peut songer à prendre un mouvement arbitraire x , et à le laisser « vieillir »; il revient évidemment au même de considérer le même mouvement *avancé* d'une durée t , et de chercher s'il existe une limite de ce mouvement — $\underline{t}_U(x)$ lorsque t tend vers $+\infty$.

On peut de même se demander, μ étant un état statistique, s'il existe une limite μ^* à $-\underline{t}_U^+(\mu)$, et si celle-ci est un équilibre statistique.

Comment *définir* la « limite » μ_∞ de cette mesure? On peut envisager la *limite en norme* définie par

$$(17.16) \quad \left\| -\underline{t}_U^+(\mu) - \mu_\infty \right\| \rightarrow 0$$

la *limite vague*, définie par

$$(17.17) \quad -\underline{t}_U^+(\mu)(g) \rightarrow \mu_\infty(g) \quad \forall g \in \mathcal{K}(U)$$

et la *limite vague en moyenne*, définie par

$$(17.18) \quad \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T [-\underline{t}_U^+(\mu)(g)] dt \rightarrow \mu_\infty(g) \quad \forall g \in \mathcal{K}(U).$$

(17.19) Il est facile de montrer que chacune de ces conditions entraîne la suivante; en utilisant le théorème de Birkhoff ⁽²⁾, on peut formuler des conditions dans lesquelles la limite vague en moyenne existe, et constitue un état d'équilibre statistique: la masse de la limite μ_∞ est alors égale à la masse de μ — à savoir 1.

On peut aussi montrer, dans certains cas, que la *valeur moyenne du hamiltonien* h (supposé μ -intégrable) est conservée à la limite:

$$\mu_\infty(x \mapsto h) = \mu(x \mapsto h).$$

⁽¹⁾ En effet $A^+(\delta_x) = \delta_{A(x)}$ pour tout homéomorphisme A .

⁽²⁾ G. D. Birkhoff, *Dynamical Systems* (1927); voir aussi A. Blanc-Lapierre, P. Casal, A. Tortrat, *Méthodes mathématiques de la mécanique statistique* (1959).

(17.20)

Si μ est une mesure complètement continue, on vérifie facilement que l'entropie de $\underline{t}_U^+(\mu)$ est *constante* et égale à celle de μ ; mais l'entropie de la limite vague μ_∞ peut être *différente* de celle de μ ⁽¹⁾; si μ n'est pas un état d'équilibre, on peut montrer — avec certaines hypothèses — que l'entropie de μ_∞ est *plus grande* que celle de μ ⁽²⁾; nous allons seulement donner un argument heuristique (17.21) et un exemple (17.22).

(17.21)

Supposons μ complètement continue; en la mettant sous la forme

$$\mu = \mu_0 \times f$$

μ_0 désignant la mesure de Liouville de U (16.96), l'entropie de μ est par définition

$$\mu_0(H \circ f)$$

H étant la fonction définie en (16.197). Désignons par μ_T la mesure définie par

$$\mu_T(g) = \frac{1}{T} \int_0^T [-\underline{t}_U^+(\mu)(g)] dt \quad [\forall g \in \mathcal{K}(U)]$$

on voit facilement (grâce au théorème de Liouville (16.99)) que

$$\mu_T = \mu_0 \times f_T.$$

f_T étant la fonction continue définie par

$$f_T(x) = \frac{1}{T} \int_0^T f(\underline{t}_U(x)) dt;$$

on en déduit la formule

$$\mu_{2T} = \frac{1}{2}[\mu_T + \underline{T}_U(\mu_T)].$$

Il est immédiat que les mesures μ_T et $\underline{T}_U(\mu_T)$ ont la même entropie; si ces deux mesures ne sont pas égales, il en résulte que l'entropie de μ_{2T} est supérieure à celle de μ_T ; en effet, si ρ et ρ' sont les densités de deux lois de probabilité ⁽³⁾, l'entropie

de la demi-somme s'écrit $\int_U H\left(\frac{\rho + \rho'}{2}\right) \varphi(dx)$; comme H est convexe (Fig. 16. II), cette entropie est supérieure à $\int_U \frac{1}{2}[H(\rho) + H(\rho')] \varphi(dx)$, donc à la demi-somme des entropies (égales dans le cas considéré). ■

⁽¹⁾ En d'autres termes: l'entropie n'est pas continue pour la topologie vague, ainsi que va le montrer l'exemple (17.22).

⁽²⁾ Comparer avec le célèbre « théorème H » de Boltzmann, énoncé certainement faux qui affirme que l'entropie est fonction croissante du temps quand il n'y a pas d'équilibre.

⁽³⁾ φ désignera désormais la densité symplectique de U (16.94).

Exemple

(17.22) On considère une variété symplectique U de dimension 2, munie d'une carte canonique $\begin{pmatrix} p \\ q \end{pmatrix} \mapsto x$; on fait opérer R sur U au moyen du hamiltonien $h \equiv [p^2 + q^2]^2$. Soit μ une loi de probabilité sur U :

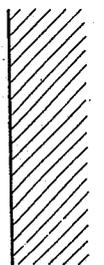
$$\mu = \mu_0 \times f,$$

μ_0 étant la densité de Liouville et f une fonction de la forme ⁽¹⁾

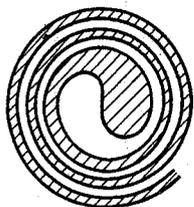
$$f(x) \equiv p^+ F(h)$$

la fonction positive F étant choisie pour que masse $(\mu) = 1$ et pour que h soit μ -intégrable (Fig. 17.1).

On peut vérifier que la limite vague de $[-\frac{t}{h}]^+(\mu)$ existe, et que c'est la loi de probabilité $\mu_\infty = \mu_0 \times f^*$ [$f^*(x) \equiv \frac{1}{\pi} h^{1/4} F(h)$]; qu'il lui correspond la même valeur moyenne de l'énergie h ; et que l'entropie de μ_∞ est obtenue en ajoutant à celle de μ le nombre positif $\log(2\pi/e)$.



support de μ



support de $[-\frac{t}{h}]^+(\mu)$

Fig. 17.1.

(17.23) — Si l'on part de l'idée préconçue que les états d'équilibres « naturels » sont ceux qui s'établissent spontanément par vieillissement, les remarques ci-dessus permettent de passer à l'idée que les équilibres sont d'autant plus « naturels » qu'ils ont une plus grande entropie; puis à formuler le principe suivant :

(17.24) Les états d'équilibre naturels d'un système constituent l'ensemble de Gibbs du groupe dynamique des translations dans le temps

⁽¹⁾ Notation (16.32) : $p^+ = p$ si $p \geq 0$, $p^+ = 0$ si $p \leq 0$.

parmi les états qui attribuent à l'énergie une valeur moyenne donnée, on sait en effet que les états de Gibbs sont ceux qui ont la plus grande entropie (Théorème (16.219)).

On peut donner des motivations plus raffinées à ce principe ⁽¹⁾; quoiqu'il en soit, nous allons l'étudier en lui-même, et le comparer à l'expérience; nous en proposerons plus loin une modification (17.77). ■

(17.25) Un état d'équilibre naturel sera donc caractérisé par un élément Z de l'algèbre de Lie du groupe; Z est donc un nombre réel; nous constaterons plus loin qu'il repère la température d'équilibre. En désignant par E la valeur moyenne de l'énergie h , les résultats (16.219) nous permettent d'écrire les formules :

$\diamond z = \log \int_U e^{-Zh} \varphi(dx)$	$\heartsuit E = - \frac{dz}{dZ}$
$\clubsuit s = s + EZ$	$\spadesuit \frac{dE}{dZ} \leq 0$

(17.26)

en introduisant une variable réelle z qu'on appelle *potentiel thermodynamique de Planck*; la densité de l'état d'équilibre est

(17.27)
$$e^{-[z+Zh]} \varphi$$

φ désignant, comme dans (17.26), la densité symplectique de U .

Dans le cas général où les mouvements du système ne sont pas tous des équilibres, on peut appliquer (16.220); on voit alors que la fonction $Z \mapsto E$ est strictement décroissante, et que l'on a

(17.28)
$$\spadesuit \frac{\partial E}{\partial Z} < 0 \quad \heartsuit \frac{\partial s}{\partial E} = Z \quad \blacksquare$$

— En appliquant (16.221), on voit que :

Si on remplace le hamiltonien h par $h - h_0$ (h_0 étant une constante arbitraire), les variables (17.26) subissent la substitution

(17.29)

$Z \rightarrow Z$	$s \rightarrow s$	$z \rightarrow z - h_0 Z$	$E \rightarrow E - h_0$
-------------------	-------------------	---------------------------	-------------------------

l'ensemble des états d'équilibre naturels ne dépend pas du choix de h_0 .

⁽¹⁾ Voir notamment A. I. Kinchin, *Mathematical Foundations of Statistical Mechanics* (trad. anglaise par G. Gamow (1949)); Blanc-Lapierre, Casal, Tortrat (référence ci-dessus); E. T. Jaynes, *Phys. Rev.*, 106, 620 (1957); Mackey (référence ci-dessus).

— Comme le groupe des translations dans le temps est abélien, le théorème (16.222) montre que les éléments de ce groupe opèrent sur les états d'équilibre *sans les changer* — et sans changer la valeur des variables z, E, s ; le principe de Gibbs (17.24) définit bien des états d'équilibre, au sens (17.14).

— Voici les *équations aux dimensions* des grandeurs associées à l'équilibre (notations (13.13)) :

$$(17.30) \quad \begin{array}{|c|c|} \hline \sigma & A \\ \hline \varphi & A^N \\ \hline \end{array} \quad \begin{array}{|c|c|} \hline Z & TA^{-1} \\ \hline E & AT^{-1} \\ \hline \end{array} \quad \left. \begin{array}{l} e^z \\ e^s \end{array} \right\} A^{-N}$$

$2N$ étant la dimension de la variété symplectique U .

On constate que Z n'a pas la même équation aux dimensions T qu'une translation infinitésimale ε dans le temps (13.13), et que les variables z et s n'ont pas d'équation aux dimensions, tant qu'on n'a pas fait $A = 1$, c'est-à-dire tant qu'on n'a pas choisi une unité d'action (comme nous le ferons plus loin en mécanique quantique). Par contre E a toujours l'équation aux dimensions d'une énergie.

— Supposons qu'un système soit composé de plusieurs systèmes sans interactions; on sait que l'espace des mouvements U est le *produit direct symplectique* des espaces des mouvements U_j (12.146), ce qui entraîne la possibilité de choisir comme hamiltonien du système composé la somme des hamiltoniens partiels h_j ; on sait aussi que la mesure de Liouville de U est le produit tensoriel des mesures de Liouville des U_j (16.98); en appliquant (16.143) et (16.114), on voit que l'intégrale I_0 (notations (16.201)) pourra s'écrire sous la forme

$$(17.31) \quad I_0 = \prod_j \int_{U_j} e^{-Zh_j} \varphi_j(dx_j)$$

d'où résulte immédiatement l'énoncé :

(17.32) Tout état d'équilibre naturel d'un système composé de systèmes sans interactions s'obtient en faisant le *produit tensoriel* d'états d'équilibre naturel des systèmes composants correspondant tous à la même température ⁽¹⁾; les variables z, E, s du système composé s'obtiennent en ajoutant les variables z_j, E_j, s_j correspondantes des systèmes composants.

(17.33) Dans tout équilibre naturel d'un système composé, on peut donc dire que les mouvements des systèmes composants sont des *variables aléatoires*

⁽¹⁾ Voir (17.25).

indépendantes (Définition (16.188)); réciproquement, si on donne à chaque système un équilibre naturel, on définit (par produit tensoriel) un équilibre du système composé; cet équilibre n'est naturel que si les températures sont toutes égales.

Ce fait pourrait faire douter de la validité physique du principe de Gibbs (17.24); il faut cependant remarquer, expérimentalement, que l'existence d'une interaction, si faible soit-elle, empêche l'établissement d'équilibre non isotherme (exemple : la conductivité thermique d'une plaque solide séparant deux milieux); notons aussi que les traités de thermodynamique physique définissent l'égalité de deux températures par le fait que la juxtaposition des deux systèmes en équilibre est encore un système en équilibre; ce fait est d'ailleurs le principe de fonctionnement de la plupart des *thermomètres*.

GAZ PARFAITS

(17.34) Soit à étudier l'équilibre naturel d'un gaz enfermé dans une boîte. Cette étude suppose évidemment une connaissance détaillée des interactions entre molécules pour pouvoir calculer l'intégrale (notation (16.201))

$$(17.35) \quad I_0 = \int_U e^{-Zh} \varphi(dx)$$

Choisissons une date t ; une *approximation* consistera à remplacer cette intégrale par l'intégrale prise sur l'ensemble des mouvements qui ne comportent pas de collision à cet instant t (17.3) ⁽¹⁾, ensemble qui est égal à val (A_t) (notation (17.5)); en utilisant le fait que A_t est un symplectomorphisme (15.64), on remplacera donc I_0 par I'_0 :

$$(17.36) \quad I'_0 = \int_{E_t} e^{-Zh} \varphi_*(dx_*)$$

φ_* étant la densité symplectique de l'espace des mouvements *libres*.

(17.37) Une *seconde approximation* consistera à négliger le volume propre des molécules, c'est-à-dire, de façon précise, à remplacer E_t par l'ensemble E'_t des mouvements libres dans lesquels chaque molécule est à l'intérieur de la boîte à la date t ⁽²⁾; on est alors dans l'approximation dite du *gaz parfait*.

On voit alors que la densité de probabilité, l'entropie, le potentiel thermodynamique, la valeur moyenne de l'énergie sont *les mêmes que si les*

⁽¹⁾ Ce qui revient à supposer que l'ensemble des mouvements dans lesquels une interaction se produit à l'instant t peut être inclus dans un ouvert de probabilité très petite (si cette probabilité peut être rendue inférieure à tout nombre positif, on dit que l'ensemble est *négligeable*).

⁽²⁾ On traite donc comme très petite la probabilité pour que des molécules se touchent ou s'interpénètrent à l'instant t donné.

molécules n'avaient aucune interaction ⁽¹⁾ — leurs mouvements étant définis par leurs positions et leurs vitesses à l'instant t choisi. En se reportant à (17.32), on voit qu'on est ramené ⁽²⁾ à l'étude de l'équilibre naturel de chaque molécule, supposée seule dans la boîte à la température donnée : il se produit donc une *décomposition statistique* de l'état d'équilibre ; mais il n'y a pas *décomposition dynamique* (l'espace des mouvements du gaz n'est, en aucune façon, le produit direct des espaces de mouvements des molécules isolées).

Gaz parfait monoatomique

(17.38) On traite dans ce cas chaque molécule comme un point matériel en mécanique classique. Les résultats sont les suivants :

— A chaque instant t , la *position* et la *vitesse* de chaque molécule sont des variables aléatoires *indépendantes*.

— La loi de probabilité de la vitesse v possède la densité

$$(17.39) \quad \left[\frac{mZ}{2\pi} \right]^{3/2} e^{-mZ\|v\|^2/2} \varphi \quad (\text{loi de répartition des vitesses de Maxwell})$$

m étant la masse, φ la densité euclidienne de R^3 ⁽³⁾.

— La densité de probabilité de la position r est ⁽⁴⁾

$$(17.40) \quad \frac{1}{V} \varphi$$

φ étant toujours la densité euclidienne de R^3 , V le volume de l'enceinte.

Gaz parfait quelconque

(17.41) On suppose que les dimensions des molécules sont petites devant celles de la boîte (c'est-à-dire essentiellement qu'une molécule est « à l'intérieur » de la boîte si son centre de gravité l'est) ; en utilisant la mécanique *non relativiste*, on sait que l'espace des mouvements libres d'une molécule est le *produit symplectique* de l'espace des mouvements du centre de gravité

⁽¹⁾ Et par conséquent pouvaient se traverser l'une l'autre sans être déviées.

⁽²⁾ Soit en supposant que l'espace des mouvements libres des molécules est le *produit direct* des espaces de mouvements de chaque molécule (voir (17.32)), soit en supposant que les molécules sont *indiscernables* et en travaillant sur le revêtement défini en (15.46).

⁽³⁾ φ coïncide avec la densité de Lebesgue (16.85) ; il s'agit donc d'une *loi normale* (voir (16.206)). En supposant, selon la coutume, m positive, on voit que l'ensemble Ω (notation (16.219)) est R^+ : Z peut prendre toutes les valeurs positives.

⁽⁴⁾ Cette formule suppose négligeables les forces extérieures, telles que le champ électrique ou la pesanteur. Dans le cas contraire, on trouverait la loi *barométrique*.

et de l'espace des mouvements propres (13.24) ; ce qui conduit aux résultats suivants :

- (17.42) [a) La *position*, la *vitesse* et le *mouvement propre* de chaque molécule sont des variables aléatoires indépendantes ⁽¹⁾.
 b) La loi de probabilité des vitesses et celle des positions sont les mêmes que dans le cas monoatomique (17.39) et (17.40) ⁽²⁾.
 c) La loi de probabilité des mouvements propres est la loi de Gibbs correspondant à la valeur donnée de Z .

Thermomètre à gaz parfait

(17.43) Reproduisons schématiquement le principe d'un appareil utilisé pour la mesure précise des températures (*thermomètre à hydrogène*) : un gaz est enfermé dans un cylindre fermé par un piston de masse M , de poids Mg , susceptible — grâce à des liaisons parfaites — de se déplacer verticalement (Fig. 17. II) ; on néglige l'effet de la pesanteur sur les molécules ⁽³⁾.

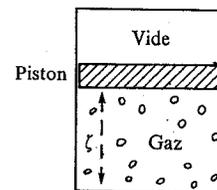


Fig. 17. II.

Z Dans ces conditions, les *positions des molécules* ne sont plus des variables aléatoires indépendantes ; la cote ζ du piston est elle-même une *variable aléatoire*, dont on peut calculer la densité de probabilité ⁽⁴⁾

$$(17.44) \quad \rho = [ZMg]^{n+1} e^{-ZMg\zeta} \frac{\zeta^n}{n!} ;$$

n désignant le *nombre de molécules* enfermées dans la boîte.

⁽¹⁾ Dans le cas de molécules indiscernables, ce résultat est valable sur le revêtement ; on passe aux lois de probabilités réelles en prenant l'image par la projection $x \mapsto \tilde{x}$ (notation (15.45)).

⁽²⁾ Il existe une vérification expérimentale directe de la loi de Maxwell (17.39) : la fréquence d'une raie spectrale émise par un atome dépend, en apparence, de sa vitesse (effet Doppler-Fizeau) ; les raies émises par un gaz chaud sont donc élargies, avec une répartition spectrale qui se déduit facilement de (17.39) ; cet effet est utilisé pour des mesures de température dans la haute atmosphère et dans les étoiles.

⁽³⁾ Simplement pour la facilité du calcul ; une étude plus complète montre que cet effet est bien négligeable dans les conditions usuelles.

⁽⁴⁾ La densité de Lebesgue de R est sous-entendue.

On en déduit immédiatement la valeur moyenne de ζ

$$(17.45) \quad \text{moyenne } \zeta = \frac{n+1}{ZMg}$$

d'où, en introduisant la *pression* exercée par le piston $p = Mg/A$ ($A =$ aire du piston) et la *variable aléatoire* V , *volume* de l'enceinte ($V \equiv A\zeta$) :

$$(17.46) \quad p \times \text{moyenne } V = \frac{n+1}{Z} \quad (1) \quad \blacksquare$$

Or l'expérience montre que les gaz réels vérifient très bien — lorsque leur densité est très petite (2) — la loi de Mariotte-Gay-Lussac-Avogadro :

$$(17.47) \quad pV = nkT$$

où T définit la *température absolue*, k une constante dite *constante de Boltzmann* :

$$(17.48) \quad k = 1,3804 \cdot 10^{-16} \text{ erg/degé Kélyn. } \blacksquare$$

Puisque le rapport $n/n+1$ est très voisin de 1 dans les expériences, la comparaison de (17.46) et (17.47) donne la relation

$$(17.49) \quad Z = \frac{1}{kT}$$

entre la *température absolue* usuelle mesurée par le thermomètre et le paramètre Z ; encore faut-il comprendre pourquoi la variable aléatoire V peut être confondue, dans une expérience, avec sa valeur moyenne.

Il est facile de déduire de (17.44) la valeur moyenne de ζ^2

$$(17.50) \quad \text{moyenne } \zeta^2 = \frac{[n+1][n+2]}{[ZMg]^2}$$

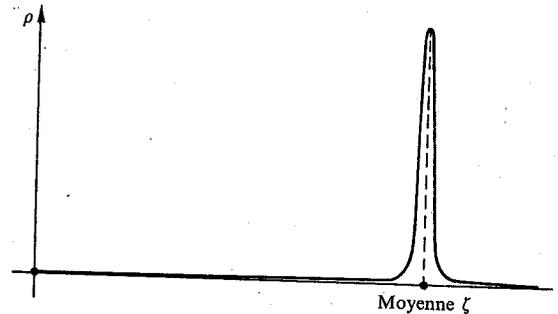
d'où découle la *variance* de ζ (Définition (16.196)); la racine carrée de cette *variance*, appelée *erreur quadratique moyenne sur* ζ , vaut

$$(17.51) \quad \Delta\zeta = \frac{\text{moyenne } \zeta}{\sqrt{n+1}}$$

(1) On rappelle que Z est nécessairement positif (17.39); donc aussi p .

(2) Les approximations de la théorie du gaz parfait (17.36) supposent évidemment des molécules écartées les unes des autres — donc un gaz peu condensé.

Cette formule montre que $\Delta\zeta/\text{moyenne } \zeta$ est *très petit* dans les circonstances usuelles (de l'ordre de 10^{-11} pour un litre d'hydrogène à la température ordinaire); on constate d'ailleurs que le graphe de la loi de probabilité (17.44) possède un *pic très aigu*; la valeur (17.51) de $\Delta\zeta$ donne une estimation de la *largeur* de pic (la Fig. (17.III) n'en donne qu'une allure qualitative).



(17.52) On doit donc admettre, comme règle d'interprétation de la mécanique statistique, qu'une variable aléatoire dont la loi de probabilité est très concentrée se confond — expérimentalement — avec sa valeur moyenne; mathématiquement, on peut noter que la loi de probabilité de la variable ζ tend vers la mesure de Dirac δ_1 , au sens de la topologie vague, lorsque n tend vers l'infini. \blacksquare

CHALEUR ET TRAVAIL

Nous savons que l'entropie s d'un système peut s'exprimer en fonction de l'énergie E , et que l'on a (17.28c)

$$(17.53) \quad \frac{ds}{dE} = Z.$$

(17.54) Il se trouve que Z est toujours positif dans les conditions que nous connaissons (voir (17.39) et (17.49)); on peut donc exprimer inversement E en fonction de s et, éventuellement, d'une variable u exprimant les caractéristiques du système — u étant prise, pour fixer les idées, dans une variété;

si $\left(\frac{a}{s} \right) \rightarrow E$ est différentiable, nous pourrions donc écrire (à cause de (17.53))

$$(17.55) \quad dE \equiv \frac{ds}{Z} + \omega(du)$$

ω étant une 1-forme (appelée, elle aussi, *potentiel thermodynamique* !); en utilisant (17.49) on écrira

$$(17.56) \quad \boxed{dE = T dS + \omega(du)} \quad \text{avec} \quad \boxed{S \equiv ks}$$

on dit que T et ω sont les variables de « tension » associées aux variables de « position » S et u ; c'est généralement S que l'on appelle « entropie » en thermodynamique; elle coïncide avec s si on a soin de choisir une unité de température telle que la constante de Boltzmann soit égale à 1.

Si nous faisons évoluer l'équilibre en fonction d'un paramètre, et si nous désignons par d la dérivation associée, la formule (17.56) s'écrit généralement

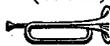
$$(17.57) \quad \boxed{dE = dQ + dW}$$

avec

$$(17.58) \quad \boxed{dQ = T dS} \quad \boxed{dW = \omega(du)}$$

la variable Q s'appelle la *chaleur* et la variable W le *travail*; notons que Q et W ne sont pas des fonctions de S et u , mais des *intégrales d'action* (au sens du calcul des variations), dépendant du programme d'évolution choisi.

(17.59) Divers types d'évolution méritent d'être examinées séparément : évolution *adiabatique* ($S = \text{Cte}$, pas d'échange de chaleur, on peut prendre $W = E$); évolution *isotherme* ($T = \text{Cte}$, on peut prendre $Q = TS$, $W = E - TS$); *chauffage* ou *réfrigération* ($u = \text{Cte}$; on peut prendre $Q = E$; c'est le cas où sont valables les formules (17.26)).

(17.60)  — Les formules (17.57), (17.58) constituent une formulation des deux principes de la thermodynamique (équivalence chaleur-travail; dQ/T est une « différentielle totale »). Donnons un premier exemple de calcul thermodynamique.

Sur la variété parcourue par le couple $\begin{pmatrix} S \\ u \end{pmatrix}$, le théorème de Poincaré

$$\nabla \nabla E = 0 \text{ s'écrit :}$$

$$\diamond \quad \frac{dT}{T} \delta Q - \frac{\delta T}{T} dQ + d[\omega(\delta u)] - \delta[\omega(du)] - \omega([d, \delta] u) \equiv 0$$

(relation de Maxwell généralisée); en choisissant $\delta T = \text{Cte}$, $\delta u = 0$, $dT = 0$, du indépendant de T , on trouve la relation de *Clapeyron-Clausius*

$$\heartsuit \quad dQ = -T \frac{\partial \omega}{\partial T}(du)$$

valable pour toute transformation infinitésimale isotherme repérée par du ; la vérification est bonne, par exemple, dans le cas de la liquéfaction d'un gaz.

CHALEURS SPÉCIFIQUES

Considérons un gaz enfermé dans une enceinte de volume ajustable V ; on peut prendre V comme paramètre u ; la formule (17.56) s'écrit dans ce cas

$$(61) \quad dE = T dS - p dV$$

p s'appelant la *pression* ⁽¹⁾.

(62) On appelle *capacité calorifique* du gaz la quantité dQ/dT ; elle dépend évidemment du programme d'évolution choisi.

— En particulier, la capacité calorifique à volume constant est

$$(63) \quad \boxed{c \equiv T \frac{dS}{dT} \equiv -kZ^2 \frac{dE}{dZ} \equiv -kZ \frac{ds}{dZ}}$$

l'inégalité (17.28) montre que c est toujours positive.

(64) — Dans l'évolution à pression constante, on peut évidemment prendre $W \equiv -pV$, d'où $Q \equiv E + pV$ (cette variable $E + pV$ s'appelle l'*enthalpie*); nous noterons C la capacité calorifique correspondante.

Exemple

Supposons ⁽²⁾ que le potentiel thermodynamique de Planck d'un fluide se mette sous la forme

$$z \equiv f(Z) + g(V).$$

⁽¹⁾ Elle coïncide avec le résultat de la mesure décrite en (17.43), à condition de supposer négligeable, a priori, la variance du volume (nous avons calculé cette variance dans le cas d'un gaz parfait).

⁽²⁾ C'est ainsi que l'on définit les *fluides parfaits* dans les traités de thermodynamique; ou d'une façon équivalente, par le fait que l'énergie est indépendante du volume (formule (17.66), expérience de Joule).

En utilisant les formules (17.26), (17.61), (17.58) on obtient facilement l'énergie

$$(17.66) \quad E = -f'(Z)$$

l'entropie

$$(17.67) \quad s = f(Z) - Zf'(Z) + g(V)$$

la pression

$$(17.68) \quad p = \frac{g'(V)}{Z}$$

les capacités calorifiques

$$(17.69) \quad c = kZ^2 f''(Z), \quad C = k \left[Z^2 f''(Z) - \frac{g'(V)^2}{g''(V)} \right]$$

— Le calcul montre que le potentiel thermodynamique d'un gaz parfait est de la forme (17.65), avec

$$(17.70) \quad \begin{cases} f(Z) = -\frac{3n}{2} \log(Z) + \sum_j \left[\frac{3}{2} \log(2\pi m_j) + z_j \right] \\ g(V) = n \log(V) \end{cases}$$

les m_j étant les masses des n molécules, les z_j les potentiels thermodynamiques de leurs mouvements propres ⁽¹⁾.

On voit que la capacité calorifique (à volume constant) c dépend du modèle choisi pour les molécules; on démontre facilement les relations suivantes — qui sont indépendantes du modèle :

$$(17.71, 72) \quad \boxed{c_m \geq \frac{3}{2}k} \quad \boxed{C_m - c_m = k} \quad (2)$$

où c_m et C_m désignent les chaleurs spécifiques moléculaires (quotients de c et C par le nombre n des molécules); d'où résulte, pour le rapport γ des chaleurs spécifiques :

$$(17.73) \quad \boxed{\gamma = \frac{C}{c} = 1 + \frac{k}{c_m} \leq \frac{5}{3}}$$

⁽¹⁾ Voir (17.42). Si les molécules sont indiscernables, il faut retrancher à $f(Z)$ la constante $\log(n!)$, ce qui ne change pas p , c , C .

⁽²⁾ Relation de Robert von Mayer.

ces relations (17.71, 72, 73) sont bien vérifiées par les fluides réels suffisamment peu condensés.

Le tableau suivant donne la valeur de z , c_m , γ dans divers modèles ⁽¹⁾ :

Modèle	z	c_m	γ
Point matériel	0	$3k/2$	$5/3$
Particule à spin	0	$3k/2$	$5/3$
Solide rectiligne	$\log T + \text{Cte}$	$5k/2$	$7/5$
Solide non rectiligne	$\frac{3}{2} \log T + \text{Cte}$	$3k$	$4/3$
Deux points liés par une force en r^p	$\left[\frac{3}{2} + \frac{3}{p+1} \right] \log T + \text{Cte}$	$\left[3 + \frac{3}{p+1} \right] k$	$< 4/3$

L'expérience montre, pour les gaz monoatomiques, que γ est voisin de $5/3$; le modèle constitué par un point matériel ou une particule à spin est donc acceptable. Pour un gaz bi-atomique, γ est une fonction décroissante de T , qui reste généralement comprise entre 1,6 et 1,3; on voit que les modèles ci-dessus ne conviennent pas, puisqu'ils donnent une valeur constante à c_m et γ ; la difficulté de trouver un modèle moléculaire convenable a été l'une des motivations importantes de la mécanique quantique.

— La chaleur spécifique d'un solide se calcule immédiatement en supposant que les atomes qui le constituent peuvent osciller autour d'une position d'équilibre; dans l'approximation linéarisée (voir (12.150)), l'espace des mouvements est un espace vectoriel de dimension $6n$ (n étant le nombre d'atomes considérés comme des points matériels), et le hamiltonien une forme quadratique positive; alors la loi de probabilité de l'équilibre naturel est une loi normale; la formule (16.207) donne immédiatement $z \equiv -3n \log(Z) + \text{Cte}$, d'où la chaleur spécifique atomique $c_{at} = 3k$; l'accord avec l'expérience est bon si la température n'est pas trop basse (loi de Dulong et Petit).

MÉCANIQUE STATISTIQUE COVARIANTE

Le groupe des translations dans le temps (7.9) est un sous-groupe du groupe de Galilée; mais ce n'est pas un sous-groupe invariant, ainsi que le

⁽¹⁾ On notera que l'existence du spin ne change pas les résultats, comme pourrait le faire croire le prétendu « principe de l'équipartition de l'énergie ».

montre un calcul trivial. Si un système dynamique est *conservatif* dans un repère d'inertie, il en résulte qu'il peut *ne plus être conservatif dans un autre*. La formulation (17.24) du principe de Gibbs doit donc être élargie, pour devenir compatible avec la relativité galiléenne.

Nous proposons donc le principe suivant :

(17.77) Si un système dynamique est invariant par un sous-groupe de Lie G' du groupe de Galilée, les équilibres naturels du système constituent l'ensemble de Gibbs du groupe dynamique G' .

Soit \mathcal{G}' l'algèbre de Lie G' ; on sait que \mathcal{G}' est une sous-algèbre de Lie de celle de G , notée \mathcal{G} ; un équilibre du système sera caractérisé par un élément Z de \mathcal{G}' , donc de \mathcal{G} ; on pourra écrire

$$(17.78) \quad Z = \begin{bmatrix} j(\omega) & \beta & \gamma \\ 0 & 0 & \varepsilon \\ 0 & 0 & 0 \end{bmatrix}$$

en utilisant les notations (13.4); Z parcourt l'ensemble Ω défini en (16.219); à chaque valeur de Z est associé un élément M du dual \mathcal{G}'^* de \mathcal{G}' , valeur moyenne du moment μ ; on peut appliquer les formules (16.219), (16.220), qui généralisent les relations thermodynamiques (17.26), (17.27), (17.28).

(17.79) On voit que c'est Z (17.78) qui généralise la « température »; le théorème d'isothermie (17.32) s'étend immédiatement : l'équilibre d'un système composé de plusieurs parties sans interactions s'obtient en attribuant à chaque composante un équilibre *correspondant à la même valeur de Z* ; l'entropie s , le potentiel de Planck z et le moment moyen M sont *additifs*. W

(17.80) Nous allons donc — comme dans le cas classique — interpréter Z au moyen d'un « thermomètre » constitué par une *bulle connexe de gaz parfait*; on applique encore la théorie de la diffusion aux mouvements libres des molécules, et on suppose que le diffuseur admet G' comme *groupe de symétrie* (Définition (15.69)); en appliquant la théorie des groupes de symétrie (15.73), l'approximation des gaz parfaits (17.36), la décomposition barycentrique (13.14 à 13.36), et en portant dans la définition (16.219) de l'ensemble de Gibbs, on constate que le mouvement x du *centre de gravité* de chaque molécule (considéré comme point matériel de masse m) et le *mouvement propre* x' de chaque molécule sont des variables aléatoires *indépendantes*, ayant respectivement comme densités de probabilité

$$(17.81) \quad C e^{-\psi(x).Z} \varphi,$$

$$(17.82) \quad C' e^{-\psi'(x').Z} \varphi',$$

$\psi(x)$ et $\psi'(x')$ étant les moments (dans le dual \mathcal{G}^* de l'algèbre de Lie du groupe de Galilée), φ et φ' les densités symplectiques correspondantes, C et C' des constantes dépendant de la bulle. ■

Pour interpréter la formule (17.81), nous allons effectuer quelques changements de variable.

En utilisant le théorème de Noether, on peut écrire $\psi(x) \equiv \psi(y)$ (Fig. 12.III)

$y \equiv \begin{pmatrix} t \\ \mathbf{r} \\ \mathbf{v} \end{pmatrix}$ étant un élément de l'espace d'évolution V du *centre de gravité*, ψ étant

l'application

$$(17.83) \quad \psi(y) \equiv m \{ \mathbf{r} \times \mathbf{v}, \mathbf{r} - \mathbf{v}t, \mathbf{v}, \frac{1}{2} \|\mathbf{v}\|^2 \} \quad (\text{formule (13.36)}).$$

Puisque le groupe de Galilée opère sur V , on peut *poser*

$$(17.84) \quad y_0 \equiv \begin{pmatrix} t_0 \\ \mathbf{r}_0 \\ \mathbf{v}_0 \end{pmatrix} \equiv \exp \left(-\frac{t}{\varepsilon} Z \right) (y);$$

on a d'ailleurs

$$(17.85) \quad \begin{pmatrix} \mathbf{r}_0 \\ t_0 \\ 1 \end{pmatrix} = \exp \left\{ - \begin{bmatrix} j(\omega t/\varepsilon) & \beta t/\varepsilon & \gamma t/\varepsilon \\ 0 & 0 & t \\ 0 & 0 & 0 \end{bmatrix} \right\} \cdot \begin{pmatrix} \mathbf{r} \\ t \\ 0 \end{pmatrix}$$

(Théorème (6.12)), ce qui permet notamment de vérifier que

$$(17.86) \quad t_0 \equiv 0.$$

la formule (12.116) nous donne d'autre part

$$(17.87) \quad \begin{pmatrix} \mathbf{v}_0 \\ 1 \\ 0 \end{pmatrix} = \exp \left\{ -\frac{t}{\varepsilon} Z \right\} \begin{pmatrix} \mathbf{v} \\ 1 \\ 0 \end{pmatrix}.$$

En posant $a \equiv \exp \left(\frac{t}{\varepsilon} Z \right)$, on voit que

$$\begin{aligned} \psi(y).Z &\equiv \psi(a_V(y_0)).Z \equiv [a_{\mathcal{G}'} \psi(y_0) + \theta(a)].Z \quad (11.17 \diamond) \\ &\equiv \psi(y_0) a^{-1}_{\mathcal{G}'}(Z) + \theta(a).Z \quad (11.15); \end{aligned}$$

en appliquant la formule intégrale (11.22 ♠), comme le cocycle θ est symplectique, on constate que $\theta(a).Z \equiv 0$; on voit facilement que $a^{-1}_{\mathcal{G}'}(Z) \equiv Z$; on a donc

$$(17.88) \quad \psi(y).Z \equiv \psi(y_0).Z$$

en utilisant (17.83), (17.86) et (12.122), on trouve alors

$$(17.89) \quad \psi(y) \cdot Z \equiv -m\varepsilon \left[\frac{1}{2} \| \mathbf{v}_0 - \mathbf{U}^* \|^2 + f(\mathbf{r}_0) \right],$$

avec

$$(17.90) \quad \mathbf{U}^* \equiv \frac{\boldsymbol{\omega}}{\varepsilon} \times \mathbf{r}_0 + \frac{\boldsymbol{\gamma}}{\varepsilon},$$

$$(17.91) \quad f(\mathbf{r}_0) \equiv \left\langle \mathbf{r}_0, \frac{\boldsymbol{\beta}}{\varepsilon} \right\rangle - \frac{1}{2} \| \mathbf{U}^* \|^2.$$

La formule (17.85) montre que

$$(17.92) \quad \mathbf{r}_0 \equiv A(\mathbf{r} + \mathbf{R}) \quad \left[A = \exp \left(j \left(\frac{-\boldsymbol{\omega} t}{\varepsilon} \right) \right) \in \text{SO}(3); \quad \mathbf{R} \in R^3 \right],$$

A et R étant fonctions de t , on peut interpréter (17.92) en disant que \mathbf{r}_0 est la position de la molécule dans un *nouveau référentiel*, en général *accélééré* (il suffit de reconnaître la formule (12.59)). Nous poserons maintenant

$$(17.93) \quad \boxed{\mathbf{r}^* \equiv \mathbf{r}_0 \quad \mathbf{v}^* \equiv \text{vitesse relative au nouveau référentiel } ^{(1)}}$$

$$(17.94) \quad \boxed{\boldsymbol{\omega}^* = \frac{\boldsymbol{\omega}}{\varepsilon} \quad \boldsymbol{\beta}^* = \frac{\boldsymbol{\beta}}{\varepsilon} \quad \boldsymbol{\gamma}^* = \frac{\boldsymbol{\gamma}}{\varepsilon}}.$$

On a alors

$$(17.95) \quad \boxed{\mathbf{U}^* \equiv \boldsymbol{\omega}^* \times \mathbf{r}^* + \boldsymbol{\gamma}^* \quad f(\mathbf{r}^*) \equiv \langle \boldsymbol{\beta}^*, \mathbf{r}^* \rangle - \frac{\| \mathbf{U}^* \|^2}{2}}$$

la vitesse \mathbf{U} et l'accélération $\boldsymbol{\Gamma}$ d'entraînement (12.66) sont données par

$$(17.96) \quad \boxed{\begin{pmatrix} \mathbf{U} \\ 1 \\ 0 \end{pmatrix} \equiv \exp \left(\frac{t}{\varepsilon} Z \right) \cdot \begin{pmatrix} \mathbf{U}^* \\ 1 \\ 0 \end{pmatrix} \quad \begin{pmatrix} \boldsymbol{\Gamma} \\ 0 \\ 0 \end{pmatrix} \equiv \exp \left(\frac{t}{\varepsilon} Z \right) \cdot \begin{pmatrix} \boldsymbol{\omega}^* \times \mathbf{U}^* + \boldsymbol{\beta}^* \\ 0 \\ 0 \end{pmatrix}}$$

La *vitesse relative* (formule (12.67)) :

$$(17.97) \quad \mathbf{v}^* = \exp(j(-\boldsymbol{\omega}' t)) [\mathbf{v} - \mathbf{U}^*],$$

peut s'écrire, grâce à (17.87) et (17.96)

$$(17.98) \quad \boxed{\mathbf{v}^* = \mathbf{v}_0 - \mathbf{U}^*};$$

⁽¹⁾ Ne pas confondre \mathbf{v}^* et \mathbf{v}_0 .

d'autre part les formules (12.44) et (12.67) montrent qu'il apparaît dans le nouveau référentiel une force, dite *force d'inertie* ⁽¹⁾ :

$$(17.99) \quad \mathbf{F}^* \equiv \mathbf{E}^* - \mathbf{B}^* \times \mathbf{v}^*$$

avec

$$(17.100) \quad \mathbf{E}^* \equiv -m[\boldsymbol{\omega}^* \times \mathbf{U}^* + \boldsymbol{\beta}^*]$$

$$(17.101) \quad \mathbf{B}^* \equiv 2m\boldsymbol{\omega}^*;$$

on vérifie immédiatement qu'elles dérivent d'un *potentiel* v^* et d'un *potentiel-vecteur* \mathbf{A}^* (Définitions (12.91)) donnés par

$$(17.102) \quad v^* \equiv mf(\mathbf{r}^*) \quad \mathbf{A}^* = m\boldsymbol{\omega}^* \times \mathbf{r}^*$$

ce qui permet de construire le *hamiltonien* (formule (12.93))

$$(17.103) \quad \boxed{h^* = \frac{1}{2} m \| \mathbf{v}^* \|^2 + mf(\mathbf{r}^*)}$$

En comparant cette formule (17.103) à (17.89), tenant compte de (17.98), on peut donc mettre la densité de probabilité (17.81) du centre de gravité d'une molécule sous la forme

$$(17.104) \quad C e^{eh^*} \varphi^*$$

φ^* désignant la densité symplectique dans le nouveau référentiel (puisque le changement de référentiel conserve la forme symplectique σ (12.68), il conserve aussi la densité symplectique) ; ce qui fournit l'énoncé suivant :

(Notations (17.78), (17.94)).

(17.105) L'équilibre naturel d'un gaz parfait [au sens général (17.77)] correspondant à une valeur donnée de Z apparaît comme un équilibre naturel [au sens classique (17.24)], avec la température absolue

$$\diamond \quad T = -\frac{1}{k\varepsilon} \quad (k = \text{Cte de Boltzmann})$$

dans le référentiel où la position \mathbf{r}^* est donnée par la formule

$$\heartsuit \quad \begin{pmatrix} \mathbf{r} \\ t \\ 1 \end{pmatrix} \equiv \exp \left\{ t \begin{bmatrix} j(\boldsymbol{\omega}^*) & \boldsymbol{\beta}^* & \boldsymbol{\gamma}^* \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \right\} \cdot \begin{pmatrix} \mathbf{r}^* \\ 0 \\ 1 \end{pmatrix} \quad (2).$$

⁽¹⁾ Les deux termes de l'expression (17.99) de \mathbf{F}^* s'appellent respectivement *force centrifuge* et *force de Coriolis*.

⁽²⁾ Cette formule constitue l'équation des orbites spatio-temporelles du groupe de dimension 1 dont l'algèbre de Lie est engendrée par Z (ces orbites étant indexées par \mathbf{r}^*).

Remarques :

7.106) — Nous avons supposé $\varepsilon \neq 0$ en faisant le changement de variables (17.84); l'interprétation (17.105◇) montre que ε est négatif dans les conditions usuelles.

7.107) — Si un système est décomposé en systèmes sans interactions, l'équilibre est obtenu en attribuant la même valeur de Z aux systèmes composants (17.79); ainsi les systèmes composants auront, non seulement la même température absolue T (17.105◇), mais aussi le même référentiel d'équilibre apparent [(17.105♡) et (17.94)].

7.108) — Nous avons supposé, pour simplifier, qu'il n'y avait pas de forces appliquées directement aux molécules (exemple : l'attraction terrestre pour l'atmosphère); il suffit, dans le cas général, d'ajouter le potentiel correspondant au hamiltonien h^* .

Exemples

7.109) *Vent* ($\omega = 0, \beta = 0$).
On trouve alors

$$\mathbf{r} \equiv \mathbf{r}^* + \gamma^* t; \quad f(\mathbf{r}^*) = \text{Cte}$$

le nouveau référentiel est un référentiel d'inertie; dans ce référentiel la probabilité de présence est constante, la répartition des vitesses est celle de Maxwell (17.39); le vecteur γ^* s'interprète donc, dans l'ancien référentiel, comme *vitesse du vent*. ■

7.110) *Fusée accélérée* ($\omega = 0, \gamma = 0$).

La matrice figurant dans (17.105♡) étant alors nilpotente (son cube est nul), l'exponentielle se calcule immédiatement par développement en série (6.17); on trouve

$$\diamond \quad \mathbf{r} = \mathbf{r}^* + \beta^* \frac{t^2}{2}$$

le nouveau référentiel a donc un *mouvement de translation uniformément accéléré*, d'accélération β^* (on peut penser par exemple à la cabine d'une fusée dont le fonctionnement est réglé ad hoc); la formule (17.100) montre qu'il y règne une *pesanteur artificielle* — β^* ⁽¹⁾; le potentiel intérieur est $v^* = m \langle \beta^*, \mathbf{r}^* \rangle$; la densité de probabilité de présence est donc propor-

⁽¹⁾ Cet exemple a été choisi par Einstein pour illustrer le principe d'équivalence des forces de pesanteur et d'inertie.

tionnelle à $\exp\left(-\frac{m}{kT} \langle \beta^*, \mathbf{r}^* \rangle\right)$; on trouve la loi barométrique même obtenue en supposant l'accélération de la pesanteur terrestre constante et égale à $-\beta^*$. ■

Centrifugeuse ($\beta = 0, \gamma = 0$).

Alors

$$(17.111) \quad \mathbf{r} \equiv \exp(j(\omega^* t)) \mathbf{r}^*$$

le nouveau référentiel a donc un mouvement de *rotation uniforme*, ω^* étant le *vecteur rotation* ⁽¹⁾.

La probabilité de présence du gaz est proportionnelle à

$$(17.112) \quad \exp\left(\frac{m}{2kT} \|\omega^* \times \mathbf{r}^*\|^2\right);$$

la présence de m dans cette expression montre — dans le cas d'un gaz non homogène — que *la concentration relative des divers constituants varie avec la distance à l'axe de rotation*, suivant une formule immédiate à déduire de (17.112); cet effet est bien vérifié expérimentalement; on l'utilise pour l'enrichissement de l'uranium. ■

Appliquons à ce cas la formule (17.82), en supposant le gaz composé de *particules à spin* (exemple : un plasma d'hydrogène ionisé); en appliquant (13.35) et (14.50), on constate que l'« élément de probabilité » du vecteur unitaire \mathbf{u} portant le spin est

$$(17.113) \quad C \exp\left(\frac{s}{kT} \langle \mathbf{u}, \omega^* \rangle\right) \varphi(d\mathbf{u})$$

C étant une constante, s le spin, $\varphi(d\mathbf{u})$ l'élément d'angle solide (16.174). On voit que l'orientation la plus probable du spin est celle de ω^* ; le rapport de sa probabilité à celle de l'orientation opposée est $\exp\left(\frac{2s \|\omega^*\|}{kT}\right)$; la vérification expérimentale de cet effet semble difficile, parce que ce nombre est très voisin de 1 dans les circonstances habituelles. ■

⁽¹⁾ (17.107) nous montre que la centrifugeuse doit tourner à la même vitesse angulaire que le gaz et avoir la même température.

ÉQUILIBRE STATISTIQUE D'UN SYSTÈME ISOLÉ

7.114) Il semble naturel, a priori, d'appliquer le principe (17.77) à un système isolé; mais cette tentative se heurte à une difficulté de principe : de tels équilibres *n'existent pas*.

On peut parvenir à cette conclusion par une méthode purement algébrique; *valable, donc, quel que soit le modèle du système*; la formule (b) du théorème (16.224) s'écrit

$$7.115) \quad m f_0(Z) + M \cdot \text{Ad}(Z) = 0$$

m étant la *masse* du système, f_0 le cocycle symplectique défini en (12.130), Z l'élément de l'algèbre de Lie définissant l'équilibre, M la valeur moyenne du moment galiléen. Or on peut vérifier que cette équation impose à Z des relations qui ne sont vérifiées dans *aucun ouvert*; ce qui est incompatible avec les hypothèses (16.219) définissant l'ensemble de Gibbs.

7.116) Une seconde méthode pour parvenir à cette conclusion fait appel à la *décomposition barycentrique* (13.24) : si le système est en équilibre naturel, son centre de gravité (assimilé à un point matériel doué de toute la masse) l'est aussi; or on vérifie directement qu'un point matériel libre ne possède pas d'équilibre selon le groupe de Galilée ⁽¹⁾. ■

7.117) Il est clair que la difficulté qui se présente est d'ordre *cosmologique* : l'univers galiléen est impropre à posséder un équilibre statistique; bien que la cosmologie galiléenne soit fort naïve, l'observation confirme d'ailleurs ce fait : *l'univers est en expansion*.

Il est donc naturel d'éluder ce problème cosmologique, en recherchant si le *mouvement propre* ⁽²⁾, seul, peut posséder un équilibre conforme au principe (17.77). On sait que le groupe de Galilée opère sur le mouvement propre par l'intermédiaire du groupe G' constitué par les matrices

$$7.118) \quad \begin{bmatrix} A & 0 & 0 \\ 0 & 1 & e \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{array}{l} A \in \text{SO}(3) \\ e \in R \end{array}$$

⁽¹⁾ On sait d'ailleurs que de tels équilibres définiraient des équilibres pour les fluides parfaits; or l'expérience, comme le calcul, montre bien que les fluides parfaits ne possèdent d'équilibre (avec ou sans accélération du type (17.105)) que *s'ils sont enfermés dans une boîte*.

⁽²⁾ Définition (13.6).

(voir (13.30)); ce qui conduit à postuler que chaque équilibre naturel du mouvement propre est défini par un élément

$$(17.119) \quad Z = \begin{bmatrix} j(\omega) & 0 & 0 \\ 0 & 0 & \varepsilon \\ 0 & 0 & 0 \end{bmatrix}$$

de l'algèbre de Lie de G' .

(17.120) On se trouve donc dans le cas étudié en (17.111) : il existe un référentiel [doué d'un mouvement de rotation uniforme autour d'un axe passant par le centre de gravité 0, avec le vecteur rotation $\omega^* = \omega/\varepsilon$] par rapport auquel on a les apparences d'un équilibre usuel, à la température $T = -1/k\varepsilon$.

La formule (16.224b) s'écrit

$$\sigma \times \omega^* = 0$$

σ désignant la valeur moyenne du moment cinétique propre (13.35); il existe donc un nombre I tel que

$$(17.121) \quad \sigma = I\omega^*$$

on exprime cette propriété en langage classique en disant que l'axe de rotation est un *axe principal d'inertie* du système, et que I est le *moment d'inertie* autour de cet axe.

Par ailleurs, l'inégalité (16.224c) devient simplement

$$(17.122) \quad I > 0.$$

L'observation confirme bien ces dernières conclusions (rotation des corps célestes), alors que différents phénomènes (radioactivité, rayonnement) s'opposent à l'équilibre thermique.

MÉCANIQUE STATISTIQUE RELATIVISTE

On obtient évidemment un énoncé relativiste en remplaçant le groupe de Galilée par le groupe de Poincaré dans (17.77) :

(17.123) Si un système dynamique est invariant par un sous-groupe de Lie G' du groupe de Poincaré G , les équilibres naturels du système constituent l'ensemble de Gibbs du groupe dynamique G' .

Ce principe semble concorder de façon satisfaisante avec l'expérience; nous allons en étudier quelques conséquences.

Rappelons les notations : l'élément de l'algèbre de Lie \mathcal{G} de G qui fixe l'équilibre sera noté (comme en (13.54))

$$7.124) \quad Z = \begin{bmatrix} A & \Gamma \\ 0 & 0 \end{bmatrix}$$

Γ étant un vecteur de l'espace de Minkowski E_4 , A un opérateur anti-hermitien de cet espace ($\bar{A} = -A$).

7.125) — Il n'existe pas de décomposition barycentrique en mécanique relativiste ; mais on peut cependant montrer — en utilisant le théorème (16.224) et le fait que la cohomologie symplectique du groupe de Poincaré est nulle (13.73) qu'un système isolé n'a pas d'équilibre naturel — sauf dans le cas dépourvu d'interprétation physique où le groupe de Poincaré laisse invariants tous les mouvements du système ⁽¹⁾.

Il faut donc se limiter à un sous-groupe G' strictement plus petit que G . Nous allons examiner quelques exemples.

GAZ PARFAIT RELATIVISTE

Considérons un gaz parfait composé de points matériels (particules sans spin).

7.126) Supposons d'abord $A = 0$ (notation (17.124)); on peut montrer que l'intégrale I_0 (notation (16.201)) n'est convergente que si Γ est un vecteur du genre passé (la masse des particules étant supposée positive); on peut alors choisir un référentiel d'inertie \mathcal{R} par rapport auquel le vecteur Γ prend la forme $\mathcal{R} \begin{pmatrix} 0 \\ \varepsilon \end{pmatrix}$, $\varepsilon < 0$ (notation (13.55)); on constate que Z est

une translation infinitésimale dans le temps, ce qui permet d'interpréter, comme en (17.105), ε par la formule

$$7.127) \quad \varepsilon = -\frac{1}{kT},$$

T étant la température absolue, k la constante de Boltzmann.

Repérons une particule au moyen de son impulsion \mathbf{p} (13.79) et de sa position \mathbf{r} (pour $t = 0$, dans le référentiel \mathcal{R}); grâce à (13.60) et à la description (14.24) des particules sans spin, on trouve la probabilité élémentaire (pour une particule)

$$7.128) \quad \exp\left(-z - \frac{\sqrt{\|\mathbf{p}\|^2 + m^2}}{kT}\right) \varphi(d\mathbf{p}) \varphi(d\mathbf{r})$$

⁽¹⁾ Bien entendu, on peut nommer « vide » un tel système, et considérer qu'il constitue un équilibre naturel pour tout Z . Cette remarque cesse d'être purement verbale en mécanique quantique.

φ désignant la densité euclidienne et z le potentiel thermodynamique de Planck ; la formule (16.203) donne

$$(17.129) \quad z = \log \left(4 \pi V m^3 F \left(\frac{m}{kT} \right) \right)$$

V étant le volume de la boîte contenant le gaz (mesuré dans le référentiel \mathcal{R}) et F la fonction définie par

$$(17.130) \quad F(x) = \int_0^\infty e^{-x \operatorname{ch} \varphi} [\operatorname{sh} \varphi]^2 \operatorname{ch} \varphi \, d\varphi ;$$

on a d'ailleurs

$$(17.131) \quad F(x) = \frac{K_2(x)}{x}$$

K_2 désignant une fonction de Bessel ⁽¹⁾.

On déduit de (17.128) et (17.129) les lois de répartition des vitesses et de compressibilité qui remplacent les lois non relativistes de Maxwell et de Mariotte-Gay-Lussac ; la vitesse moyenne ⁽²⁾ est nulle, le gaz semble au repos dans le référentiel \mathcal{R} .

(17.132) On peut exprimer la fonction F par des formules d'approximation valables à basse température ⁽³⁾; on en déduit que la différence entre les lois relativistes et non relativistes reste négligeable tant que la quantité m/kT est grande devant l'unité; ce qui est le cas dans tous les exemples connus (pour l'hydrogène, m/k est voisine de 10^{13} degrés Kelvin ; la température au centre des étoiles les plus chaudes semble inférieure à 10^{10} degrés). ■

— En changeant de référentiel d'inertie, on observera un « vent relativiste » ; la température T relative au référentiel propre est évidemment un invariant relativiste, défini d'ailleurs par la formule

$$(17.133) \quad T = \frac{1}{k \sqrt{\langle \Gamma, \Gamma \rangle}}$$

— Comment peut-on définir la « température T' relative au nouveau référentiel » ? Si on tient à choisir une définition qui ne fasse intervenir

⁽¹⁾ Voir G. N. Watson, *A treatise on the Theory of Bessel Functions* (1944), Cambridge University Press.

⁽²⁾ La vitesse s'exprime en fonction de \mathbf{p} par $\mathbf{v} = \frac{\mathbf{p}}{\sqrt{\|\mathbf{p}\|^2 + m^2}}$ (14.27).

⁽³⁾ Voir le livre de Sygne, *The relativistic gas* (North Holland, 1957).

que la composante temporelle du vecteur Γ , et qui redonne la température propre dans le cas du repère propre, on est conduit à la formule de Planck

$$7.134) \quad T' = T \sqrt{1 - v^2} \quad (v = \text{vitesse du vent})$$

formule qui a été récemment contestée par divers auteurs; il nous semble que cette contestation est un peu académique, puisque la formule (17.134) n'est qu'une définition — assez inutile par surcroît. Il suffit de noter que le vecteur température Γ (1) définit à la fois la température propre, la vitesse du vent et le sens d'écoulement du temps.

Dans le cas général $\Lambda \neq 0$, le gaz ne peut se trouver qu'aux points X de l'espace-temps E_4 tels que le vecteur

$$7.136) \quad \Gamma_X \equiv Z_{E_4}(X) \equiv \Lambda \cdot X + \Gamma$$

soit du genre *passé*; au voisinage de ce point, la répartition de vitesse est la même que dans le cas précédent, Γ_X étant le vecteur-température.

On décrira par exemple une *centrifugeuse relativiste* en choisissant $\beta = \gamma = 0$ dans la définition (13.56) de Z , ce qui donne le vecteur température

$$7.137) \quad \Gamma_X = \mathcal{R} \begin{pmatrix} \boldsymbol{\omega} \times \mathbf{r} \\ \varepsilon \end{pmatrix}$$

ce qui montre que la vitesse du vent est

$$7.138) \quad \mathbf{v} = \boldsymbol{\omega}' \times \mathbf{r} \quad \left(\boldsymbol{\omega}' = \frac{\boldsymbol{\omega}}{\varepsilon} \right)$$

on trouve donc le champ des vitesses d'un solide tournant *non relativiste* (2).

— La formule (17.137) indique aussi l'existence d'un effet nouveau : la température propre T du fluide *n'est pas constante*; elle est donnée, en fonction de sa valeur T_0 sur l'axe, par

$$7.139) \quad T = \frac{T_0}{\sqrt{1 - \|\mathbf{v}\|^2}}$$

ÉQUILIBRES STATISTIQUES DE PHOTONS

7.140) Etudions, dans un référentiel \mathcal{R} , une particule de masse nulle (14.29); on sait définir la trajectoire de la particule relative à \mathcal{R} (14.33); le point

(1) Le vecteur température défini par Van Dantzig, Bergmann, Costa de Beauregard vaut $-kT$.

(2) On rappelle qu'il n'existe pas de définition relativiste d'un solide accéléré (voir (13.69)).

$X = \mathcal{R} \begin{pmatrix} \mathbf{r} \\ 0 \end{pmatrix}$ qui appartient à cette trajectoire définit la « position » \mathbf{r} de la particule à la date $t = 0$; on peut repérer tout mouvement x par \mathbf{r} et par l'impulsion \mathbf{p} (13.79); en appliquant (13.59), (13.87), (13.90), (14.29), (14.33), on trouve

$$7.141, 142) \quad \begin{array}{|c|c|} \hline I \equiv \mathcal{R} \begin{pmatrix} \mathbf{p}\eta \\ \|\mathbf{p}\| \end{pmatrix} & J \equiv \mathcal{R} \begin{pmatrix} \mathbf{p}\eta \\ -\|\mathbf{p}\| \end{pmatrix} \frac{1}{2\|\mathbf{p}\|^2} \\ \hline \Omega \equiv \mathcal{R} \begin{pmatrix} j \left(\mathbf{p} \frac{\eta}{\|\mathbf{p}\|} \right) & 0 \\ 0 & 0 \end{pmatrix} \mathcal{R}^{-1} & P \equiv \mathcal{R} \begin{pmatrix} \mathbf{p} \\ \eta \|\mathbf{p}\| \end{pmatrix} \\ \hline \end{array}$$

7.143, 144)

d'où, en portant dans (14.29) :

7.145)

$$\sigma(dx) (\delta x) \equiv \langle d\mathbf{p}, \delta\mathbf{r} \rangle - \langle \delta\mathbf{p}, d\mathbf{r} \rangle - \frac{\chi s \eta}{\|\mathbf{p}\|^3} \langle \mathbf{p}, d\mathbf{p} \times \delta\mathbf{p} \rangle$$

η étant le signe de l'énergie, χ l'hélicité, s le spin.

Pour calculer la densité symplectique φ (16.94), il est commode de remarquer que — pour toute base S — $\varphi(S)$ est la racine carrée du déterminant de la matrice des composantes σ_{jk} de la forme de Lagrange dans la base S ; en appliquant (17.145), on en déduit

7.146)

$$\varphi(dx) \equiv \varphi(d\mathbf{p}) \varphi(d\mathbf{r})$$

φ désignant au second membre la densité euclidienne de R^3 ; on voit que cette expression est la même que si $\begin{pmatrix} \mathbf{p} \\ \mathbf{r} \end{pmatrix} \mapsto x$ était une carte canonique (ce qui n'est pas le cas), et que les nombres χ, s, η n'interviennent pas dans cette formule.

Notons que l'énergie E de la particule est donnée par la formule d'Einstein

7.147)

$$E = \eta \|\mathbf{p}\|$$

qui résulte immédiatement de la confrontation de (13.59) et (17.144), ou encore de (14.34).

Nous pouvons alors étudier l'équilibre statistique d'une telle particule à la température T : en utilisant le principe (17.123), avec les notations (17.124), (17.126), (17.127), on trouve la « probabilité élémentaire »

$$\exp\left(-z - \frac{n \|\mathbf{p}\|}{kT}\right) \varphi(d\mathbf{p}) \varphi(d\mathbf{r})$$

z étant le potentiel thermodynamique que nous calculerons plus loin ; on voit qu'on aura convergence de l'intégrale si la particule est dans une boîte de volume fini V , et si les nombres η et T sont de même signe ⁽¹⁾ ; nous supposons bien entendu $T > 0$; on voit que la possibilité de l'équilibre thermodynamique *exclut l'existence de particules d'énergie négative* ; par contre l'hélicité est arbitraire, les particules peuvent aussi bien être polarisées à droite qu'à gauche.

En remplaçant maintenant η par 1, on déduit facilement de (17.148) la densité de probabilité de l'énergie ⁽²⁾

$$\frac{1}{2[kT]^3} e^{-E/kT} E^2 dE \quad (E > 0) . \blacksquare$$

Essayons de décrire le rayonnement à l'intérieur d'une boîte maintenue à la température T (*rayonnement du corps noir*) par un équilibre thermodynamique de photons, traités comme ci-dessus. On peut, comme dans le cas d'un gaz parfait, traiter le problème par la théorie de la diffusion, ce qui permet de ne pas tenir compte de la structure intime des interactions entre les photons et les parois de la boîte ; mais il y a, avec le cas des gaz, une différence essentielle : on ne peut pas pomper de lumière dans la boîte, *le nombre des photons est imposé par la nature*.

Suivant quelle loi ? Nous allons faire l'hypothèse que la lumière est un système dynamique dont la variété des mouvements U_ϕ est composée, au sens (16.8) ; elle contient non seulement la variété U_1 (de dimension 6) qui correspond à un photon (14.29), mais aussi toutes les variétés U_N , de dimension $6N$, décrivant le mouvement de N photons, étudiées en (15.45).

Nous admettrons même — à titre d'hypothèse — la variété U_0 , de dimension 0, réduite à un seul point, décrivant le *vide photonique*. Nous appellerons *variété de Fock* la variété U_ϕ ainsi composée.

⁽¹⁾ Notons la cohérence de ce résultat avec (14.77) : l'inversion temporelle, qui change le signe de l'énergie, change aussi celui de la température (17.135).

⁽²⁾ L'énergie E et la vitesse $\frac{\mathbf{p}}{E}$ sont des variables aléatoires indépendantes (la vitesse étant répartie uniformément sur la sphère unité).

(17.153) Puisque chaque variété pure U_N est symplectique, on sait définir sa mesure de Liouville (pour $N = 0$, on choisira la mesure définie de (16.70)) ; on définira la mesure de Liouville μ de U_ϕ par la formule (notation (16.100))

$$\mu_{U_N} = \text{mesure de Liouville de } U_N \quad (\forall N)$$

ce qui permet d'appliquer le principe de Gibbs (17.24) ou (17.123) ; en tenant compte du fait que la variété U_1 comporte deux composantes connexes, correspondant aux deux valeurs de l'hélicité, on trouve

$$(17.154) \quad I_0 = \sum_{N=0}^{\infty} \frac{[16 \pi V k^3 T^3]^N}{N!}$$

d'où — en reconnaissant le développement en série d'une exponentielle —

$$(17.155) \quad z = 16 \pi V [kT]^3$$

les formules (17.26) permettent d'en déduire la valeur moyenne \tilde{E} de l'énergie totale contenue dans la boîte

$$(17.156) \quad \tilde{E} = - \frac{dz}{d[1/kT]} = 48 \pi V [kT]^4$$

et l'entropie

$$(17.157) \quad s = z + \frac{\tilde{E}}{kT} = 64 \pi V [kT]^3 ;$$

comme en (17.54), on peut calculer l'énergie moyenne en fonction de l'entropie et du volume :

$$(17.158) \quad \tilde{E} = \frac{3}{16} \pi^{-1/3} s^{4/3} V^{-1/3}$$

d'où

$$(17.159) \quad dQ = kT ds \quad dW = - 16 \pi [kT]^4 dV$$

Q désignant la chaleur, W le travail (formules (17.57), (17.58)); on voit apparaître une *pression de radiation* (cf. (17.61)) :

$$p = 16 \pi [kT]^4 = \frac{1}{3} \frac{\tilde{E}}{V}.$$

Notons que le nombre N des photons est une variable aléatoire; la probabilité associée à chaque valeur de N est

$$p_N = e^{-z} \frac{z^N}{N!};$$

on vérifie que la valeur moyenne de N est

$$\tilde{N} = z \quad (1).$$

On peut aussi calculer la probabilité $p_{n,n'}$ pour qu'il y ait $n + n'$ photons, dont n polarisés à gauche, n' polarisés à droite :

$$p_{n,n'} = \frac{e^{-z}}{n! n'!} \left[\frac{z}{2} \right]^{n+n'}.$$

En se reportant à (17.30), on constate que ces formules ne sont pas homogènes du point de vue dimensionnel : elles dépendent de l'unité choisie pour l'action; comme on reconnaît dans (17.156) la loi de Stefan-Boltzmann

$$\frac{\tilde{E}}{VT^4} = \text{Cte}$$

on peut donc utiliser la valeur expérimentale de la constante de Stefan pour choisir cette unité; ce qui donne la valeur

$$6,45 \cdot 10^{-27} \text{ g cm}^2 \text{ sec}^{-1};$$

alors la formule (17.160) qui donne la pression de radiation est, elle aussi, conforme à l'expérience.

(1) La loi de probabilité (17.161) s'appelle *loi de Poisson*.

On pourrait donc présumer que les autres formules ci-dessus sont bonnes, et calculer par exemple le nombre moyen de photons contenus dans une boîte d'un litre à la température ordinaire (300 °K)

$$(17.166) \quad \tilde{N} = z \approx 5 \cdot 10^{11},$$

la probabilité pour que tous ces photons soient polarisés à gauche

$$(17.167) \quad p = e^{-z/2} \approx 10^{-10^{11}}$$

etc.

(17.168) Mais la formule (17.150), qui donne la répartition d'énergie des photons individuels (1) n'est pas conforme à l'expérience, qui vérifie de façon très précise la loi proposée par Planck; la figure (17.IV), donne la densité de probabilité de la variable $u = E/kT$, pour la statistique ci-dessus ($\frac{1}{2} u^2 e^{-u}$) et pour la loi de Planck ($\frac{15}{\pi^4} u^3 / [e^u - 1]$) (2).

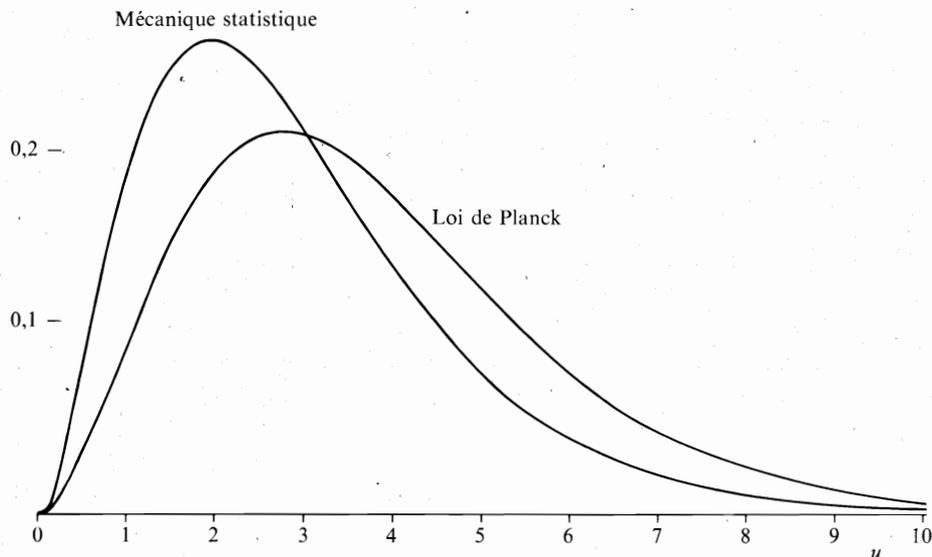


Fig. 17.IV.

(1) L'effet photoélectrique fournit, en principe, un moyen pour mesurer directement l'énergie d'un photon; la mécanique quantique permet de la mesurer indirectement par la formule $E = h\nu$, h étant la constante de Planck et ν la fréquence de la lumière associée; on voit que la figure (17.IV) est un *spectre*.

(2) Lorsque Planck proposa cette formule (décembre 1900), la notion de photon n'existait pas encore (elle fut créée par Einstein en 1905); la physique ne proposait alors, pour le rayonnement du corps noir, que des modèles hybrides utilisant à la fois la mécanique statistique classique et les équations de Maxwell (qui représentent des photons déjà quantifiés); telle la formule de Rayleigh-Jeans qui conduit à une *catastrophe ultraviolette*.

(17.169) Comme dans le cas des gaz parfaits et des solides, la comparaison de la mécanique statistique et de l'expérience, sans mettre en cause les modèles classiques proposés pour les particules ⁽¹⁾, montre donc la nécessité d'un traitement quantique des phénomènes.

⁽¹⁾ Si on ne postulait pas l'*indiscernabilité des photons*, au sens (15.47), le terme $N!$ disparaîtrait du dénominateur de (17.154) ; ce qui conduirait à remplacer la loi de Stefan-Boltzmann (17.156) par

$$\tilde{E} = \frac{48 \pi V [kT]^4}{1 - 16 \pi V [kT]^3}$$

loi complètement inacceptable, car elle conduit à une énergie infinie pour des températures de l'ordre du °K (dans le cas $V = 1 \text{ cm}^3$).