

Chapitre premier
Notions de Géométrie différentielle

1. Introduction : systèmes et modélisation mathématique

1.1. Généralités. — On désigne ici par *système* un objet du monde réel, pouvant éventuellement interagir avec d'autres objets, dont l'état évolue au cours du temps. Par exemple, une galaxie, un système planétaire, sont des systèmes étudiés en Astronomie. Un haut fourneau en fonctionnement, un gisement d'hydrocarbures en exploitation, un avion en vol sont des systèmes dont l'étude fait partie du domaine d'activité de l'ingénieur. L'atmosphère de la Terre, les océans, la croûte terrestre, sont des systèmes étudiés par le météorologue, l'océanologue, le géologue ou le géophysicien. Un organisme vivant, une vigne de Bourgogne, la société française, un cerveau humain peuvent aussi, dans une certaine mesure, être considérés comme des systèmes, si l'on admet qu'il existe des lois (encore très incomplètement connues) qui régissent leur évolution.

Depuis toujours, l'homme cherche à comprendre l'évolution des systèmes qui l'entourent, et dans certains cas tente de contrôler cette évolution. Une des formes prises par cet effort de compréhension est la modélisation mathématique. Elle consiste à formuler en termes mathématiques les lois qui régissent l'évolution d'un système, puis à utiliser des méthodes mathématiques pour étudier cette évolution. Bien entendu, la modélisation mathématique n'est pas applicable à tous les systèmes que l'homme s'efforce de comprendre : elle ne peut être employée que pour ceux dont l'évolution est régie par des lois suffisamment bien connues pour être formulées en termes mathématiques. Cependant, son champ d'application couvre déjà une large partie des domaines d'activité des Astronomes, Physiciens et Ingénieur, et ne cesse de s'étendre.

Les modèles mathématiques utilisés en Mécanique et en Physique font de plus en plus souvent appel à des notions de géométrie différentielle qui, malheureusement, ne sont pas toujours enseignées dans les second cycles des Universités ou les écoles d'ingénieurs. Nous allons présenter certaines de ces notions, en nous efforçant de faire comprendre les raisons de leur importance.

1.2. L'espace des états d'un système ; quelques exemples. — La première étape d'une modélisation mathématique est généralement le choix d'un espace

représentant l'ensemble des états possibles du système étudié. Le choix judicieux de cet espace suffit parfois à rendre évidente la réponse à une question apparemment très difficile (voir l'exemple dû à N. Konstantinov, présenté par V. Arnol'd dans son livre "Équations différentielles ordinaires", pages 12-13). Ainsi qu'on le verra ci-dessous, les variétés différentiables apparaissent très naturellement comme espaces d'états possibles de systèmes.

Dans de nombreux cas, l'état d'un système peut être caractérisé par les valeurs d'une ou plusieurs grandeurs qui, une fois des unités choisies, sont représentées par des nombres réels. Donnons quelques exemples.

a) Point matériel mobile. — L'état dynamique (position et vitesse) d'un mobile assimilé à un point matériel peut, moyennant le choix d'un repère de l'espace physique et d'une unité de temps, être représenté par six nombres réels : les trois coordonnées du mobile et les trois composantes de sa vitesse.

b) Réacteur chimique. — L'état d'un réacteur chimique de volume fixé, contenant un mélange, supposé bien homogène, de n constituants, est décrit par la donnée de la pression, de la température et des concentrations massiques en $n - 1$ constituants (puisque la somme des concentrations massiques en tous les constituants est égale à 1); l'unité de pression et l'échelle de température une fois choisies, l'état du réacteur est donc décrit par la donnée de $2 + n - 1 = n + 1$ nombres réels.

Dans ces deux exemples, on peut prendre pour espace des états du système une partie de l'espace numérique \mathbb{R}^p (ensemble des multiuplets de p nombres réels); on a $p = 6$ dans le premier exemple, et $p = n + 1$ dans le second. Cependant, dans le premier exemple, on remarque que l'emploi de repères de l'espace physique et de l'espace des vitesses a quelque chose d'arbitraire; en procédant ainsi, on laisse de côté certaines propriétés géométriques du système (par exemple, l'invariance par changement de repère galiléen) qu'il pourrait être souhaitable d'introduire dans son modèle mathématique dès le choix de l'espace des états. De même, le choix d'unités, qui permet d'utiliser des nombres réels pour représenter les valeurs de grandeurs physiques, est dans une large mesure arbitraire et risque de masquer certaines propriétés intéressantes du système, telles que l'invariance vis-à-vis des changements d'unités (dont l'étude fait l'objet de l'analyse dimensionnelle). C'est pourquoi il est à la fois plus naturel et plus efficace de renoncer à décrire l'état d'un système uniquement au moyen de nombres réels, et d'utiliser pour cette description des éléments d'espaces ayant une structure plus riche. Bien entendu, après le choix de repères, d'unités et d'autres objets définis plus loin (les *cartes* des variétés différentiables) les éléments de ces espaces pourront eux-mêmes être représentés par un ou plusieurs nombres réels; mais cette représentation dépendra du choix des repères, des unités et des cartes.

Souvent, l'espace des états d'un système, indépendamment de tout choix de repères

ou d'unités, est une partie d'un espace vectoriel ou affine de dimension finie. C'est le cas dans les deux exemples mentionnés ci-dessus. Mais il en va tout autrement dans d'autres exemples guère plus compliqués. Nous allons en décrire quelques-uns, afin de montrer l'utilité de la notion de variété différentiable, que nous définirons ensuite.

c) *Pendule plan, configuration.* — Considérons un pendule plan, c'est-à-dire un point matériel soumis à la pesanteur, se déplaçant sur un cercle, dans un plan vertical. L'ensemble des positions du pendule est évidemment le cercle sur lequel se déplace le point matériel (qu'on peut identifier au cercle trigonométrique standard S^1). On peut, bien entendu, repérer chaque point de ce cercle au moyen d'un nombre réel (par exemple au moyen de l'angle au centre, ou de l'abscisse curviligne, à partir d'une position de référence). Si l'on utilise l'angle au centre (resp., l'abscisse curviligne), deux valeurs différant d'un multiple de 2π (resp., d'un multiple de la longueur de la circonférence) correspondent à la même position du pendule. Le cercle peut être identifié à un intervalle de \mathbb{R} , par exemple en choisissant la détermination de l'angle au centre appartenant à l'intervalle semi-ouvert $[0, 2\pi[$, mais au prix de l'introduction d'une discontinuité, ce qui est peu satisfaisant. On voit bien que les propriétés globales du cercle sont tout à fait différentes de celles d'un segment de droite.

d) *Pendule plan, états cinématiques.* — Intéressons-nous maintenant à l'ensemble des *états cinématiques* du pendule plan de l'exemple précédent. Un état cinématique est constitué par une position du pendule et une valeur de sa vitesse angulaire. On voit aisément que l'ensemble des vitesses angulaires possibles est un espace vectoriel de dimension 1, qu'on peut identifier à la droite réelle \mathbb{R} . L'ensemble des états cinématiques du pendule plan s'identifie donc au produit $S^1 \times \mathbb{R}$; c'est un cylindre.

e) *Pendule sphérique.* — Considérons maintenant un pendule sphérique, c'est-à-dire un point matériel pesant astreint à se déplacer sur une sphère. Son espace de configuration, c'est-à-dire l'ensemble de ses positions possibles, est évidemment cette sphère, qu'on peut identifier à la sphère standard S^2 . Un état cinématique du pendule sphérique est constitué par une position possible du point matériel, et une vitesse possible de ce point lorsqu'il occupe cette position. Celle-ci peut être identifiée (moyennant le choix d'une unité de temps) à un vecteur tangent à la sphère au point représentant la position du pendule. L'ensemble des états cinématiques du système est donc l'ensemble des couples (m, \vec{v}) formés par un point m de la sphère S^2 , et un vecteur \vec{v} de l'espace à trois dimensions dans lequel cette sphère est plongée, tangent à la sphère S^2 au point m , c'est-à-dire orthogonal à la droite qui joint m au centre de la sphère. Contrairement au cas du pendule plan, on remarque que l'espace des états cinématiques du pendule sphérique n'est

pas le produit de son espace de configuration et d'un autre espace, car le plan auquel doit appartenir le vecteur \vec{v} dépend du point m . On verra plus loin que cet espace est le *fibré tangent* à la sphère S^2 .

f) *Toupie de Lagrange*. — Considérons maintenant une toupie dont la pointe est fixée, c'est-à-dire un solide tridimensionnel mobile autour d'un point fixe O . La manière la plus simple d'identifier l'ensemble de ses positions possibles consiste à en choisir une de référence P_0 , et à remarquer que pour toute autre position P , il existe une rotation unique de l'espace autour du point fixe O qui envoie P_0 sur P . L'ensemble des positions du solide s'identifie donc, moyennant le choix d'une position de référence, à l'ensemble des rotations de l'espace physique autour du point O . C'est un groupe (car on peut composer deux rotations, et prendre l'inverse d'une rotation donnée) qu'on note habituellement $\mathbf{SO}(3)$. On peut donc dire que l'espace de configuration de la toupie (ensemble de ses positions possibles) est le groupe $\mathbf{SO}(3)$.

Un *état cinématique* du système est la donnée d'une position possible et d'un champ de vitesses possible de la toupie dans cette position. On verra plus loin que moyennant le choix d'une unité de temps, l'ensemble des états cinématiques de la toupie est le *fibré tangent* au groupe $\mathbf{SO}(3)$.

Le groupe $\mathbf{SO}(3)$ peut être paramétré au moyen de nombres réels de diverses manières. L'une des plus connues fait appel aux angles d'Euler, ainsi définis. On attache un repère orthonormé, d'origine O , au corps solide; on note, $(O, \vec{X}, \vec{Y}, \vec{Z})$ et $(O, \vec{x}, \vec{y}, \vec{z})$ les positions de ce repère dans l'espace physique lorsque le solide occupe la position de référence P_0 , et une position courante P , respectivement. Soit θ l'angle formé par les demi-axes $O\vec{Z}$ et $O\vec{z}$ (à valeurs dans l'intervalle fermé $[0, \pi]$). Si $0 < \theta < \pi$, les plans $O\vec{X}\vec{Y}$ et $O\vec{x}\vec{y}$ se coupent selon une droite bien définie; soit \vec{u} un vecteur unitaire parallèle à leur intersection. Si θ est égal à 0 ou à 2π , on choisit pour \vec{u} un vecteur unitaire quelconque parallèle au plan $O\vec{X}\vec{Y}$. Soient φ l'angle $(O\vec{X}, O\vec{u})$ et ψ l'angle $(O\vec{u}, O\vec{x})$. Ce sont des angles de demi-droites, qu'on peut considérer comme des réels définis modulo 2π (les plans ayant pour repères (O, \vec{X}, \vec{Y}) et (O, \vec{x}, \vec{y}) étant munis de l'orientation définie par ces repères). Les angles d'Euler sont θ , φ et ψ . Leur donnée détermine de manière unique la position P , lorsqu'on connaît la position de référence P_0 . Mais réciproquement, une même position P peut correspondre à plusieurs valeurs des angles d'Euler (puisque \vec{u} peut être remplacé par son opposé, que dans certains cas, lorsque $\vec{z} = \pm\vec{Z}$, il peut même être choisi de manière quelconque dans le plan orthogonal à \vec{Z} , et que φ et ψ sont définis modulo 2π).

Une autre paramétrisation du groupe $\mathbf{SO}(3)$ consiste à utiliser, pour représenter la rotation qui applique le repère orthonormé $(O\vec{X}\vec{Y}\vec{Z})$ sur le repère orthonormé de même orientation $(O\vec{x}\vec{y}\vec{z})$, la matrice M , à 3 lignes et 3 colonnes, qui représente

cette transformation : c'est la matrice dont les colonnes sont constituées par les composantes des vecteurs \vec{x} , \vec{y} et \vec{z} dans la base formée par les vecteurs \vec{X} , \vec{Y} , \vec{Z} . On sait que la matrice M est orthogonale et de déterminant 1 ; ses 9 coefficients ne sont donc pas indépendants, ils sont astreints à vérifier 6 égalités (exprimant que les vecteurs \vec{x} , \vec{y} et \vec{z} sont unitaires et deux à deux orthogonaux) et 1 inégalité (exprimant que le déterminant de M est positif).

Il existe encore d'autres paramétrisations du groupe $\mathbf{SO}(3)$ utiles en physique et en mécanique, par exemple celle utilisant des quaternions unitaires, et celle utilisant le fibré unitaire tangent à une sphère, que nous allons décrire. Le repère $(O, \vec{x}, \vec{y}, \vec{z})$, donc aussi l'élément du groupe $\mathbf{SO}(3)$ qui transforme le repère de référence $(O, \vec{X}, \vec{Y}, \vec{Z})$ en ce repère, sont déterminés lorsqu'on connaît les vecteurs unitaires \vec{z} et \vec{x} , puisque le vecteur \vec{y} est orthogonal au plan formé par les deux précédents, et de sens tel que les trièdres $(\vec{X}, \vec{Y}, \vec{Z})$ et $(\vec{x}, \vec{y}, \vec{z})$ soient de même orientation. Le vecteur \vec{z} étant unitaire, son extrémité est un point m de la sphère S^2 , centrée sur O et de rayon 1. Pour une position donnée de \vec{z} , ou, ce qui est équivalent, de son extrémité m , le vecteur unitaire \vec{x} est orthogonal à \vec{z} ; il peut donc être considéré comme un vecteur unitaire tangent en m à la sphère S^2 . On voit ainsi que le groupe $\mathbf{SO}(3)$ peut être identifié à l'ensemble des couples (m, \vec{x}) formés d'un point m de la sphère S^2 et d'un vecteur unitaire \vec{x} tangent en m à cette sphère. Cet ensemble est appelé *fibré unitaire tangent* à la sphère S^2 . On remarque que $\mathbf{SO}(3)$ ne s'identifie pas au produit $S^2 \times S^1$, bien que l'ensemble des positions possibles de l'extrémité m de \vec{z} soit la sphère S^2 et que, pour chaque position fixée de m , l'ensemble des positions possibles du vecteur \vec{x} (unitaire et orthogonal à \vec{z}) puisse être identifié au cercle S^1 : cette identification nécessiterait le choix, de manière continue, d'une direction de référence dans le plans tangent à la sphère S^2 en chacun de ses points, ce qui est impossible ; on montre en effet qu'il ne peut exister sur S^2 de champ de vecteurs continu qui ne s'annule en aucun point.

2. Variétés différentiables

Les exemples du paragraphe précédent nous ont montré que la prise en compte des propriétés géométriques de l'espace des états d'un système ne pouvait être faite de manière satisfaisante en n'utilisant que des espaces numériques \mathbb{R}^p . Les notions de *variété différentiable* et d'*espace fibré*, que nous allons maintenant introduire, vont nous fournir des objets mathématiques beaucoup plus aptes à représenter l'espace des états de nombreux systèmes.

2.1. Définitions. —

1. Une variété topologique de dimension n (n entier positif) est un espace topologique M dont tout point possède un voisinage ouvert homéomorphe à un ouvert de \mathbb{R}^n .

2. Une carte de la variété topologique M , de dimension n , est un couple (U, φ) , où U est un ouvert de M , et φ un homéomorphisme de U sur un ouvert $V = \varphi(U)$ de \mathbb{R}^n . Un atlas de M est une famille de cartes (U_i, φ_i) , $i \in I$, ensemble d'indices, dont les domaines U_i recouvrent M , c'est-à-dire telle que $\bigcup_{i \in I} U_i = M$.

La notion de variété topologique est un peu trop générale pour les besoins les plus courants, où l'on souhaite pouvoir utiliser le calcul différentiel. On introduit donc la notion de variété différentiable, définie ci-dessous.

2.2. Définitions. —

1. Un atlas (U_i, φ_i) , $i \in I$, d'une variété topologique M de dimension n est dit différentiable de classe C^p (p entier ≥ 1 , ou $p = \infty$) si pour tout couple $(i, j) \in I^2$ tel que $U_i \cap U_j \neq \emptyset$, l'application (dite changement de carte) $\varphi_j \circ \varphi_i^{-1}$ est différentiable de classe C^p . Cet atlas est dit analytique ou de classe C^ω si les applications de changement de carte sont analytiques.

2. Deux atlas (U_i, φ_i) , $i \in I$, et (U'_j, φ'_j) , $j \in J$, de classe C^p , de la même variété topologique M , sont dits C^p -équivalents si leur réunion est encore un atlas de classe C^p , c'est-à-dire si pour tout couple $(i, j) \in I \times J$ tel que $U_i \cap U'_j \neq \emptyset$, les applications $\varphi'_j \circ \varphi_i^{-1}$ et $\varphi_i \circ \varphi'_j^{-1}$ sont différentiables de classe C^p .

3. Une variété différentiable de dimension n et de classe C^p (avec p entier ≥ 1 , $p = \infty$ ou $p = \omega$) est une variété topologique de dimension n munie d'une classe d'équivalence (au sens de la C^p -équivalence définie ci-dessus) d'atlas de classe C^p . Une variété différentiable de classe C^ω est aussi appelée variété analytique.

4. Une variété différentiable étant donnée, les cartes appartenant à un atlas faisant partie de la classe d'équivalence qui définit sa structure différentiable sont dites cartes admissibles (pour cette structure).

2.3. Commentaires et exemples

a) On impose souvent aux variétés différentiables d'être des espaces topologiques séparés (c'est-à-dire tels que pour tout couple (x, y) de points distincts de cet espace, il existe un voisinage de x et un voisinage de y disjoints). Cette propriété ne résulte pas des définitions ci-dessus (bien que celles-ci montrent que tout point d'une variété topologique, ou différentiable, possède un voisinage séparé). Sauf mention explicite contraire, nous supposons dans ce qui suit que les variétés considérées sont séparées.

b) Les variétés que nous considérerons seront toujours supposées avoir un atlas fini ou dénombrable. Ces variétés sont alors des espaces topologiques à base dénombrable d'ouverts.

c) Soit M une variété différentiable de dimension n , et (U, φ) une carte admissible. L'application φ , à valeurs dans \mathbb{R}^n , a n composantes; on la notera donc $\varphi = (x^1, \dots, x^n)$. Les n fonctions x^i (définies sur l'ouvert U de M et à valeurs réelles) sont appelées *coordonnées locales* associées à la carte (U, φ) . Soit (V, ψ) une autre carte admissible de M , telle que $U \cap V \neq \emptyset$, et (y^1, \dots, y^n) les coordonnées locales associées. L'application changement de carte $\psi \circ \varphi^{-1}$ n'est autre que l'application $(x^1, \dots, x^n) \mapsto (y^1, \dots, y^n)$ exprimant les coordonnées locales associées à la carte (V, ψ) en fonction des coordonnées locales associées à la carte (U, φ) .

d) Un ouvert de \mathbb{R}^n est une variété analytique de dimension n : il suffit en effet de l'équiper de l'atlas comportant une carte unique, l'application identique. De même, un ouvert d'un espace vectoriel (ou affine) de dimension n est une variété analytique de dimension n : le choix d'un repère nous ramène en effet immédiatement au cas précédent.

e) Un ouvert V d'une variété différentiable M de dimension n et de classe C^p est une variété différentiable de même dimension et de même classe. Pour toute carte admissible (U, φ) de M telle que $U \cap V$ soit non vide, $(U \cap V, \varphi|_{U \cap V})$ est une carte admissible de V .

f) Pour définir une structure de variété différentiable de classe C^p et de dimension n sur un ensemble M , sans nécessairement avoir préalablement défini une structure topologique sur M , on peut utiliser une famille de parties U_i ($i \in I$, ensemble d'indices) de M dont la réunion est M , et, pour chaque i , une application bijective φ_i de U_i sur un ouvert $V_i = \varphi_i(U_i)$ de \mathbb{R}^n . On doit alors vérifier que les applications de changement de carte $\varphi_j \circ \varphi_i^{-1}$ sont des difféomorphismes de classe C^p . On montre que lorsque c'est le cas, il existe sur M une unique structure de variété différentiable de classe C^p admettant les (U_i, φ_i) pour cartes admissibles. C'est la méthode employée dans les exemples (g) et (i) ci-dessous.

g) Considérons le cercle trigonométrique S^1 , c'est-à-dire l'ensemble des points (x, y) du plan \mathbb{R}^2 tels que $x^2 + y^2 = 1$. Soit (x_0, y_0) un point de ce cercle. Choisissons un réel θ_0 tel que $x_0 = \cos \theta_0$, $y_0 = \sin \theta_0$. Ce réel est une détermination particulière de l'angle polaire du point (x_0, y_0) ; il n'est pas unique, puisqu'on peut lui ajouter un multiple quelconque de 2π ; on choisit arbitrairement une de ses valeurs possibles. Considérons alors l'application de l'intervalle ouvert $]\theta_0 - \pi, \theta_0 + \pi[$ dans le cercle, qui au réel θ associe le point $(\cos \theta, \sin \theta)$. Elle est injective, et a pour image le cercle S^1 privé d'un seul point, le point diamétralement opposé à (x_0, y_0) . On notera $U_{(x_0, y_0)}$ cette image, et $\varphi_{(x_0, y_0)}$ l'application inverse de celle définie ci-dessus. L'application $\varphi_{(x_0, y_0)}$ associe à chaque point (x, y) de

$U_{(x_0, y_0)}$ une détermination particulière θ de l'angle polaire de ce point, cette détermination étant continue sur $U_{(x_0, y_0)}$ et prenant la valeur θ_0 au point (x_0, y_0) . On définit la structure de variété différentiable de S^1 en prenant pour cartes les $(U_{(x_0, y_0)}, \varphi_{(x_0, y_0)})$, pour plusieurs choix distincts du point (x_0, y_0) (au moins deux, afin que les domaines de ces cartes recouvrent S^1). Les applications de changement de carte sont simplement des changements de détermination de l'angle polaire ; elles sont de la forme $\theta' = \theta + 2k\pi$, avec $k \in \mathbf{Z}$, entier fixé. Ce sont des applications analytiques. Le cercle S^1 est donc une variété analytique de dimension 1.

h) Si M et M' sont des variétés différentiables de classe C^p , de dimensions respectives n et n' , le produit $M \times M'$ est une variété de classe C^p , de dimension $n + n'$. En effet, si les (U_i, φ_i) , $i \in I$ forment un atlas de M et les (V_j, ψ_j) , $j \in J$, un atlas de M' , les $(U_i \times V_j, (\varphi_i, \psi_j))$, $(i, j) \in I \times J$, forment un atlas de $M \times M'$. Par exemple, le cylindre $S^1 \times \mathbb{R}$ est une variété différentiable de dimension 2.

i) La sphère S^2 est l'ensemble des points (x, y, z) de \mathbb{R}^3 qui vérifient $x^2 + y^2 + z^2 = 1$. On note U_+ , V_+ et W_+ (resp., U_- , V_- , W_-) l'ensemble des points (x, y, z) de S^2 tels que, respectivement, $x > 0$, $y > 0$, $z > 0$ (resp., $x < 0$, $y < 0$, $z < 0$). On note φ_+ , χ_+ et ψ_+ les applications, respectivement de U_+ , V_+ et W_+ dans \mathbb{R}^2 qui, au point (x, y, z) associent, respectivement, (y, z) , (x, z) et (x, y) . On définit, de la même manière, φ_- , χ_- et ψ_- , définies, respectivement, sur U_- , V_- et W_- . La structure différentiable de S^2 est définie par l'atlas formé par les cartes (U_+, φ_+) , (V_+, χ_+) , (W_+, ψ_+) , (U_-, φ_-) , (V_-, χ_-) , (W_-, ψ_-) . Les applications de changement de carte sont analytiques : en effet, $\psi_+ \circ \varphi_-^{-1}$, par exemple, a pour expression

$$(y, z) \mapsto (-\sqrt{1 - y^2 - z^2}, y).$$

La sphère S^2 est donc une variété analytique de dimension 2.

Dans la suite de ce paragraphe nous considérerons, pour simplifier, seulement des variétés différentiables de classe C^∞ . Mais les concepts que nous introduirons gardent leur sens pour les variétés différentiables de classe C^p . Nous laissons au lecteur le soin d'adapter les énoncés pour ce cas.

2.4. Définitions. — Soient M et N deux variétés différentiables de classe C^∞ , de dimensions m et n , respectivement, et f une application continue de M dans N .

1. On dit que f est différentiable de classe C^∞ au voisinage d'un point a de M si, pour une carte admissible (U, φ) de M telle que $a \in U$ et une carte admissible (V, ψ) de N telle que $f(a) \in V$, l'application $\psi \circ f \circ \varphi^{-1}$ (définie sur l'ouvert $\varphi(U)$ de \mathbf{R}^m et à valeurs dans \mathbf{R}^n) est différentiable de classe C^∞ au voisinage du point $\varphi(a)$.

2. On dit que f est différentiable de classe C^∞ si elle l'est au voisinage de tout point de M .

3. On dit que f est un *difféomorphisme de classe C^∞* si elle est différentiable de classe C^∞ , bijective, et d'inverse $f^{-1} : N \rightarrow M$ différentiable de classe C^∞ (ce qui implique $m = n$).

2.5. Commentaire. — La notion d'application différentiable est illustrée sur la figure 1.

Figure 1. Application différentiable d'une variété M dans une variété N .

L'application $\psi \circ f \circ \varphi^{-1}$ est appelée *expression* de l'application f dans les cartes (U, φ) et (V, ψ) . Il importe de remarquer que la notion d'application différentiable ne dépend pas du choix des cartes admissibles (U, φ) et (V, ψ) . Soient en effet (U', φ') et (V', ψ') des cartes admissibles, respectivement de M et de N , avec $a \in U'$, $f(a) \in V'$. On peut écrire

$$\psi' \circ f \circ \varphi'^{-1} = (\psi' \circ \psi^{-1}) \circ (\psi \circ f \circ \varphi^{-1}) \circ (\varphi \circ \varphi'^{-1}).$$

Les applications de changement de carte $\psi' \circ \psi^{-1}$ et $\varphi \circ \varphi'^{-1}$ sont différentiables de classe C^∞ . Donc si $\psi \circ f \circ \varphi^{-1}$ est différentiable de classe C^∞ au voisinage du point $\varphi(a)$, $\psi' \circ f \circ \varphi'^{-1}$ est différentiable de classe C^∞ au voisinage du point $\varphi' \circ \varphi^{-1}(\varphi(a)) = \varphi'(a)$, comme composée d'applications de classe C^∞ .

Dans ce qui suit, nous sous-entendons souvent le qualificatif “de classe C^∞ ” pour parler des variétés et applications différentiables. Nous dirons donc “différentiable” pour dire “différentiable de classe C^∞ ”.

2.6. Théorème. — Soient M , N et P trois variétés différentiables, f une application différentiable de M dans N et g une application différentiable de N dans P . L'application composée $g \circ f$, de M dans P , est différentiable.

La démonstration, qui ne présente pas de difficulté, est laissée à la bonne volonté du lecteur.

2.7. Définitions. — Soient M et N deux variétés différentiables de dimensions respectives m et n , et f une application différentiable de M dans N .

1. Soit a un point de M , (U, φ) une carte admissible de M et (V, ψ) une carte admissible de N telles que $a \in U$, $f(a) \in V$. On appelle *rang* de l'application différentiable f au point a le rang de la différentielle de $\psi \circ f \circ \varphi^{-1}$ au point $\varphi(a)$. On vérifie aisément qu'il ne dépend pas du choix des cartes (U, φ) et (V, ψ) . Il est évidemment inférieur ou égal à $\inf(m, n)$.

2. On dit que f est une *immersion* (resp., une *submersion*) au point a si son rang en ce point est égal à m (resp., à n). On dit que f est une *immersion* (resp., une *submersion*) si c'est une immersion (resp., une submersion) en tout point de M .

3. Une application qui est à la fois une immersion et une submersion est dite *étale*; on dit aussi que c'est un *difféomorphisme local*; on a bien entendu dans ce cas $m = n$.

2.8. Propriétés et exemples

a) Si l'application différentiable $f : M \rightarrow N$ est de rang constant k sur M , pour tout point a de M , il existe une carte admissible (U, φ) de M avec $a \in U$ et une carte admissible (V, ψ) de N avec $f(a) \in V$, telles que $\psi \circ f \circ \varphi^{-1}$ soit l'application

$$(x^1, \dots, x^m) \mapsto (y^1 = x^1, \dots, y^k = x^k, y^{k+1} = 0, \dots, y^n = 0).$$

En particulier, si f est une immersion, c'est-à-dire si $k = m \leq n$, $\psi \circ f \circ \varphi^{-1}$ est de la forme

$$(x^1, \dots, x^m) \mapsto (y^1 = x^1, \dots, y^m = x^m, y^{m+1} = 0, \dots, y^n = 0).$$

Si f est une submersion, c'est-à-dire si $k = n \leq m$, $\psi \circ f \circ \varphi^{-1}$ est de la forme

$$(x^1, \dots, x^m) \mapsto (y^1 = x^1, \dots, y^n = x^n).$$

b) Une application différentiable étale bijective est un difféomorphisme.

c) Soient M_1 et M_2 deux variétés différentiables. Les projections $p_1 : (x_1, x_2) \mapsto x_1$ et $p_2 : (x_1, x_2) \mapsto x_2$ du produit $M_1 \times M_2$, respectivement, sur M_1 et sur M_2 , sont des submersions.

La notion de surface, dans l'espace euclidien usuel de dimension 3, est sans doute familière au lecteur. Ainsi que nous allons le voir, c'est un cas particulier de la notion de *sous-variété* d'une variété différentiable.

2.9. Définition. — Soit M une variété différentiable de dimension m . Un sous-ensemble N de M est une *sous-variété* de M , de dimension n (avec $n \leq m$) s'il existe un atlas admissible (U_i, φ_i) , $i \in I$, de M ayant la propriété suivante : pour

tout $i \in I$ tel que $U_i \cap N \neq \emptyset$, $U_i \cap N$ est l'ensemble des points $m \in M$ dont les $m - n$ dernières coordonnées locales $x^k(m)$ ($n + 1 \leq k \leq m$), pour le système de coordonnées locales associé à la carte (U_i, φ_i) sont nulles. Un atlas de M ayant cette propriété, et une carte de cet atlas, sont dits adaptés à la sous-variété N .

2.10. Commentaire et exemples

a) Soient n et m deux entiers tels que $1 \leq n \leq m$. L'ensemble des points de \mathbb{R}^m dont les $m - n$ dernières coordonnées sont nulles est une sous-variété de \mathbb{R}^m , de dimension n , qui s'identifie à \mathbb{R}^n .

b) Une sous-variété N d'une variété différentiable M est une variété différentiable, car d'un atlas (U_i, φ_i) de M adapté à N , on déduit un atlas $(U_i \cap N, \varphi_i|_{U_i \cap N})$ de N . Pour cette structure de variété différentiable sur N , l'injection canonique de N dans M est une application différentiable, et plus précisément une immersion.

c) Soit S^2 la sphère unité de \mathbb{R}^3 , c'est-à-dire l'ensemble des points $(x, y, z) \in \mathbb{R}^3$ tels que $x^2 + y^2 + z^2 = 1$. Nous savons déjà (2.3.i) que c'est une variété différentiable. Il est facile de vérifier que c'est une sous-variété de \mathbb{R}^3 . Soient en effet \tilde{U}_+ , \tilde{V}_+ , \tilde{W}_+ les demi-espaces ouverts de \mathbb{R}^3 sur lesquels on a, respectivement, $x > 0$, $y > 0$ et $z > 0$. De même, soient \tilde{U}_- , \tilde{V}_- , \tilde{W}_- les demi-espaces ouverts sur lesquels on a, respectivement, $x < 0$, $y < 0$ et $z < 0$. On définit sur \tilde{U}_+ l'application $\tilde{\varphi}_+ : (x, y, z) \mapsto (u, v, w)$, en posant

$$u = y, \quad v = z, \quad w = x^2 + y^2 + z^2 - 1.$$

Il est facile de voir que $(\tilde{U}_+, \tilde{\varphi}_+)$ est une carte admissible de \mathbb{R}^3 et que $U_+ = \tilde{U}_+ \cap S^2$ est l'ensemble des points de \tilde{U}_+ dont la troisième coordonnée w , dans cette carte, est nulle. De plus, la restriction de $\tilde{\varphi}_+$ à $\tilde{U}_+ \cap S^2$ n'est autre (si l'on omet la troisième coordonnée locale w , qui prend la valeur 0) que l'application φ_+ définie en 2.3.i. En procédant de même avec les cinq autres demi-espaces, on obtient un atlas de \mathbb{R}^3 adapté à la sous-variété S^2 .

2.11. Définition. — On appelle *plongement* d'une variété différentiable N , de dimension n , dans une variété différentiable M de dimension $m \geq n$ une immersion (notion définie en 2.7) injective f de N dans M qui est aussi un homéomorphisme (c'est-à-dire une bijection continue d'inverse continue) de N sur $f(N)$ muni de la topologie induite par celle de M .

2.12. Propriétés et exemples

a) L'injection canonique d'une sous-variété N de M dans M est un plongement. Réciproquement, si $f : N \rightarrow M$ est un plongement, $f(N)$ est une sous-variété de M et f , considérée comme application de N sur $f(N)$, est un difféomorphisme.

b) Toute immersion injective d'une variété différentiable compacte dans une autre variété différentiable est un plongement.

c) D'après un célèbre théorème de Whitney, toute variété différentiable de dimension n peut être plongée dans un espace numérique \mathbb{R}^N , pour N assez grand ($N \geq 2n + 1$).

d) Il existe des immersions injectives qui ne sont pas des plongements. C'est le cas par exemple de l'application de \mathbb{R} dans \mathbb{R}^2 :

$$t \mapsto (\cos(2\operatorname{Arctg} t), \sin(2\operatorname{Arctg} t)),$$

car l'image de cette application est le cercle de centre 0 et de rayon 1 privé du point $(-1, 0)$, qui n'est pas une partie fermée de \mathbb{R}^2 . C'est également le cas de l'application $t \mapsto (x, y, z)$ de \mathbb{R} dans \mathbb{R}^3 , définie par

$$x = (2 + \cos t) \cos(\sqrt{2}t), \quad y = (2 + \cos t) \sin(\sqrt{2}t), \quad z = \sin t.$$

L'image de cette application n'est en effet pas fermée dans \mathbb{R}^3 ; son adhérence est le tore de dimension 2, ensemble des points (x, y, z) de la forme

$$x = (2 + \cos \theta) \cos \varphi, \quad y = (2 + \cos \theta) \sin \varphi, \quad z = \sin \theta, \quad 0 \leq \theta, \quad \varphi \leq 2\pi,$$

qu'on obtient en faisant tourner le cercle d'équations $x = 2 + \cos \theta, y = 0, z = \sin \theta$ autour de l'axe Oz .

e) Soit M une variété différentiable de dimension m . D'après 2.3.h, le produit $M \times M$ est une variété différentiable de dimension $2m$. L'application $x \mapsto (x, x)$ est un plongement de M dans $M \times M$, dont l'image est la diagonale Δ de $M \times M$, c'est-à-dire l'ensemble des points (x, y) de $M \times M$ tels que $x = y$. Celle-ci est une sous-variété de dimension m de $M \times M$.

Nous indiquons ci-dessous, sans démonstration, deux résultats très souvent employés pour reconnaître une sous-variété.

2.13. Théorème. — Soient M et P deux variétés différentiables, de dimensions respectives m et p ($m \geq p$), et $f : M \rightarrow P$ une application différentiable.

1. Soit b un point de P tel que $M_b = f^{-1}(b)$ soit non vide. Si, en tout point de M_b , le rang de f est égal à p , M_b est une sous-variété de M de dimension $m - p$.

2. Plus généralement, soit Q une sous-variété de dimension q de P ($q \leq p$) telle que $M_Q = f^{-1}(Q)$ soit non vide. Si, en tout point de M_Q , le rang de f est égal à p , M_Q est une sous-variété de M de dimension $m - p + q$.

2.14. Commentaires et exemples

a) La partie 1 du théorème précédent est un cas particulier de la partie 2, car un point b de P peut être considéré comme une sous-variété de dimension 0 de P .

b) Pour tout entier $n \geq 1$, la sphère de dimension n , définie par

$$S^n = \left\{ (x^1, \dots, x^{n+1}) \in \mathbb{R}^{n+1} \mid \sum_{i=1}^{n+1} (x^i)^2 = 1 \right\}$$

est une sous-variété de dimension n de \mathbb{R}^{n+1} . Cela résulte en effet du théorème précédent, en prenant pour f l'application de \mathbb{R}^{n+1} dans \mathbb{R} : $f(x^1, \dots, x^{n+1}) = \sum_{i=1}^{n+1} (x^i)^2$.

c) Soient $f_1 : E_1 \rightarrow M$ et $f_2 : E_2 \rightarrow M$ deux submersions, respectivement des variétés différentiables E_1 de dimension k_1 et E_2 de dimension k_2 , sur une variété M de dimension n . On appelle *produit fibré de E_1 et E_2 au dessus de M* , et on note $E_1 \times_M E_2$, le sous-ensemble du produit $E_1 \times E_2$

$$E_1 \times_M E_2 = \left\{ (x_1, x_2) \in E_1 \times E_2 \mid f_1(x_1) = f_2(x_2) \right\}.$$

Le théorème 2.13 permet de montrer que $E_1 \times_M E_2$ est une sous-variété de $E_1 \times E_2$, de dimension $k_1 + k_2 - n$. En effet, c'est l'image réciproque, par la submersion (f_1, f_2) de $E_1 \times E_2$ sur $M \times M$, de la diagonale Δ de $M \times M$ (qui, d'après 2.12.e, est une sous-variété de dimension n de $M \times M$). Pour tout $m \in M$, $f_1^{-1}(m) \times f_2^{-1}(m)$ est une sous-variété de $E_1 \times_M E_2$; le produit fibré $E_1 \times_M E_2$ n'est autre que la réunion de toutes ces sous-variétés, deux à deux disjointes, m parcourant M . Les deux composantes de l'application (f_1, f_2) sont égales sur $E_1 \times_M E_2$ de $E_1 \times E_2$, et déterminent une application $f : E_1 \times_M E_2 \rightarrow M$, qui est une submersion, et on a, pour tout $m \in M$, $f^{-1}(m) = f_1^{-1}(m) \times f_2^{-1}(m)$.

3. Les fibrés tangent et cotangent

3.1. Préliminaires. — La notion de *vecteur* en un point m de l'espace physique est sans doute familière au lecteur : on l'imagine comme un segment de droite orienté $\vec{v} = \overrightarrow{mn}$ ayant pour origine le point m et pour extrémité un autre point n de l'espace physique. Cette représentation des vecteurs est légitime dans l'espace physique habituel, qui a une structure d'espace affine. Rappelons en effet qu'un espace affine E , d'espace vectoriel associé \vec{E} , est un ensemble sur lequel l'espace vectoriel \vec{E} considéré comme groupe abélien (la loi de composition étant l'addition) opère de manière simplement transitive. On a donc une application de $\vec{E} \times E$ dans E , notée $(\vec{v}, m) \mapsto m + \vec{v}$, qui au couple formé d'un *vecteur* $\vec{v} \in \vec{E}$ et d'un *point* $m \in E$ associe un autre point $m + \vec{v}$, déduit de m en le translatant par \vec{v} . Cette application est telle que pour tout point $m \in E$ fixé, l'application $\vec{v} \mapsto m + \vec{v}$ est une bijection de \vec{E} sur E (revenant à identifier l'espace affine E à l'espace vectoriel \vec{E} qui lui est associé en choisissant une origine, le point m). C'est cette propriété qu'on exprime en disant que l'action de \vec{E} sur E est *simplement transitive*. Par

conséquent, si (m, n) est un couple ordonné de points de E , il existe un vecteur unique $\vec{v} \in \vec{E}$ tel que $m + \vec{v} = n$. On dit parfois qu'un élément \vec{v} de \vec{E} est un *vecteur libre*, et qu'un couple (m, \vec{v}) formé d'un point $m \in E$ et d'un vecteur $\vec{v} \in \vec{E}$ est un *vecteur lié* attaché au point m (ou d'origine m). Si $n = m + \vec{v}$, le vecteur lié (m, \vec{v}) est souvent noté $\overrightarrow{m\vec{n}}$.

Au lieu de l'espace physique (ou d'un espace affine) E , considérons maintenant une variété différentiable M . Il est très souhaitable de pouvoir définir la notion de vecteur attaché à un point m de la variété M (on dira *vecteur tangent* à M au point m). Cela permettra, par exemple, de considérer des équations différentielles sur la variété M . Pour définir cette notion, nous ne pouvons plus procéder comme ci-dessus, car nous ne disposons plus de la notion d'espace vectoriel associé à M , cette variété n'étant pas, en général, une partie d'un espace affine. En d'autres termes, nous ne disposons plus de la notion de vecteur libre. Nous allons voir qu'en procédant d'une autre manière, au fond très naturelle, nous pouvons définir directement la notion de *vecteur tangent* à la variété M en un de ses points m , qui correspond, très exactement, à la notion de vecteur lié d'origine m . Nous définirons aussi, par la même méthode, la notion duale de *covecteur* attaché à un point d'une variété. Cette méthode consiste à définir les vecteurs et covecteurs comme *classes d'équivalence* d'applications, pour des relations d'équivalence convenablement choisies. Un vecteur, attaché à un point m d'une variété M , apparaîtra comme la *vitesse* d'un point mobile sur M à l'instant où il passe en m ; un covecteur attaché au point m , comme le *gradient* en ce point d'une fonction à valeurs réelles.

3.2. Définition. — Soient M_1 et M_2 deux variétés différentiables de dimensions respectives n_1 et n_2 , a un point de M_1 et b un point de M_2 . Soient f et g deux applications différentiables définies sur des voisinages ouverts du point a dans M_1 (pas nécessairement le même voisinage pour les deux), à valeurs dans M_2 , appliquant toutes deux le point a sur le point b ($f(a) = g(a) = b$). On dit que f et g sont *tangentes en a à l'ordre 1* si, pour une carte admissible (U_1, φ_1) de M_1 et une carte admissible (U_2, φ_2) de M_2 telles que $a \in U_1$, $b \in U_2$, les applications $\varphi_2 \circ f \circ \varphi_1^{-1}$ et $\varphi_2 \circ g \circ \varphi_1^{-1}$ (définies sur un voisinage ouvert de $\varphi_1(a)$ dans \mathbb{R}^{n_1} et à valeurs dans \mathbb{R}^{n_2}) ont même différentielle au point $\varphi_1(a)$.

3.3. Commentaires

a) La propriété définie ci-dessus ne dépend pas du choix des cartes admissibles (U_1, φ_1) et (U_2, φ_2) . Remplaçons en effet celles-ci par d'autres cartes admissibles (U'_1, φ'_1) et (U'_2, φ'_2) . Nous pouvons écrire

$$\varphi'_2 \circ f \circ \varphi'_1{}^{-1} = (\varphi'_2 \circ \varphi_2^{-1}) \circ (\varphi_2 \circ f \circ \varphi_1^{-1}) \circ (\varphi_1 \circ \varphi_1^{-1}),$$

$$\varphi'_2 \circ g \circ \varphi'_1{}^{-1} = (\varphi'_2 \circ \varphi_2^{-1}) \circ (\varphi_2 \circ g \circ \varphi_1^{-1}) \circ (\varphi_1 \circ \varphi_1^{-1}).$$

En appliquant la formule de différentiation d'une application composée, on voit aisément que si les applications $\varphi_2 \circ f \circ \varphi_1^{-1}$ et $\varphi_2 \circ g \circ \varphi_1^{-1}$ ont même différentielle au point $\varphi_1(a)$, les applications $\varphi'_2 \circ f \circ \varphi'_1^{-1}$ et $\varphi'_2 \circ g \circ \varphi'_1^{-1}$ ont même différentielle au point $\varphi'_1(a)$.

b) On vérifie aisément que le fait d'être tangentes en a à l'ordre 1 est une *relation d'équivalence* sur l'ensemble des applications différentiables définies sur un voisinage ouvert de a dans M_1 et à valeurs dans M_2 qui appliquent a sur b .

c) On définit de même, pour tout entier $p \geq 1$, la notion d'applications tangentes en a à l'ordre p ; il suffit de remplacer les différentielles des applications $\varphi_2 \circ f \circ \varphi_1^{-1}$ et $\varphi_2 \circ g \circ \varphi_1^{-1}$ au point $\varphi_1(a)$ par leurs développements de Taylor à l'ordre p en ce point. Le fait d'être tangentes en a à l'ordre p est aussi une relation d'équivalence.

3.4. Définitions. — Soit M une variété différentiable de dimension n , et m un point de M .

1. On appelle *vecteur tangent* à M au point m une classe d'équivalence (pour la relation d'équivalence d'être tangentes en 0 à l'ordre 1) d'applications différentiables d'un intervalle ouvert de \mathbb{R} contenant l'origine 0 dans M qui appliquent 0 sur m .

2. On appelle *espace tangent* à la variété M au point m et on note $T_m M$ l'ensemble des vecteurs tangents en m à M .

3. On appelle *fibré tangent* à la variété M , et on note TM , la réunion $\bigcup_{m \in M} T_m M$ des espaces tangents à M en tous ses points, considérés comme deux à deux disjoints.

4. On appelle *covecteur* attaché au point m de la variété M une classe d'équivalence (pour la relation d'équivalence d'être tangentes en m à l'ordre 1) d'applications différentiables, définies sur un voisinage ouvert de m dans M et à valeurs dans \mathbb{R} , qui appliquent m sur l'origine 0.

5. On appelle *espace cotangent* à la variété M au point m et on note $T_m^* M$ l'ensemble des covecteurs attachés au point m de M .

6. On appelle *fibré cotangent* à la variété M , et on note $T^* M$, la réunion $\bigcup_{m \in M} T_m^* M$ des espaces cotangents à M en tous ses points, considérés comme deux à deux disjoints.

3.5. Commentaires

a) Soit χ une application différentiable d'un intervalle ouvert de \mathbb{R} contenant l'origine dans la variété M , appliquant 0 sur le point m . Le vecteur tangent à M au point m , classe d'équivalence de l'application χ , est souvent noté $\left. \frac{d\chi(s)}{ds} \right|_{s=0}$, ou aussi $\chi'(t)|_{t=0}$.

Plus généralement, soit $\varphi : I \rightarrow M$ une application différentiable d'un intervalle ouvert I de M dans \mathbb{R} . Pour tout $t \in I$, on note $\frac{d\varphi(t)}{dt}$, ou $\varphi'(t)$, le vecteur tangent à M au point $\varphi(t)$, classe d'équivalence de l'application $s \mapsto \varphi(t+s)$. En composant l'application φ avec la translation $s \mapsto t+s$ (t étant considéré comme fixé), nous obtenons une application qui applique l'origine 0 de \mathbb{R} sur le point $\varphi(t)$, et cela nous a permis d'utiliser la définition 3.4 d'un vecteur tangent.

De même, soit f une application différentiable d'un voisinage ouvert du point a de la variété M dans \mathbb{R} . Le covecteur en m , classe d'équivalence de l'application $x \mapsto f(x) - f(a)$, est noté $df(a)$. La justification de cette notation apparaîtra plus loin.

b) Nous allons voir comment sont déterminés un vecteur v tangent à la variété M au point m , et un covecteur η attaché au point m , lorsqu'on a choisi une carte (U, φ) de M , telle que $m \in U$. Notons (x^1, \dots, x^n) les coordonnées locales associées à cette carte, c'est-à-dire les n composantes de l'application φ .

Le vecteur v est une classe d'équivalence d'applications différentiables d'intervalles de \mathbb{R} contenant l'origine dans M qui appliquent 0 sur m . Soit f un *représentant* (c'est-à-dire un élément particulier) de cette classe. L'application $\varphi \circ f$ a n composantes, qui sont les applications $x^i \circ f$. Sa dérivée à l'origine a donc n composantes, les $v^i = \left. \frac{d(x^i(f(s)))}{ds} \right|_{s=0}$, $1 \leq i \leq n$. Le vecteur v est déterminé par ces n nombres réels. Ceux-ci ne dépendent pas du choix du représentant f de v , en raison de la définition même de la propriété, pour deux applications, d'être tangentes en 0 à l'ordre 1. On verra plus loin que l'espace $T_m M$ tangent en m à la variété M est un espace vectoriel de dimension n , et que le choix d'une carte (U, φ) de M , telle que $m \in U$, détermine une base de cet espace. Les n réels v^i sont les composantes du vecteur v dans cette base.

De même, le covecteur η est une classe d'équivalence d'applications différentiables, définies sur un voisinage ouvert de m dans M et à valeurs dans \mathbb{R} , qui appliquent m sur l'origine 0. Soit h un représentant de cette classe. L'application $h \circ \varphi^{-1}$ est une fonction, à valeurs réelles, de n variables réelles, les coordonnées usuelles sur \mathbb{R}^n . Celles-ci, par un léger abus de notations, seront désignées par x^1, \dots, x^n , comme les coordonnées locales attachées à la carte (U, φ) .

Remarquons au passage que pour éviter cet abus de notations, il nous faudrait utiliser des notations différentes pour les coordonnées locales sur U et pour les coordonnées usuelles sur \mathbb{R}^n . Nous pourrions par exemple noter φ^i les coordonnées locales sur U et x^i les coordonnées usuelles sur \mathbb{R}^n ($1 \leq i \leq n$). Nous aurions alors, bien entendu, $\varphi^i = x^i \circ \varphi$.

La différentielle de $h \circ \varphi^{-1}$ au point $\varphi(m)$ est la suite, comportant n termes,

de ses dérivées partielles au point $\varphi(m)$, $\eta_i = \frac{\partial((h \circ \varphi^{-1})(x^1, \dots, x^n))}{\partial x^i} \Big|_{\varphi(m)}$. Le covecteur η est déterminé par les n nombres réels η_i . Ceux-ci ne dépendent pas du choix du représentant h de η . On verra plus loin que l'espace T_m^*M cotangent en m à la variété M est un espace vectoriel de dimension n , et que le choix d'une carte (U, φ) de M , telle que $m \in U$, détermine une base de cet espace. Les n réels η_i sont les composantes du covecteur η dans cette base.

c) Voyons comment changent les n nombre v^i et les n nombre η_i qui représentent, respectivement, le vecteur v et le covecteur η dans la carte (U, φ) , lorsqu'on remplace celle-ci par une autre carte admissible (U', φ') telle que $m \in U'$. On note y^1, \dots, y^n les coordonnées locales associées.

L'application changement de carte $\varphi' \circ \varphi^{-1}$ est l'application $(x^1, \dots, x^n) \mapsto (y^1, \dots, y^n)$ qui exprime les coordonnées locales y^i associées à la carte (U', φ') en fonction des coordonnées locales x^j associées à la carte (U, φ) . Les composantes w^i du vecteur v dans la base de $T_m M$ associée à la carte (U', φ') ont pour expression

$$w^i = \frac{d(y^i(f(s)))}{ds} \Big|_{s=0} = \sum_{j=1}^n \frac{\partial y^i}{\partial x^j} \Big|_{\varphi(m)} \frac{d(x^j(f(s)))}{ds} \Big|_{s=0} = \sum_{j=1}^n \frac{\partial y^i}{\partial x^j} \Big|_{\varphi(m)} v^j.$$

Cette formule montre que les composantes w^i de v dans la base de $T_m M$ associée à la carte (U', φ') s'expriment linéairement au moyen des composantes v^i de ce vecteur dans la base associée à la carte (U, φ) . La matrice ayant pour coefficients les $\frac{\partial y^i}{\partial x^j} \Big|_{\varphi(m)}$ (i indexant les lignes et j les colonnes), qui apparaît dans cette formule, n'est autre que la matrice exprimant la différentielle, au point $\varphi(m)$, de l'application changement de carte $\varphi' \circ \varphi^{-1}$.

De même, l'application changement de carte $\varphi \circ \varphi'^{-1}$, inverse de $\varphi' \circ \varphi^{-1}$ est l'application $(y^1, \dots, y^n) \mapsto (x^1, \dots, x^n)$ qui exprime les coordonnées locales x^i associées à la carte (U, φ) en fonction des coordonnées locales y^j associées à la carte (U', φ') . Les composantes ζ_i du covecteur η dans la base de T_m^*M associée à la carte (U', φ') ont pour expression

$$\begin{aligned} \zeta_i &= \frac{\partial(h \circ \varphi'^{-1}(y^1, \dots, y^n))}{\partial y^i} = \frac{\partial((h \circ \varphi^{-1}) \circ (\varphi \circ \varphi'^{-1})(y^1, \dots, y^n))}{\partial y^i} \\ &= \sum_{j=1}^n \frac{\partial(h \circ \varphi^{-1}(x^1, \dots, x^n))}{\partial x^j} \frac{\partial x^j}{\partial y^i} \Big|_{\varphi'(m)} \\ &= \sum_{j=1}^n \eta_j \frac{\partial x^j}{\partial y^i} \Big|_{\varphi'(m)}. \end{aligned}$$

Cette formule montre que les composantes ζ_i de η dans la base de T_m^*M associée à la carte (U', φ') s'expriment linéairement au moyen des composantes η_j de ce

covecteur dans la base associée à la carte (U, φ) . La matrice ayant pour coefficients les $\frac{\partial x^j}{\partial y^i} \Big|_{\varphi'(m)}$ (i indexant les lignes et j les colonnes), qui apparaît dans cette formule, est la transposée de la matrice exprimant la différentielle, au point $\varphi'(m)$, de l'application changement de carte $\varphi \circ \varphi'^{-1}$. On sait que cette matrice est l'inverse de la matrice des $\frac{\partial y^i}{\partial x^j} \Big|_{\varphi(m)}$.

d) Il est maintenant facile de voir pourquoi l'espace $T_m M$ tangent en m à la variété M est un espace vectoriel de dimension n . En utilisant la carte (U, φ) , on a pu identifier cet espace à \mathbb{R}^n , puisqu'on a pu associer à tout vecteur $v \in T_m M$ la suite de ses composantes (v^1, \dots, v^n) dans la base de $T_m M$ associée à cette carte. On peut donc transporter sur $T_m M$ la structure d'espace vectoriel de \mathbb{R}^n . La structure ainsi obtenue ne dépend pas du choix de la carte (U, φ) , puisque lorsqu'on remplace cette carte par une autre carte admissible (U', φ') , les composantes (w^1, \dots, w^n) du vecteur v dans la base associée à cette nouvelle carte se déduisent des composantes (v^1, \dots, v^n) de ce vecteur dans la base associée à l'ancienne carte (U, φ) par une transformation linéaire inversible.

Le même raisonnement montre que l'espace $T_m^* M$, cotangent en m à la variété M , est aussi un espace vectoriel de dimension n . De plus, les formules exprimant les effets d'un changement de carte sur les composantes d'un vecteur et d'un covecteur, permettent de penser que les espaces $T_m M$ et $T_m^* M$ sont en dualité, chacun s'identifiant au dual de l'autre. Nous mettons ci-dessous en évidence ce couplage par dualité de manière directe.

e) En composant f , représentant du vecteur v , et h , représentant du covecteur η , on obtient une fonction différentiable $h \circ f$, définie sur un voisinage de l'origine 0 de \mathbb{R} , et à valeurs réelles. Posons

$$\langle \eta, v \rangle = \frac{d(h \circ f)}{ds} \Big|_{s=0}.$$

On note, comme ci-dessus, v^1, \dots, v^n et η_1, \dots, η_n les composantes, respectivement, du vecteur v et du covecteur η dans les bases de $T_m M$ et de $T_m^* M$ associées à la carte (U, φ) de M . Un calcul simple montre que

$$\langle \eta, v \rangle = \sum_{i=1}^n \eta_i v^i.$$

Par suite, $\langle \eta, v \rangle$ ne dépend pas du choix des représentants f de v et h de η , ni du choix de la carte (U, φ) (qui d'ailleurs n'intervient pas dans sa définition), mais seulement de v et de η , et définit un couplage entre les deux espaces $T_m M$ et $T_m^* M$, qui permet d'identifier chacun de ces espaces au dual de l'autre. On voit de plus que les bases de $T_m M$ et de $T_m^* M$ associées à une même carte (U, φ) de M

sont duales l'une de l'autre. On les note, respectivement, $(\frac{\partial}{\partial x^1}(m), \dots, \frac{\partial}{\partial x^n}(m))$ et $(dx^1(m), \dots, dx^n(m))$.

f) Les raisonnements qui précèdent montrent aussi que les fibrés TM et T^*M , respectivement tangent et cotangent à la variété différentiable M , sont des variétés différentiables de dimension $2n$ (double de la dimension de M). Traitons par exemple le cas de TM . À toute carte (U, φ) de M est naturellement associée une carte, notée $(TU, T\varphi)$, de TM dont le domaine est TU (ensemble des vecteurs tangents à M en un point appartenant à l'ouvert U). Dans cette carte, les coordonnées locales d'un élément v de TU sont les n coordonnées locales x^1, \dots, x^n du point m auquel v est attaché et les n composantes v^1, \dots, v^n de v dans la base de $T_m M$ associée à la carte (U, φ) .

On voit qu'il existe une application différentiable p de TM dans M et une application différentiable q de T^*M dans M , qui associent à chaque vecteur $v \in TM$ (resp., à chaque covecteur $\eta \in T^*M$) le point m de M auquel ce vecteur (resp., ce covecteur) est attaché. Pour tout point $m \in M$, $p^{-1}(m) = T_m M$ est l'espace tangent et $q^{-1}(m) = T_m^* M$ l'espace cotangent en m à M . Ce sont des espaces vectoriels de dimension n , dont chacun s'identifie au dual de l'autre. De plus, tout point $m \in M$ possède un voisinage U (le domaine d'une carte admissible contenant m) tel que $p^{-1}(U)$ et $q^{-1}(U)$ soient difféomorphes au produit $U \times \mathbb{R}^n$. Ces propriétés sont celles d'un *fibré vectoriel*. Nous définissons cette notion ci-dessous.

3.6. Définition. — Soit F un espace vectoriel de dimension finie. On appelle *espace fibré vectoriel* (ou, plus simplement, *fibré vectoriel*) *localement trivial de fibre-type* F un triplet (E, π, M) , où E et M sont des variétés différentiables et π une application différentiable de E sur M , vérifiant les propriétés suivantes.

1. Pour tout point $m \in M$, $\pi^{-1}(m)$ est muni d'une structure d'espace vectoriel de même dimension que F .
2. Pour tout point $m_0 \in M$, il existe un voisinage ouvert U de m_0 dans M et un difféomorphisme χ de $\pi^{-1}(U)$ sur $U \times F$, tels que, pour tout $m \in U$, la restriction de χ à $\pi^{-1}(m)$ soit un isomorphisme d'espaces vectoriels, de $\pi^{-1}(m)$ muni de sa structure d'espace vectoriel définie en 1, sur $\{m\} \times F$.

Les variétés E et M , l'application π et la dimension k de la fibre-type F sont appelées, respectivement, *espace total*, *base*, *projection canonique* et *rang* du fibré vectoriel (E, π, M) . Pour tout $m \in M$, $\pi^{-1}(m)$ est appelé *fibre* au point m (ou *fibre au dessus de* m). Un couple (U, χ) tel que celui intervenant dans 2 ci-dessus est appelé *trivialisations locale* de (E, π, M) au voisinage du point m_0 . Les trivialisations locales jouent, pour les fibrés vectoriels, un rôle semblable à celui que jouent les cartes pour les variétés.

Nous verrons dans le paragraphe suivant qu'on peut effectuer sur les fibrés vectoriels ayant pour base une variété donnée M les mêmes opérations que sur les espaces vectoriels. On peut notamment définir la *somme directe*, le *produit tensoriel* de deux ou plusieurs fibrés vectoriels de même base, ainsi que le *dual* d'un fibré vectoriel donné. Pour un exposé plus approfondi de ces notions, le lecteur pourra se reporter par exemple au livre de C. Godbillon, *Géométrie et Mécanique*. Nous nous bornerons ici à indiquer, sans démonstration, un critère permettant de reconnaître que deux fibrés vectoriels sont en dualité, afin de pouvoir l'appliquer aux fibrés tangent et cotangent.

3.7. Proposition. — Soient (E, π, M) et (F, ϖ, M) deux fibrés vectoriels de même base M . Soit $E \times_M F$ le produit fibré de E et de F au dessus de M , muni de la structure de sous-variété de $E \times F$ définie en 2.14.c. On suppose qu'il existe une fonction différentiable, notée $(x, y) \mapsto \langle x, y \rangle$, définie sur $E \times_M F$, à valeurs réelles, telle que, pour tout $m \in M$, la restriction de cette fonction à $E_m \times F_m$ soit un couplage par dualité entre ces deux espaces vectoriels, permettant d'identifier chacun d'eux au dual de l'autre (c'est-à-dire une application bilinéaire de $E_m \times F_m$ dans \mathbb{R} telle que, pour tout $x \in E_m$ non nul, il existe $y \in F_m$ tel que $\langle x, y \rangle \neq 0$, et que pour tout $y \in F_m$ non nul, il existe $x \in E_m$ tel que $\langle x, y \rangle \neq 0$). On a posé $E_m = \pi^{-1}(m)$, $F_m = \varpi^{-1}(m)$. On dit alors que l'application $(x, y) \mapsto \langle x, y \rangle$ est un couplage par dualité entre les fibrés vectoriels (E, π, M) et (F, ϖ, M) , et chacun de ces deux fibrés s'identifie au dual de l'autre.

Nous pouvons donc énoncer :

3.8. Théorème. — Soit M une variété différentiable de dimension n .

1. Pour tout point $m \in M$, les espaces $T_m M$ et $T_m^* M$, respectivement tangent et cotangent à M en m , sont des espaces vectoriels de dimension n en dualité (chacun d'eux s'identifie au dual de l'autre).

2. Le fibré tangent TM est l'espace total d'un fibré vectoriel localement trivial de fibre-type \mathbb{R}^n , de base M et de projection canonique p (application associant à chaque vecteur le point de M auquel il est attaché). Le fibré cotangent T^*M est l'espace total d'un fibré vectoriel, dual du précédent, dont la projection canonique est notée q (application associant à chaque covecteur le point de M auquel il est attaché).

3. La donnée d'une carte admissible (U, φ) de M détermine automatiquement une carte admissible $(TU, T\varphi)$ et une carte admissible $(T^*U, \hat{\varphi})$, respectivement de TM et de T^*M , dont les domaines sont l'ensemble TU des vecteurs tangents à M en un point de U et l'ensemble T^*U des covecteurs attachés à un point de U .

3.9. Remarque. — Par abus de langage, on désigne souvent un fibré vectoriel par son espace total. C'est pourquoi on dit que TM est le *fibré tangent* et T^*M le fibré cotangent à la variété M , alors qu'en toute rigueur ces termes désignent les triplets (TM, p, M) et (T^*M, q, M) .

4. Prolongement d'une application différentiable aux vecteurs

4.1. Préliminaires. — Soit f une application différentiable d'un ouvert U de \mathbb{R}^n dans \mathbb{R}^p (n et p étant des entiers ≥ 1). On notera f^i ($1 \leq i \leq p$ les composantes de l'application f ; chacune d'elles est une fonction de n variables réelles x^1, \dots, x^n , les coordonnées usuelles sur \mathbb{R}^n . Le lecteur est sans doute familiarisé avec la notion de *différentielle* de l'application f en un point $a = (a^1, \dots, a^n)$ de \mathbb{R}^n : c'est l'application linéaire $Df(a)$ de \mathbb{R}^n dans \mathbb{R}^p qui applique le vecteur $v = (v^1, \dots, v^n)$ sur le vecteur $w = (w^1, \dots, w^p)$ défini par

$$w^i = \sum_{j=1}^n \frac{\partial f^i(x^1, \dots, x^n)}{\partial x^j} \Big|_{x=a} v^j.$$

La matrice de cette application linéaire, ayant pour coefficients les dérivées partielles $\frac{\partial f^i(x^1, \dots, x^n)}{\partial x^j} \Big|_{x=a}$ des composantes de f au point a (i indexant les lignes et j les colonnes), est appelée *matrice jacobienne* de f au point a .

La généralisation naturelle de cette notion dans le contexte des variétés différentiables, est la notion d'*application linéaire tangente* en un point à une application différentiable, que nous allons définir. La définition est illustrée par la figure 2.

Figure 2. Application linéaire tangente.

4.2. Définition. — Soit f une application différentiable d'une variété différentiable M dans une autre variété différentiable N .

1. Soit a un point de M . On appelle *application linéaire tangente* en a à l'application f , et on note $T_a f$, l'application linéaire de l'espace $T_a M$ tangent en a à M dans l'espace $T_{f(a)} N$ tangent en $f(a)$ à N , ainsi définie. Soit $v \in T_a M$. D'après 2.5.4.1, v est une classe d'équivalence d'applications différentiables d'un

intervalle ouvert de \mathbb{R} contenant l'origine dans M , appliquant l'origine sur a . Soit χ un représentant (c'est-à-dire un élément particulier) de cette classe d'équivalence. L'application composée $f \circ \chi$ est une application différentiable d'un intervalle de \mathbf{R} contenant l'origine dans N , qui applique l'origine sur $f(a)$. Sa classe d'équivalence (pour la relation d'équivalence définie en 2.5.2) est un vecteur tangent en $f(a)$ à N . Par définition, ce vecteur est $T_a f(v)$.

2. On appelle *prolongement aux vecteurs* de l'application f , et on note Tf , l'application de TM dans TN dont la restriction à chaque fibre $T_a M$ de TM ($a \in M$) est l'application linéaire tangente à f en a .

4.3. Commentaires

a) Pour rendre la première définition légitime, on doit vérifier que la classe d'équivalence de $f \circ \chi$ ne dépend que de f et de v , non du choix du représentant χ de v . Le lecteur n'aura aucune difficulté à le faire en s'inspirant de 3.5.e. Il vérifiera de même que l'application $T_a f$ est bien linéaire.

b) La nécessaire rigueur des énoncés, qui rend un peu lourde la définition de l'application linéaire tangente, ne doit pas masquer l'idée sous-jacente qui est très simple et naturelle. Un vecteur tangent à M au point a est la vitesse d'un mobile, qui parcourt une courbe paramétrée $s \mapsto \chi(s)$ sur la variété M , à l'instant $s = 0$ où il passe par le point a . On le note $v = \left. \frac{d\chi(s)}{ds} \right|_{s=0}$. On peut penser à l'application $f : M \rightarrow N$ comme à un dispositif formant une image, dans la variété N , de ce qui se passe dans M (comme une chaîne de télévision qui donne sur un écran une image de ce qui se passe dans le monde réel). L'application composée $f \circ \chi$ décrit le mouvement de l'image du mobile qui sert à définir le vecteur v . La vitesse à l'instant 0 de cette image, $\left. \frac{d(f \circ \chi(s))}{ds} \right|_{s=0}$, est par définition le vecteur $T_a f(v)$.

c) Soient (U, φ) et (V, ψ) des cartes admissibles, respectivement de M et de N . On note (x^1, \dots, x^m) et (y^1, \dots, y^n) les coordonnées locales dans ces cartes. L'expression $\tilde{f} = \psi \circ f \circ \varphi^{-1}$ de f dans ces cartes s'écrit

$$(x^1, \dots, x^m) \mapsto (y^1 = \tilde{f}^1(x^1, \dots, x^m), \dots, y^n = \tilde{f}^n(x^1, \dots, x^m)).$$

Soient $(TU, T\varphi)$ et $(TV, T\psi)$ les cartes, respectivement, de TM et de TN , associées aux cartes (U, φ) et (V, ψ) , $(x^1, \dots, x^m, v^1, \dots, v^m)$ et $(y^1, \dots, y^n, w^1, \dots, w^n)$ les coordonnées locales correspondantes. Le lecteur vérifiera aisément que l'expression $\tilde{Tf} = T\psi \circ Tf \circ (T\varphi)^{-1}$ de Tf dans ces cartes est

$$(x^1, \dots, x^m, v^1, \dots, v^m) \mapsto (y^1, \dots, y^n, w^1, \dots, w^n),$$

avec

$$y^i = \tilde{f}^i(x^1, \dots, x^m), \quad w^i = \sum_{j=1}^m \frac{\partial \tilde{f}^i(x^1, \dots, x^m)}{\partial x^j} v^j.$$

Cette expression montre que Tf est une application différentiable (de classe C^∞ si f est de classe C^∞ , de classe C^{p-1} si f est de classe C^p).

On voit que pour tout $a \in U$, l'expression (dans les cartes considérées) de l'application linéaire $T_a f$, tangente à f au point a , s'identifie à la différentielle $D\tilde{f}(\varphi(a))$ de l'expression \tilde{f} de f au point $\varphi(a)$. Ce résultat confirme le fait, annoncé en 2.6.1, que la notion d'application linéaire tangente généralise celle de différentielle d'une application en un point.

4.4. Quelques propriétés

a) La notation $(Tu, T\varphi)$ pour désigner la carte de TM associée à une carte (U, φ) d'une variété M est maintenant justifiée : en remarquant que le fibré tangent à \mathbb{R}^m est $\mathbf{R}^m \times \mathbb{R}^m$, on voit que $T\varphi$ est bien le prolongement aux vecteurs de φ .

b) Soient M, N et P trois variétés différentiables, f une application différentiable de M dans N et g une application différentiable de N dans P . On a

$$T(g \circ f) = Tg \circ Tf.$$

Cette propriété généralise la formule classique de différentiation d'une application composée.

c) Soit f un difféomorphisme d'une variété différentiable M sur une autre variété différentiable N . Son prolongement aux vecteurs Tf est alors un isomorphisme de TM sur TN , au dessus du difféomorphisme f . Cela signifie que Tf est un difféomorphisme et que, pour tout point x de M , sa restriction à la fibre $T_x M$ est un isomorphisme de cette fibre sur la fibre correspondante $T_{f(x)} N$. De plus, on a

$$T(f^{-1}) = (Tf)^{-1}.$$

En particulier, si f est l'application identique de M , Tf est l'application identique de TM .

4.5. Remarque. — Plaçons-nous dans les hypothèses de la définition 4.2. L'application linéaire $T_a f : T_a M \rightarrow T_{f(a)} N$ a pour transposée l'application linéaire ${}^t(T_a f) : T_{f(a)}^* N \rightarrow T_a^* M$, du dual de $T_{f(a)} N$ dans le dual de $T_a M$, définie par

$$\langle {}^t(T_a f)(\eta), v \rangle = \langle \eta, T_a f(v) \rangle, \quad \eta \in T_{f(a)}^* N, v \in T_a M.$$

On pourrait songer à utiliser cette propriété pour définir une application ${}^t(Tf)$, qui irait de $T^* N$ dans $T^* M$. Mais ce n'est en général pas possible, car un point $b \in N$ étant donné, il peut n'exister aucun point $a \in M$ tel que $f(a) = b$, ou au contraire en exister plusieurs. On ne sait donc pas toujours dans quelle fibre de $T^* M$ doit s'appliquer l'espace $T_b^* N$ cotangent en b à N .

Toutefois, dans le cas particulier où f est un difféomorphisme, l'application ${}^t(Tf) : T^* N \rightarrow T^* M$ est parfaitement définie. On vérifie aisément que c'est un isomorphisme de fibrés vectoriels (au sens où cette notion a été définie en 2.6.4.3

à propos des fibrés tangents). L'isomorphisme inverse sera noté $\widehat{f} : T^*M \rightarrow T^*N$. Cela explique la notation $(T^*U, \widehat{\varphi})$ que nous avons utilisée pour désigner la carte de T^*M associée à une carte (U, φ) de M .

5. Les fibrés des tenseurs et des formes extérieures

5.1. Rappels d'algèbre linéaire. — Nous rappelons dans ce paragraphe quelques notions d'algèbre linéaire qui nous seront utiles plus loin. Le lecteur déjà familiarisé avec ces notions pourra se contenter de parcourir rapidement ce paragraphe pour se familiariser avec les notations.

Dans ce qui suit E, F, E_i, F_j (i et j indices quelconques) désignent des espaces vectoriels réels de dimension finie. On traiterait de même le cas d'espaces vectoriels complexes. On peut, de manière naturelle, associer à ces espaces divers autres espaces vectoriels que nous allons définir.

a) Dual d'un espace vectoriel. — On appelle *dual* de l'espace vectoriel E , et on note E^* , l'ensemble des formes linéaires sur E , c'est-à-dire des applications linéaires de E dans \mathbb{R} . C'est un espace vectoriel de même dimension que E .

Soit $x \in E, \alpha \in E^*$. Afin de souligner la similitude des rôles des deux arguments, on écrit souvent $\langle \alpha, x \rangle$ au lieu de $\alpha(x)$ pour désigner la valeur de la forme α au point x .

La donnée d'une base (e_1, \dots, e_n) de E détermine automatiquement une base $(\epsilon^1, \dots, \epsilon^n)$ de E^* , dite *duale* de (e_1, \dots, e_n) , dont les éléments ϵ^i sont déterminés par

$$\langle \epsilon^i, e_j \rangle = \delta_j^i = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{si } i \neq j. \end{cases}$$

Le dual de E^* s'identifie, de manière canonique, à l'espace E .

b) Produit tensoriel de deux espaces. — Nous ne donnerons pas ici la définition abstraite du produit tensoriel utilisable dans les cas les plus généraux; le lecteur la trouvera par exemple dans le livre de L. Schwartz sur les tenseurs. Nous nous contenterons d'une définition relativement concrète qui lui est équivalente dans le cas, qui est celui que nous considérons, d'espaces vectoriels réels de dimension finie.

On appelle *produit tensoriel* des deux espaces vectoriels E et F , et on note $E \otimes F$, l'ensemble des formes bilinéaires sur $E^* \times F^*$, c'est-à-dire l'ensemble des applications $t : (\alpha, \beta) \mapsto t(\alpha, \beta)$ du produit $E^* \times F^*$ dans \mathbb{R} qui sont linéaires en chacun des arguments $\alpha \in E^*$ et $\beta \in F^*$. Si $\dim E = n, \dim F = m, \dim E \otimes F = nm$.

Il existe une application bilinéaire naturelle de $E \times F$ dans $E \otimes F$, notée $(x, y) \mapsto$

$x \otimes y$, définie par

$$(x \otimes y)(\alpha, \beta) = \langle \alpha, x \rangle \langle \beta, y \rangle.$$

Soient (e_1, \dots, e_n) une base de E et (f_1, \dots, f_m) une base de F . Les $e_i \otimes f_j$, $1 \leq i \leq n$, $1 \leq j \leq m$, forment une base de $E \otimes F$, dite *associée* aux bases considérées de E et de F .

c) *Produit tensoriel de plusieurs espaces.* — On définit de même le *produit tensoriel* de p espaces vectoriels réels de dimension finie E_1, \dots, E_p : c'est l'ensemble des formes p -multilinéaires sur $E_1^* \times \dots \times E_p^*$. On le note $E_1 \otimes \dots \otimes E_p$, ou encore $\otimes_{i=1}^p E_i$.

Il existe une application p -multilinéaire de $E_1 \times \dots \times E_p$ dans $E_1 \otimes \dots \otimes E_p$, notée $(x_1, \dots, x_p) \mapsto x_1 \otimes \dots \otimes x_p$, définie comme dans le cas du produit tensoriel de deux espaces.

La donnée d'une base de chacun des espaces facteurs E_i ($1 \leq i \leq p$) détermine une base de leur produit tensoriel, définie comme dans le cas du produit tensoriel de deux espaces.

d) *Cas d'un espace et de son dual.* — Nous allons appliquer les définitions qui précèdent au cas où tous les espaces vectoriels E_i sont identiques à l'un d'entre eux E ou à son dual E^* . Un élément de $E \otimes \dots \otimes E$ (p facteurs) est appelé *tenseur p fois contravariant* sur E , et un élément de $E^* \otimes \dots \otimes E^*$ (q facteurs), *tenseur q fois covariant* sur E . On appelle *tenseur de type (p, q)* , ou *tenseur p fois contravariant et q fois covariant*, un élément de $E \otimes \dots \otimes E \otimes E^* \otimes \dots \otimes E^*$ (avec p facteurs E et q facteurs E^*).

En imposant certaines propriétés de symétrie du tenseur vis-à-vis des permutations des arguments dont il est fonction, on peut définir certains sous-espaces des espaces de tenseurs. Ainsi par exemple, un tenseur deux fois contravariant Λ sur l'espace vectoriel E est dit

- *symétrique* si, pour tous α et $\beta \in E^*$, $\Lambda(\beta, \alpha) = \Lambda(\alpha, \beta)$,
- *antisymétrique* si, pour tous α et $\beta \in E^*$, $\Lambda(\beta, \alpha) = -\Lambda(\alpha, \beta)$.

Nous donnons ci-dessous quelques rappels sur les tenseurs entièrement antisymétriques, qui jouent un rôle particulièrement important en géométrie différentielle.

e) *Formes extérieures.* — On appelle *forme extérieure de degré p* ou, en abrégé, *p -forme extérieure* sur un espace vectoriel E , un tenseur p fois covariant entièrement antisymétrique sur cet espace, c'est-à-dire une forme p -multilinéaire η sur $E^p = E \times \dots \times E$ (p facteurs) entièrement antisymétrique, c'est-à-dire vérifiant, pour tous $(x_1, \dots, x_p) \in E^p$ et toute permutation σ de $\{1, \dots, p\}$,

$$\eta(x_{\sigma(1)}, \dots, x_{\sigma(p)}) = \epsilon(\sigma)\eta(x_1, \dots, x_p),$$

où $\epsilon(\sigma)$, appelé *signature* de la permutation σ , vaut 1 si σ est paire (c'est-à-dire s'obtient en composant un nombre pair de transpositions) et -1 si σ est impaire. Soit n la dimension de E . Pour tout entier p vérifiant $1 \leq p \leq n$, l'ensemble des p -formes extérieures sur E est un sous-espace vectoriel de l'espace $E^* \otimes \cdots \otimes E^*$ (p facteurs) des tenseurs p fois covariants sur E , de dimension $\frac{n!}{p!(n-p)!}$, noté $\bigwedge^p E^*$.

Pour $p > n$, toute p -forme extérieure sur E est identiquement nulle; on a donc $\bigwedge^p E^* = \{0\}$. Pour $p < 0$ on pose aussi, par convention, $\bigwedge^p E^* = \{0\}$. Par convention aussi, on pose souvent $\bigwedge^0 E^* = \mathbb{R}$. L'espace $\bigwedge^p E^*$ est ainsi bien défini pour tout entier $p \in \mathbb{Z}$; mais il est non nul seulement pour $0 \leq p \leq n$.

Il existe une application p -multilinéaire entièrement antisymétrique naturelle de $(E^*)^p$ dans $\bigwedge^p E^*$, notée $(\alpha_1, \dots, \alpha_p) \mapsto \alpha_1 \wedge \dots \wedge \alpha_p$, définie par

$$(\alpha_1 \wedge \dots \wedge \alpha_p)(v_1, \dots, v_p) = \sum_{\sigma} \epsilon(\sigma) \alpha_1(x_{\sigma(1)}) \cdots \alpha_p(x_{\sigma(p)}),$$

où x_1, \dots, x_p sont des éléments de E . La somme du membre de droite porte sur l'ensemble des permutations σ de $\{1, \dots, p\}$, et $\epsilon(\sigma)$ désigne la signature de la permutation σ . On remarque que

$$\alpha_1 \wedge \cdots \wedge \alpha_p = \sum_{\sigma} \epsilon(\sigma) \alpha_{\sigma(1)} \otimes \cdots \otimes \alpha_{\sigma(p)}.$$

La notation $\alpha_1 \wedge \cdots \wedge \alpha_p$ est justifiée par la notion de produit extérieur rappelée ci-dessous.

f) Bases des espaces de formes extérieures. — Soit (e_1, \dots, e_n) une base de E , $(\epsilon^1, \dots, \epsilon^n)$ la base duale de E^* . Pour tout entier p vérifiant $1 \leq p \leq n$, on obtient une base de $\bigwedge^p E^*$ en prenant les $\epsilon^{i_1} \wedge \cdots \wedge \epsilon^{i_p}$, où (i_1, \dots, i_p) parcourt l'ensemble des multiuplets de p entiers vérifiant $1 \leq i_1 < i_2 < \cdots < i_p \leq n$. Cette base est dite *associée* à la base considérée de E .

g) Produit extérieur de deux formes. — Pour tout couple d'entiers p et q vérifiant $0 \leq p \leq n$, $0 \leq q \leq n$, il existe une application bilinéaire de $\bigwedge^p E^* \times \bigwedge^q E^*$ dans $\bigwedge^{p+q} E^*$, appelée *produit extérieur*, notée $(\alpha, \beta) \mapsto \alpha \wedge \beta$, ainsi définie :

— Si $p = 0$ (resp., si $q = 0$), α (resp., β) est un réel, et $\alpha \wedge \beta$ est le produit usuel de β (resp., de α) par ce nombre réel.

— Si p et q sont tous deux ≥ 1 , on définit $\alpha \wedge \beta$ en posant, pour tous $(x_1, \dots, x_{p+q}) \in E^{p+q}$,

$$(\alpha \wedge \beta)(x_1, \dots, x_{p+q}) = \sum_{\sigma} \epsilon(\sigma) \alpha(x_{\sigma(1)}, \dots, x_{\sigma(p)}) \beta(x_{\sigma(p+1)}, \dots, x_{\sigma(p+q)}),$$

la somme figurant au membre de gauche portant sur l'ensemble des permutations σ de $\{1, \dots, p+q\}$ qui vérifient

$$\sigma(1) < \sigma(2) < \cdots < \sigma(p) \quad \text{et} \quad \sigma(p+1) < \sigma(p+2) < \cdots < \sigma(p+q).$$

Ces permutations sont celles qu'effectue un joueur lorsqu'il bat un jeu de $p + q$ cartes en le séparant en deux paquets, formés respectivement par les p premières et les q dernières cartes, puis en entremêlant les deux paquets. En effet, l'ordre relatif des cartes appartenant à chacun des deux paquets est maintenu dans une telle opération. Comme précédemment, $\epsilon(\sigma)$ désigne la signature de la permutation σ .

Le produit extérieur possède les propriétés suivantes :

— il est associatif,

$$\alpha \wedge (\beta \wedge \gamma) = (\alpha \wedge \beta) \wedge \gamma,$$

— il vérifie, si $\alpha \in \bigwedge^p E^*$ et $\beta \in \bigwedge^q E^*$,

$$\beta \wedge \alpha = (-1)^{pq} \alpha \wedge \beta.$$

On appelle *algèbre extérieure de E* , et on note $\bigwedge E^*$, la somme directe $\bigoplus_{p \in \mathbb{Z}} \bigwedge^p E^*$ de tous les $\bigwedge^p E^*$. Cette dénomination est justifiée par le fait que le produit extérieur fait de cet espace vectoriel une algèbre associative.

h) Multivecteurs. — En suivant les mêmes lignes que ci-dessus, mais en échangeant les rôles de E et de son dual E^* , on peut définir l'algèbre $\bigwedge E$ des tenseurs contravariants entièrement antisymétriques sur E . Celle-ci joue, en géométrie différentielle, un rôle dont l'importance a beaucoup augmenté ces dernières années.

5.2. Tenseurs de type (p, q) . — Soient M une variété différentiable, x un point de M , p et q deux entiers ≥ 0 . On appelle *espace des tenseurs de type (p, q) au point x* , et on note $(\otimes^p(T_x M)) \otimes (\otimes^q(T_x^* M))$, le produit tensoriel de p exemplaires de l'espace tangent $T_x M$ et de q exemplaires de l'espace cotangent $T_x^* M$. Un élément de cet espace, appelé *tenseur de type (p, q) au point x* , est une application $(p + q)$ -multilinéaire de $(T_x^* M)^p \times (T_x M)^q$ dans \mathbb{R} .

La réunion, pour tous les points x de M , des espaces des tenseurs de type (p, q) au point x est appelée *fibré des tenseurs de type (p, q) sur la variété M* , et notée $(\otimes^p(TM)) \otimes (\otimes^q(T^*M))$. On montre que c'est l'espace total d'un fibré vectoriel localement trivial de base M . On sait qu'à toute carte (U, φ) de M sont associées une carte $(TU, T\varphi)$ de TM et une carte $(T^*U, \widehat{\varphi})$ de T^*M . Il est facile d'en déduire une carte de $(\otimes^p(TM)) \otimes (\otimes^q(T^*M))$, dite *associée* à la carte considérée de M . En effet, les cartes de TM et de T^*M s'obtiennent en remarquant que la carte (U, φ) détermine, pour tout point x de U , une base de $T_x M$ et une base de $T_x^* M$, duale de la précédente. Ainsi que nous l'avons vu au paragraphe 5.1.c, les produits tensoriels de p éléments de cette base de $T_x M$ et de q éléments de cette base de $T_x^* M$ forment une base de l'espace des tenseurs de type (p, q) au point x . On en déduit une trivialisatation locale du fibré $(\otimes^p(TM)) \otimes (\otimes^q(T^*M))$ au dessus de U , donc une carte de son espace total.

5.3. Formes extérieures de degré p . — Comme ci-dessus, M est une variété différentiable, et x un point de M . Pour tout entier $p \geq 1$, on appelle *espace des formes extérieures de degré p au point x* , et on note $\bigwedge^p(T_x^*M)$, l'espace des formes extérieures de degré p sur l'espace tangent T_xM . On sait que si $p > \dim M$, $\bigwedge^p(T_x^*M) = \{0\}$. Pour $p < 0$ on pose, par convention, $\bigwedge^p(T_x^*M) = \{0\}$ et, pour $p = 0$, $\bigwedge^0(T_x^*M) = \mathbb{R}$.

La réunion des espaces $\bigwedge^p(T_x^*M)$ pour tous les points x de M est appelée *fibré des formes extérieures de degré p sur M* , et notée $\bigwedge^p(T^*M)$. C'est l'espace total d'un fibré vectoriel localement trivial de base M . Comme ci-dessus, à toute carte (U, φ) de M est naturellement associée une trivialisatation locale de $\bigwedge^p(T^*M)$ au dessus de l'ouvert U , donc aussi une carte. Pour $p < 0$ et pour $p > \dim M$, $\bigwedge^p(T^*M)$ est le fibré nul de base M et, pour $p = 0$, $\bigwedge^0(T^*M)$ est le fibré trivial $M \times \mathbb{R}$.

Chapitre II
Calcul différentiel sur une variété

1. Champs de vecteurs, de tenseurs et formes différentielles

Le lecteur est sans doute déjà familiarisé avec la notion de champ de vecteurs ou de tenseurs sur un espace affine ou vectoriel. Ces notions sont d'un usage courant en physique et en mécanique. Ainsi par exemple, l'ensemble des vitesses des différents points d'un fluide en mouvement est naturellement décrit par un champ de vecteurs; le champ électromagnétique peut être décrit par un champ de tenseurs (ou par une forme différentielle de degré 2).

Ces notions s'adaptent sans difficulté au cas d'une variété différentiable : un champ de vecteurs (resp., de tenseurs de type (p, q)) sur une variété M est tout simplement une application qui associe à chaque point de M un vecteur (resp., un tenseur de type (p, q)) attaché à ce point.

1.1. Définitions. —

1. Soit (E, π, M) un fibré vectoriel localement trivial ayant pour base une variété différentiable M . On appelle *section* de ce fibré une application s de M dans E qui applique chaque point x de M dans la fibre $E_x = \pi^{-1}(x)$ au dessus de ce point, c'est-à-dire qui vérifie $\pi \circ s = \text{id}_M$.

2. En particulier, une section du fibré des tenseurs de type (p, q) sur une variété différentiable M est appelée *champ de tenseurs* (ou souvent, plus simplement, *tenseur*) de type (p, q) sur M . Pour $p = 1, q = 0$, une section du fibré tangent TM est appelée *champ de vecteurs* sur la variété M . Pour $p = 0, q = 1$, une section du fibré cotangent T^*M est appelée *forme différentielle de degré 1*, ou aussi *forme de Pfaff* sur la variété M . Une section du fibré des formes extérieures de degré p sur M est appelée *forme différentielle de degré p* sur cette variété.

Par convention, les sections de fibrés vectoriels (donc les champs de vecteurs ou de tenseurs et les formes différentielles) seront toujours supposées différentiables de classe C^∞ .

1.2. Quelques exemples

a) *Forme élément de volume*. — Soit M une variété différentiable de dimension n . On appelle *forme élément de volume* sur M une forme différentielle de degré n

ne s'annulant en aucun point de M . L'existence d'une forme élément de volume sur une variété n'est pas toujours assurée : une telle forme existe si et seulement si la variété est orientable, c'est-à-dire si et seulement s'il existe un atlas dont les applications changement de carte ont, en tout point, une matrice jacobienne à déterminant positif.

Les formes élément de volume interviennent en mécanique des milieux continus pour l'expression d'équations exprimant la conservation de la masse, ou, dans le cas de fluides incompressibles, du volume.

b) Métriques riemanniennes ou pseudo-riemanniennes. — Une *métrique riemannienne* (resp., *pseudo-riemannienne*) sur une variété différentiable M est un champ de tenseurs g deux fois covariants symétriques définis positifs (resp., non dégénérés), c'est-à-dire tel que, pour tout $x \in M$, $g(x)$ soit une forme bilinéaire symétrique définie positive (resp., non dégénérée) sur l'espace tangent $T_x M$. Rappelons que $g(x)$ est dite *définie positive* si pour tout $v \in T_x M$, $v \neq 0$, on a $g(x)(v, v) > 0$; elle est dite *non dégénérée* si pour tout $v \in T_x M$, $v \neq 0$, il existe $w \in T_x M$ tel que $g(x)(v, w) \neq 0$. Une forme bilinéaire symétrique définie positive est évidemment non dégénérée.

La donnée d'une métrique riemannienne (resp., pseudo-riemannienne) g sur une variété différentiable M détermine, pour tout point x de M , un *produit scalaire euclidien* (resp., *pseudo-euclidien*) $(v, w) \mapsto g(x)(v, w)$ sur $T_x M$. Si la métrique est riemannienne, on peut aussi définir une *norme* sur $T_x M$ en posant $\|v\| = \sqrt{g(x)(v, v)}$.

Les métriques pseudo-riemanniennes sont utilisées en physique notamment pour décrire la gravitation (théorie de la Relativité générale).

c) Différentielle d'une fonction. — Soit f une fonction différentiable sur la variété M . Pour tout point $a \in M$, on note $df(a)$ le covecteur, élément de $T_a^* M$, classe d'équivalence (pour la relation d'être tangentes en a à l'ordre 1) de l'application de M dans \mathbb{R} appliquant a sur $0 : x \mapsto f(x) - f(a)$. On obtient ainsi une forme différentielle de degré 1 sur M , notée df , appelée *différentielle* de la fonction f .

1.3. Expressions en coordonnées locales. — Soient M une variété différentiable de dimension n , (U, φ) une carte admissible de cette variété, et x^1, \dots, x^n les coordonnées locales associées à cette carte.

a) Expression des formes différentielles. — Les différentielles des fonctions coordonnées locales dx^1, \dots, dx^n (au sens de 1.2.c) sont n formes différentielles de degré 1 définies sur l'ouvert U de M . Leurs valeurs en chaque point x de U forment une base de l'espace cotangent $T_x^* U$; c'est cette base que nous avons utilisée pour définir la trivialisations locale et la carte $(T^* U, \hat{\varphi})$ de $T^* M$ associées à la carte (U, φ) de M . Toute forme différentielle η sur M , de degré 1, restreinte à U , peut

d'une manière unique s'écrire sous la forme

$$\eta = \sum_{i=1}^n \eta_i dx^i,$$

où les η_i sont des fonctions différentiables définies sur U , appelées *composantes* de la forme η dans la carte (U, φ) .

Plus généralement, pour tout entier p vérifiant $1 \leq p \leq n$, les valeurs en chaque point $x \in U$ des formes différentielles $dx^{i_1} \wedge \cdots \wedge dx^{i_p}$, lorsque (i_1, \dots, i_p) parcourt l'ensemble des multiuplets de p entiers vérifiant $1 \leq i_1 < i_2 < \cdots < i_p \leq n$, constituent une base de $\bigwedge^p(T_x^*M)$. Par suite, toute forme différentielle ζ de degré p sur la variété M s'exprime, en restriction à l'ouvert U , de manière unique sous la forme

$$\zeta = \sum_{1 \leq i_1 < \cdots < i_p \leq n} \zeta_{i_1 \dots i_p} dx^{i_1} \wedge \cdots \wedge dx^{i_p}.$$

Les fonctions différentiables $\zeta_{i_1 \dots i_p}$, définies sur l'ouvert U , sont appelées *composantes strictes* de la forme différentielle η associées à la carte (U, φ) .

b) Expression des champs de vecteurs. — On note $\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n}$ les champs de vecteurs, définis sur l'ouvert U , dont les valeurs en chaque point x forment la base de $T_x M$ duale de la base $(dx^1(x), \dots, dx^n(x))$ de T_x^*M . C'est cette base que nous avons utilisée pour associer à la carte (U, φ) de M une trivialisatation locale et une carte $(Tu, T\varphi)$ de TM . Tout champ de vecteurs X sur M s'exprime sur l'ouvert U de manière unique sous la forme

$$X = \sum_{i=1}^n X^i \frac{\partial}{\partial x^i}.$$

Les fonctions différentiables X^i , définies sur U , sont appelées *composantes* du champ de vecteurs X dans la carte (U, φ) .

c) Expression de la différentielle d'une fonction. — Soit f une fonction différentiable sur la variété M . Son expression locale $\tilde{f} = f \circ \varphi^{-1}$ dans la carte (U, φ) est une fonction des n coordonnées usuelles sur \mathbb{R}^n , que nous notons x^1, \dots, x^n (par le même léger abus de notations qu'au paragraphe I.2.3.c). La différentielle de la fonction f a dans U pour expression

$$df = \sum_{i=1}^n \left(\frac{\partial \tilde{f}}{\partial x^i} \circ \varphi \right) dx^i.$$

En pratique, on identifie souvent une fonction sur l'ouvert U avec sa composée avec φ^{-1} . C'est ainsi qu'on a identifié les coordonnées locales x^i avec les coordonnées usuelles dans \mathbb{R}^n . De même, on identifie f et son expression locale $\tilde{f} = f \circ \varphi^{-1}$. On identifie les fonctions $\frac{\partial \tilde{f}}{\partial x^i}$ (définies sur l'ouvert $\varphi(U)$ de \mathbb{R}^n) avec leurs composées

avec φ . C'est pourquoi on écrit

$$df = \sum_{i=1}^n \frac{\partial f}{\partial x^i} dx^i.$$

2. Opérations ponctuelles

On peut effectuer diverses opérations sur les champs de vecteurs ou de tenseurs et les formes différentielles. Nous allons décrire les plus fréquemment employées, en commençant par les plus simples, les opérations *ponctuelles* (qui se font en chaque point de la variété considérée, indépendamment des points voisins).

2.1. Définitions. — Soit M une variété différentiable, X un champ de vecteurs, α et β des formes différentielles de degrés p et q , respectivement, sur cette variété.

1. On appelle *produit extérieur* des formes différentielles α et β , et on note $\alpha \wedge \beta$, la forme différentielle de degré $p + q$ dont la valeur, en chaque point x de M , est $\alpha(x) \wedge \beta(x)$ (définie paragraphe I.5.1.g).

2. On appelle *produit intérieur* de la forme différentielle α par le champ de vecteurs X , et on note $i(X)\alpha$, la forme différentielle de degré $p - 1$ ainsi définie : pour tout point x de M et tout multipléte (v_1, \dots, v_{p-1}) de $p - 1$ éléments de $T_x M$, on a

$$i(X)\alpha(x)(v_1, \dots, v_{p-1}) = \alpha(x)(X(x), v_1, \dots, v_{p-1}).$$

Par convention, si $p = 0$, c'est-à-dire si α est une fonction différentiable sur M à valeurs réelles, on a $i(X)\alpha = 0$.

2.2. Remarque. — On définirait de même le produit tensoriel de deux champs de tenseurs, ou encore les produits contractés de divers types de champs de tenseurs et de vecteurs sur la variété M .

2.3. Quelques propriétés. — Soit M une variété différentiable de dimension n . Pour tout entier p , on note $A^p(M)$ l'espace des formes différentielles de degré p sur M . Pour $p < 0$ et pour $p > n$, on a $A^p(M) = \{0\}$. Pour $p = 0$, $A^0(M)$ est, par convention, l'espace des fonctions différentiables réelles sur M . On note $A(M) = \bigoplus_p A^p(M)$ la somme directe de tous les $A^p(M)$.

a) Le produit extérieur fait de $A(M)$ une algèbre associative graduée, appelée *algèbre des formes différentielles*, ou *algèbre extérieure* sur M .

b) Soit X un champ de vecteurs sur M . Le produit intérieur par X est une *dérivation* de degré -1 de l'algèbre $A(M)$. Cela signifie que c'est une application linéaire de $A(M)$ dans elle-même, qui applique chaque $A^p(M)$ dans $A^{p-1}(M)$, et qui vérifie, pour tous $\alpha \in A^p(M)$ et $\beta \in A^q(M)$,

$$i(X)(\alpha \wedge \beta) = (i(X)\alpha) \wedge \beta + (-1)^p \alpha \wedge (i(X)\beta).$$

2.4. Définition. — Soit f une application différentiable d'une variété différentiable M dans une autre variété différentiable N , et η une forme différentielle de degré p sur M . On appelle *image réciproque de η par f* , et on note $f^*\eta$, la forme différentielle de degré p sur M définie par

$$f^*\eta(x)(v_1, \dots, v_p) = \eta(f(x))(T_x f(v_1), \dots, T_x f(v_p)),$$

où x est un point de M , v_1, \dots, v_p des vecteurs tangents en x à M et $T_x f$ l'application linéaire tangente en x à f .

2.5. Quelques propriétés

a) L'application $\eta \mapsto f^*\eta$ est un homomorphisme de l'algèbre graduée $A(N)$ dans l'algèbre graduée $A(M)$. Cela signifie que c'est une application linéaire, qui associe à chaque forme différentielle sur N une forme différentielle de même degré sur M , et qui vérifie, pour tout couple (α, β) de formes différentielles sur N ,

$$f^*(\alpha \wedge \beta) = (f^*\alpha) \wedge (f^*\beta).$$

b) Une fonction différentiable h sur N , à valeurs réelles, peut être considérée comme une forme différentielle de degré 0. Son image réciproque f^*h n'est autre que la fonction composée $h \circ f$.

c) Soient M , N et P trois variétés différentiables, $f : M \rightarrow N$ et $g : N \rightarrow P$ deux applications différentiables. Pour toute forme différentielle η sur P , on a

$$f^*(g^*\eta) = (g \circ f)^*\eta.$$

d) Soient (U, φ) et (V, ψ) des cartes admissibles, respectivement de M et de N , x^1, \dots, x^m et y^1, \dots, y^n les coordonnées locales associées. Soit ζ une forme différentielle de degré p sur N ayant pour expression sur V

$$\zeta = \sum_{1 \leq i_1 < \dots < i_p \leq n} \zeta_{i_1 \dots i_p}(y^1, \dots, y^n) dy^{i_1} \wedge \dots \wedge dy^{i_p}.$$

L'expression $\psi \circ f \circ \varphi^{-1}$ de l'application f est l'application $(x^1, \dots, x^m) \mapsto (y^1, \dots, y^n)$. En utilisant cette application, on peut considérer les y^j ($1 \leq j \leq n$) comme des fonctions de (x^1, \dots, x^m) . L'expression de $f^*\zeta$ sur $U \cap f^{-1}(V)$ s'obtient en remplaçant, dans l'expression de ζ , les y^j et leurs différentielles dy^j par leurs expressions en fonction des x^i , c'est-à-dire par $y^i(x^1, \dots, x^m)$ et

$$d(y^j(x^1, \dots, x^m)) = \sum_{k=1}^m \frac{\partial y^j(x^1, \dots, x^m)}{\partial x^k} dx^k. \text{ On peut donc écrire}$$

$$f^*\zeta = \sum_{1 \leq i_1 < \dots < i_p \leq n} \zeta_{i_1 \dots i_p}((y^1(x^1, \dots, x^m), \dots, y^n(x^1, \dots, x^m))) \\ d(y^{i_1}(x^1, \dots, x^m)) \wedge \dots \wedge d(y^{i_p}(x^1, \dots, x^m)),$$

avec

$$d(y^{i_j}(x^1, \dots, x^m)) = \sum_{k=1}^m \frac{\partial y^{i_j}(x^1, \dots, x^m)}{\partial x^k} dx^k.$$

2.6. Remarque. — Dans les hypothèses de la définition 2.4, il n'est en général pas possible de définir l'image réciproque d'un champ de vecteurs Y sur la variété N , car l'application Tf va de TM dans TN . On pourrait plutôt envisager de définir l'image *directe* d'un champ de vecteurs X sur M . Mais cette image n'est en général pas un champ de vecteurs sur N , car un point y de N étant donné, il peut n'exister aucun point x de M tel que $f(x) = y$, ou en exister plusieurs. Ces difficultés n'apparaissent pas lorsque f est un difféomorphisme; les images directes ou réciproques de champs de vecteurs peuvent alors être définies. Lorsque f n'est pas un difféomorphisme, la notion adéquate est celle de *couple de champs de vecteurs compatibles*. Ces notions sont explicitées ci-dessous.

2.7. Définitions. — Soient f une application différentiable d'une variété M dans une variété N , X un champ de vecteurs sur M et Y un champ de vecteurs sur N .

1. On dit que le couple (X, Y) est compatible avec f si, pour tout point x de M ,

$$T_x f(X(x)) = Y(f(x)).$$

2. Lorsque f est un difféomorphisme, on appelle *image directe* du champ de vecteurs X , et on note $f_* X$, l'unique champ de vecteurs sur N tel que $(X, f_* X)$ soit compatible avec f , qui a pour expression

$$f_* X = Tf \circ X \circ f^{-1}.$$

De même, on appelle *image réciproque* du champ de vecteurs Y et on note $f^* Y$, l'unique champ de vecteurs sur M tel que $(f^* Y, Y)$ soit compatible avec f . Il n'est autre que l'image directe de Y par f^{-1} .

La propriété suivante est une conséquence immédiate des définitions.

2.8. Proposition. — Soient f une application différentiable d'une variété M dans une variété N , X un champ de vecteurs sur M , Y un champ de vecteurs sur N tels que (X, Y) soit compatible avec f , et η une forme différentielle sur N . On a

$$f^*(i(Y)\eta) = i(X)(f^*\eta).$$

3. Champs de vecteurs et équations différentielles

3.1. Rappel sur les équations différentielles. — Soit E un espace affine de dimension finie, d'espace vectoriel associé \vec{E} . On appelle *équation différentielle du premier ordre sous forme canonique* sur E , une relation de la forme

$$\frac{d\varphi(t)}{dt} = X(t, \varphi(t)), \quad (1)$$

où X est une application d'un ouvert V de $\mathbb{R} \times E$ dans \vec{E} . On appelle *solution* de cette équation différentielle toute application différentiable φ , définie sur un intervalle ouvert I de \mathbb{R} et à valeurs dans E , telle que pour tout $t \in I$, le couple $(t, \varphi(t))$ appartienne à l'ouvert V sur lequel f est définie, et que la dérivée $\frac{d\varphi(t)}{dt}$ de φ au point t soit égale à $X(t, \varphi(t))$.

Les modèles mathématiques de systèmes physiques utilisent très souvent une équation différentielle du type de l'équation (1) pour régir l'évolution de l'état du système au cours du temps. L'espace affine E est l'ensemble des états possibles du système, l'ensemble des réels \mathbb{R} représente l'ensemble des temps. L'application X est l'expression mathématique des lois qui régissent l'évolution du système : si à un instant $t \in \mathbb{R}$, le système est dans l'état représenté par le point $x \in E$, le couple (t, x) appartenant à l'ouvert V de $\mathbf{R} \times E$ sur lequel la loi d'évolution est connue, alors l'évolution de l'état du système au cours du temps doit être telle qu'à l'instant t , la vitesse du point représentant l'état du système soit $X(t, x)$. Il est commode de garder présente à l'esprit cette interprétation physique d'une équation différentielle, qui conduit à penser à la variable $t \in \mathbb{R}$ intervenant dans l'équation comme représentant le temps.

La restriction d'une solution $\varphi : I \rightarrow E$ à un sous-intervalle ouvert de son intervalle de définition est encore une solution. On dit qu'une solution, définie sur un intervalle ouvert I de \mathbb{R} , est *maximale* si elle n'est pas la restriction à I d'une solution définie sur un intervalle ouvert contenant I strictement (c'est-à-dire contenant I sans lui être égal). On montre que toute solution de l'équation différentielle (1) est restriction d'une solution maximale; on peut donc ne rechercher que les solutions maximales, car une fois celles-ci trouvées, toutes les autres seront connues aussi.

Une *donnée de Cauchy* pour l'équation différentielle (1) est un couple $(t_0, x_0) \in V$. On dit qu'une solution $\varphi : I \rightarrow E$ de cette équation *satisfait* la donnée de Cauchy (t_0, x_0) si $t_0 \in I$ et si $\varphi(t_0) = x_0$.

Lorsque l'application X est assez régulière, par exemple lorsqu'elle est différentiable de classe C^∞ , on montre (théorème de Cauchy-Lipschitz) que pour toute donnée de Cauchy (t_0, x_0) , il existe une unique solution maximale de l'équation différentielle (1) satisfaisant cette donnée de Cauchy. Ce théorème est une des raisons de l'emploi d'équations différentielles dans des modèles mathématiques de systèmes physiques, car il exprime l'idée du déterminisme scientifique : si à l'instant t_0 , on connaît l'état x_0 du système étudié, alors l'état de ce système à tous les autres instants, aussi bien antérieurs que postérieurs à t_0 , peut en principe être trouvé en résolvant l'équation différentielle régissant l'état du système pour la donnée de Cauchy (t_0, x_0) .

On dit que l'équation différentielle (1) est *autonome* si l'application X ne dépend

pas du temps, c'est-à-dire si l'ouvert V de $\mathbb{R} \times E$ est de la forme $\mathbb{R} \times U$, où U est un ouvert de E , et si l'application $X : (t, x) \mapsto X(t, x)$ ne dépend en fait que de la composante $x \in U$, non de la composante $t \in \mathbb{R}$. L'application X est alors un *champ de vecteurs* sur l'ouvert U de E , et l'équation s'écrit

$$\frac{d\varphi(t)}{dt} = X(\varphi(t)). \quad (2)$$

Les équations autonomes se rencontrent dans la modélisation de systèmes isolés, dont l'évolution ne dépend que de l'état du système lui-même.

Les solutions d'équations autonomes ont une importante propriété, l'invariance par translation dans le temps : si $\varphi :]a, b[\rightarrow U$ est une solution (maximale) de l'équation autonome (2), alors pour tout $\theta \in \mathbb{R}$, l'application de $]a + \theta, b + \theta[$ dans $U : t \mapsto \varphi(t - \theta)$ est aussi une solution (maximale) de (2). C'est pourquoi on peut, pour ces équations, se contenter d'étudier les données de Cauchy de la forme $(0, x_0)$, avec $x_0 \in U$: si φ est la solution maximale satisfaisant cette donnée, la solution maximale satisfaisant la donnée de Cauchy (t_0, x_0) (avec $t_0 \in \mathbb{R}$ quelconque) est $t \mapsto \varphi(t - t_0)$.

3.2. Équations différentielles sur une variété

Tout ce qui a été rappelé dans le paragraphe précédent se généralise sans difficulté au cas où l'espace affine de dimension finie E est remplacé par une variété différentiable M . Une *équation différentielle* sur M est une relation de la forme

$$\frac{d\varphi(t)}{dt} = X(t, \varphi(t)), \quad (3)$$

où X est un *champ de vecteurs dépendant du temps* sur M , c'est-à-dire une application de $\mathbb{R} \times M$ dans TM qui, pour tout $t \in \mathbb{R}$ et tout $x \in M$, applique (t, x) dans l'espace $T_x M$ tangent en x à M . Pour rester plus près de ce qui a été exposé dans le paragraphe précédent, on pourrait supposer X définie sur un ouvert de $\mathbb{R} \times M$ plutôt que sur $\mathbb{R} \times M$ entier ; nous ne le ferons pas pour alléger l'exposé. Une *solution* (ou *courbe intégrale*) de cette équation différentielle est une application différentiable φ , définie sur un intervalle ouvert I de \mathbb{R} et à valeurs dans M , telle que pour tout $t \in I$, la dérivée $\frac{d\varphi(t)}{dt}$ de φ au point t soit égale à $X(t, \varphi(t))$. On rappelle que le vecteur $\frac{d\varphi(t)}{dt}$, tangent à M en $\varphi(t)$, est la classe d'équivalence de l'application $s \mapsto \varphi(t + s)$, pour la relation d'équivalence d'être tangentes en 0 à l'ordre 1. La solution φ est dite *maximale* si elle n'est pas la restriction à l'intervalle I d'une solution définie sur un intervalle ouvert contenant I strictement. Toute solution est restriction d'une solution maximale. Une *donnée de Cauchy* est un élément (t_0, x_0) de $\mathbb{R} \times M$, et la solution φ *satisfait* cette donnée de Cauchy si $t_0 \in I$ et $\varphi(t_0) = x_0$. Le théorème de Cauchy-Lipschitz est toujours

applicable : lorsque l'application X est assez régulière, par exemple lorsqu'elle est différentiable de classe C^∞ , pour toute donnée de Cauchy il existe une solution maximale unique satisfaisant cette donnée.

Une équation différentielle sur la variété M est dite *autonome* si elle est de la forme

$$\frac{d\varphi(t)}{dt} = X(\varphi(t)), \quad (4)$$

où X est un champ de vecteurs sur M . Cette équation est dite *associée* au champ de vecteurs X . Les solutions de cette équation sont dites *courbes intégrales du champ de vecteurs* X . Elles vérifient la propriété d'invariance par translation du temps exposée dans le paragraphe précédent.

3.3. Le flot d'une équation différentielle

On considère l'équation différentielle (3), pas nécessairement autonome, sur la variété M , et on suppose l'application X différentiable de classe C^∞ . On sait alors que pour toute donnée de Cauchy $(t_0, x_0) \in \mathbb{R} \times M$, il existe une solution maximale unique satisfaisant cette donnée; notons-la $\varphi_{(t_0, x_0)}$, et notons $I_{(t_0, x_0)}$ son intervalle de définition. Le lecteur remarquera que cet intervalle contient t_0 . En considérant $\varphi_{(t_0, x_0)}(t)$ comme fonction des trois variables (t, t_0, x_0) , on peut regrouper toutes les solutions maximales de (3) en une application unique

$$\Phi : (t, t_0, x_0) \mapsto \Phi(t, t_0, x_0) = \varphi_{(t_0, x_0)}(t),$$

à valeurs dans M , appelée *flot* de l'équation différentielle (3). On montre que le domaine de définition de Φ est un ouvert D de $\mathbb{R} \times \mathbb{R} \times M$. L'application Φ est caractérisée par la propriété suivante : pour tous $(t_0, x_0) \in \mathbb{R} \times M$, l'ensemble des $t \in \mathbb{R}$ tels que (t, t_0, x_0) soit élément de D est l'intervalle $I_{(t_0, x_0)}$ sur lequel la solution maximale $\varphi_{(t_0, x_0)}$ satisfaisant la donnée de Cauchy (t_0, x_0) est définie, et $\Phi(t, t_0, x_0) = \varphi_{(t_0, x_0)}(t)$.

3.4. Propriétés du flot

Comme ci-dessus, on suppose le champ de vecteurs X dépendant du temps différentiable de classe C^∞ .

a) Différentiabilité. — On montre que le flot Φ de l'équation (3) est une application différentiable de classe C^∞ . L'ouvert D de $\mathbb{R} \times \mathbb{R} \times M$ sur lequel le flot Φ est défini contient $\Delta \times M$, où Δ désigne la diagonale de $\mathbb{R} \times \mathbb{R}$. On a, pour tous $t_0 \in \mathbb{R}$, $x_0 \in M$,

$$\Phi(t_0, t_0, x_0) = x_0.$$

Pour tous $(t_0, x_0) \in \mathbb{R} \times M$, l'ensemble des $t \in \mathbb{R}$ tels que $(t, t_0, x_0) \in D$ est un intervalle ouvert contenant t_0 . De plus, pour tout couple de réels $(t_0, t) \in \mathbb{R}^2$, l'ensemble des $x_0 \in M$ tels que (t, t_0, x_0) soit élément de D est un ouvert (éventuellement vide), noté $D_{(t_0, t)}$, de M , et l'application

$$x_0 \mapsto \Phi(t, t_0, x_0)$$

est un difféomorphisme de classe C^∞ de cet ouvert sur un autre ouvert de M , qui n'est autre que $D_{(t,t_0)}$. L'inverse de ce difféomorphisme n'est autre que

$$x \mapsto \Phi(t_0, t, x).$$

b) *Formule de composition.* — Le flot Φ vérifie la formule

$$\Phi(t_2, t_0, x_0) = \Phi(t_2, t_1, \Phi(t_1, t_0, x_0))$$

pour tous $t_0, t_1, t_2 \in \mathbb{R}$, $x_0 \in M$ tels que les deux membres de l'égalité soient définis. On montre même que si le membre de droite est défini, celui de gauche l'est aussi, et que si le membre de gauche est défini et que $\Phi(t_1, t_0, x_0)$ est défini, le membre de droite l'est aussi. En particulier, pour $t_2 = t_0$,

$$x_0 = \Phi(t_0, t_0, x_0) = \Phi(t_0, t_1, \Phi(t_1, t_0, x_0)).$$

Ces propriétés sont des conséquences immédiates de l'unicité de la courbe intégrale satisfaisant une donnée de Cauchy. Elles sont illustrées par la figure 3.

Figure 3. Formule de composition du flot.

3.5. Le flot réduit

On appelle *flot réduit* de l'équation différentielle (3) la restriction du flot Φ à l'ensemble des $(t, t_0, x_0) \in D$ tels que $t_0 = 0$. On l'écrit alors comme une fonction de deux variables $(t, x_0) \mapsto \Psi(t, x_0)$.

Le flot réduit est utilisé surtout dans le cas où l'équation différentielle considérée est autonome, de la forme (4). Dans ce cas en effet, la connaissance du flot réduit Ψ suffit à déterminer le flot complet Φ , car en raison de l'invariance par translation du temps, on a

$$\Phi(t, t_0, x_0) = \Psi(t - t_0, x_0).$$

Pour une équation autonome, les propriétés du flot réduit se déduisent aisément de celles du flot complet : son domaine de définition est un ouvert de $\mathbb{R} \times M$ contenant $\{0\} \times M$; pour tout $t \in \mathbb{R}$, l'ensemble des $x_0 \in M$ tels que (t, x_0) appartienne au domaine de définition de Ψ est un ouvert D_t de M (éventuellement vide), et l'application $x_0 \mapsto \Psi(t, x_0)$ est un difféomorphisme de D_t sur l'ouvert D_{-t} de M , d'inverse $x \mapsto \Psi(-t, x)$; les formules de composition deviennent

$$\Psi(t_1 + t_2, x) = \Psi(t_2, \Psi(t_1, x)), \quad x = \Psi(0, x) = \Psi(-t_1, \Psi(t_1, x)).$$

4. Différentielle extérieure

On a vu (exemple 1.2.c) qu'à toute fonction différentiable f sur une variété M on pouvait associer une forme de Pfaff df , appelée *différentielle* de cette fonction. La notion de *différentielle extérieure* d'une forme différentielle s'obtient en prolongeant, de manière naturelle, l'opérateur $d : f \mapsto df$ à l'algèbre des formes différentielles sur M . On peut énoncer :

4.1. Proposition. — *Soit M une variété différentiable, $A(M)$ l'algèbre des formes différentielles sur cette variété. Il existe une unique application linéaire de $A(M)$ dans elle-même, appelée différentielle extérieure et notée d , vérifiant les propriétés suivantes :*

1. *La différentielle extérieure est une dérivation de degré 1 de l'algèbre extérieure $A(M)$. Cela signifie que c'est une application linéaire qui, pour tout entier p , applique $A^p(M)$ dans $A^{p+1}(M)$, et qui vérifie, pour $\alpha \in A^p(M)$ et $\beta \in A^q(M)$,*

$$d(\alpha \wedge \beta) = (d\alpha) \wedge \beta + (-1)^p \alpha \wedge (d\beta).$$

2. *Elle vérifie*

$$d \circ d = 0.$$

3. *Pour toute fonction différentiable f sur M , considérée comme forme différentielle de degré 0, df est la différentielle de f au sens usuel.*

4.2. Expression en coordonnées locales. — Soit (U, φ) une carte admissible de la variété différentiable M , x^1, \dots, x^n les coordonnées locales associées. Soit η une forme différentielle de degré p ayant pour expression dans U

$$\eta = \sum_{1 \leq i_1 < \dots < i_p \leq n} \eta_{i_1 \dots i_p} dx^{i_1} \wedge \dots \wedge dx^{i_p}.$$

La différentielle extérieure de η a pour expression locale

$$d\eta = \sum_{1 \leq i_1 < \dots < i_p \leq n} d\eta_{i_1 \dots i_p} \wedge dx^{i_1} \wedge \dots \wedge dx^{i_p},$$

avec

$$d\eta_{i_1 \dots i_p} = \sum_{k=1}^n \frac{\partial \eta_{i_1 \dots i_p}}{\partial x^k} dx^k.$$

4.3. Quelques propriétés

a) Soit f une application différentiable d'une variété M dans une autre variété N , et α une forme différentielle sur N . On a

$$f^*(d\alpha) = d(f^*\alpha).$$

On exprime cette propriété en disant que la différentiation extérieure commute avec la prise d'image réciproque.

b) Une forme différentielle α de degré p sur la variété différentiable M est dite *fermée* si elle vérifie $d\alpha = 0$. Elle est dite *exacte* s'il existe une forme différentielle β , de degré $p - 1$, telle que $\alpha = d\beta$. Toute forme exacte est évidemment fermée. On appelle *p -ième espace de cohomologie réelle* (ou *de De Rham*) de la variété M , et on note $H^p(M, \mathbb{R})$, le quotient de l'espace vectoriel des p -formes fermées sur M par le sous-espace des p -formes exactes sur cette variété. La dimension de l'espace $H^p(M, \mathbb{R})$ est appelée *p -ième nombre de Betti* de la variété M . Les nombres de Betti dépendent de propriétés globales de la variété M ; ainsi par exemple, le premier nombre de Betti de \mathbb{R} est nul, tandis que celui du cercle S^1 est égal à 1.

5. Crochet de deux champs de vecteurs et dérivée de Lie

5.1. La notion générale de dérivée de Lie. — Soient M une variété différentiable, X un champ de vecteurs de classe C^∞ sur cette variété, et Ψ son flot réduit. Pour tout $t \in \mathbb{R}$, on note Ψ_t le difféomorphisme $x \mapsto \Psi_t(x) = \Psi(t, x)$, de l'ouvert D_t sur l'ouvert D_{-t} de M (voir 3.5).

a) *Cas d'une forme différentielle.* — Soit η une forme différentielle de degré p sur M . Pour chaque point $x \in M$ on peut, en choisissant le réel t assez petit en valeur absolue, faire en sorte que x soit élément de l'ouvert D_t . L'image réciproque $\Psi_t^*\eta$ de η par le difféomorphisme Ψ_t est donc bien définie au point x , et dépend différentiablement de t . On pose

$$\mathcal{L}(X)\eta(x) = \left. \frac{\partial \Psi_t^*\eta(x)}{\partial t} \right|_{t=0}.$$

Le point $x \in M$ étant quelconque, on obtient ainsi une forme différentielle $\mathcal{L}(X)\eta$, de degré p , sur la variété M , appelée *dérivée de Lie de η selon le champ de vecteurs X* .

b) *Cas d'une fonction.* — Soit f une fonction différentiable réelle sur la variété M . Cette fonction peut être considérée comme une forme différentielle de degré 0, ce qui permet de lui appliquer la définition précédente. L'application $t \mapsto \Psi_t^*f(x) = f(\Psi(t, x))$ n'est autre que la valeur de la fonction f le long de la courbe intégrale de X qui passe par le point x . On a donc

$$\mathcal{L}(X)f(x) = i(X)df(x) = \langle df(x), X(x) \rangle.$$

La dérivée de Lie $\mathcal{L}(X)f$ de la fonction f selon X est souvent appelée *dérivée de f dans la direction de X* , et notée $X.f$.

c) *Cas d'un champ de vecteurs.* — Soit Y un autre champ de vecteurs (de classe C^∞) sur M . Puisque Ψ_t est un difféomorphisme, on peut définir (voir 2.7) l'image réciproque de Y par Ψ_t , et définir la dérivée de Lie de Y selon X par

$$\mathcal{L}(X)Y(x) = \left. \frac{\partial \Psi_t^*Y(x)}{\partial t} \right|_{t=0}.$$

La dérivée de Lie $\mathcal{L}(X)Y$ de Y selon X est souvent appelée *crochet* (ou *crochet de Lie*) des champs de vecteurs X et Y , et notée $[X, Y]$. On verra plus loin qu'il est antisymétrique.

d) *Autres cas.* — Le même procédé permet de définir la dérivée de Lie, selon le champ de vecteurs X , de champs de tenseurs, ou même d'objets géométriques plus généraux.

5.2. Propriétés générales

a) Les hypothèses et notation étant les mêmes que dans le paragraphe ci-dessus on a, pour tout $t_0 \in \mathbb{R}$ et tout $x \in D_{t_0}$,

$$\left. \frac{\partial \Psi_t^* \eta(x)}{\partial t} \right|_{t=t_0} = \mathcal{L}(X)(\Psi_{t_0}^* \eta)(x) = \Psi_{t_0}^* (\mathcal{L}(X)\eta)(x).$$

Les mêmes formules restent valables lorsqu'on remplace la forme différentielle η par un champ de vecteurs Y .

b) Soient f une application différentiable d'une variété M dans une autre variété N , X et Y deux champs de vecteurs, respectivement sur M et sur N , tels que le couple (X, Y) soit compatible avec f , et η une forme différentielle sur N . On a

$$f^*(\mathcal{L}(Y)\eta) = \mathcal{L}(X)(f^*\eta).$$

De même, soient X_1, X_2 deux champs de vecteurs sur M , Y_1, Y_2 deux champs de vecteurs sur N , tels que les couples (X_1, Y_1) et (X_2, Y_2) soient compatibles avec f . Alors le couple $([X_1, X_2], [Y_1, Y_2])$ est compatible avec f .

5.3. Propriétés de la dérivée de Lie des formes différentielles. — Soit M une variété différentiable, $A(M)$ l'algèbre des formes différentielles, X, Y des champs de vecteurs sur cette variété.

a) L'opérateur $\mathcal{L}(X)$ de dérivation de Lie relativement à X est une dérivation de degré 0 de l'algèbre $A(M)$. Cela signifie que c'est une application linéaire de $A(M)$ dans elle-même, qui fait correspondre à toute forme différentielle $\eta \in A(M)$ une forme différentielle $\mathcal{L}(X)\eta$, de même degré que η , et qui agit sur le produit extérieur de deux formes différentielles $\eta \in A^p(M)$, $\zeta \in A^q(M)$ suivant la formule (identité de Leibniz)

$$\mathcal{L}(X)(\eta \wedge \zeta) = \mathcal{L}(X)\eta \wedge \zeta + \eta \wedge \mathcal{L}(X)\zeta.$$

b) On a déjà rencontré d'autres dérivations de l'algèbre $A(M)$: le produit intérieur par X , noté $i(X)$, de degré -1 ; la différentiation extérieure d , de degré 1. Il existe entre les trois dérivations $i(X)$, $\mathcal{L}(X)$ et d une relation très simple, appelée *formule de Cartan*,

$$\mathcal{L}(X) = i(X) \circ d + d \circ i(X).$$

c) On a aussi les formules, liant les dérivations $\mathcal{L}(X)$, $\mathcal{L}(Y)$, $\mathcal{L}([X, Y])$, $i(X)$, $i(Y)$ et $i([X, Y])$,

$$\begin{aligned}\mathcal{L}([X, Y]) &= \mathcal{L}(X) \circ \mathcal{L}(Y) - \mathcal{L}(Y) \circ \mathcal{L}(X), \\ i([X, Y]) &= \mathcal{L}(X) \circ i(Y) + i(Y) \circ \mathcal{L}(X).\end{aligned}$$

La première formule ci-dessus, restreinte à l'algèbre des fonctions différentiables réelles $A^0(M)$ sur la variété M , peut servir de définition au crochet de deux champs de vecteurs.

5.4. Propriétés du crochet de deux champs de vecteurs. — Soit M une variété différentiable, X, Y, Z des champs de vecteurs sur cette variété.

a) Le crochet de deux champs de vecteurs X et Y est bilinéaire (linéaire en chacun des deux arguments, X et Y) et antisymétrique

$$[Y, X] = -[X, Y].$$

b) La première formule indiquée en 5.3.c pour les formes différentielles, selon laquelle $\mathcal{L}([X, Y])$ est le commutateur de $\mathcal{L}(X)$ et de $\mathcal{L}(Y)$, est encore vraie lorsqu'on l'applique à un troisième champ de vecteurs Z . Avec la notation du crochet, et compte tenu de l'antisymétrie, elle prend la forme ci-dessous, appelée *identité de Jacobi* :

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0.$$

c) Soient η une forme différentielle de degré p , et X_i ($1 \leq i \leq p+1$) une famille de $p+1$ champs de vecteurs sur M . La formule suivante donne une expression de la différentielle extérieure $d\eta$:

$$\begin{aligned}d\eta(X_1, \dots, X_{p+1}) &= \sum_{i=1}^{p+1} (-1)^{i+1} X_i \cdot (\eta(X_1, \dots, \widehat{X}_i, \dots, X_{p+1})) \\ &+ \sum_{1 \leq i < j \leq p+1} (-1)^{i+j} \eta([X_i, X_j], X_1, \dots, \widehat{X}_i, \dots, \widehat{X}_j, \dots, X_{p+1}),\end{aligned}$$

où le signe $\widehat{}$ au dessus d'un terme indique l'omission de ce terme.

Chapitre III

Espaces vectoriels symplectiques

1. Forme bilinéaire antisymétrique sur un espace vectoriel

1.1. Généralités et notations. — Soit V un espace vectoriel réel de dimension finie et $\Omega : V \times V \rightarrow \mathbb{R}$ une forme bilinéaire antisymétrique sur V . On note $\Omega^b : V \rightarrow V^*$ l'application linéaire de V dans son dual définie par

$$\Omega^b(x) = -i(x)\Omega,$$

c'est-à-dire

$$\langle \Omega^b(x), y \rangle = -\Omega(x, y), \quad x \text{ et } y \in V.$$

On appelle *rang* de Ω la dimension de l'image de Ω^b , et *noyau* de Ω , noté $\ker \Omega$, le noyau de Ω^b .

Puisque Ω est antisymétrique, l'application transposée de Ω^b est $-\Omega^b$. Comme, d'une manière générale, l'annulateur du noyau d'une application linéaire est l'image de l'application transposée, on voit que

$$\Omega^b(V) = (\ker \Omega)^0 = \{ \alpha \in V^* \mid \langle \alpha, x \rangle = 0 \text{ pour tout } x \in \ker \Omega \}.$$

On a

$$\dim \ker \Omega + \dim \Omega^b(V) = \dim V.$$

1.2. Proposition. — *Le rang de Ω est un entier pair $2p$. Si $p \neq 0$ (c'est-à-dire si Ω n'est pas identiquement nulle) il existe $2p$ éléments f^1, \dots, f^{2p} de $\Omega^b(V)$, qui forment une base de ce sous-espace de V^* , tels que*

$$\Omega = f^1 \wedge f^2 + \dots + f^{2p-1} \wedge f^{2p}.$$

Preuve : Si $\Omega = 0$, son rang est nul, donc pair, et il n'y a rien de plus à démontrer. Faisons l'hypothèse de récurrence suivante : toute forme bilinéaire antisymétrique B de rang $\leq 2k$ est de rang pair $2p$ (avec $p \leq k$), et il existe $2p$ éléments de $B^b(V)$, notés $\alpha^1, \dots, \alpha^{2p}$, formant une base de $B^b(V)$, tels que

$$B = \alpha^1 \wedge \alpha^2 + \dots + \alpha^{2p-1} \wedge \alpha^{2p}.$$

Supposons Ω de rang $2k + 1$ ou $2k + 2$, donc non nulle. Il existe alors $x_1 \in V$ tel que $f^1 = \Omega^b(x_1) \neq 0$. On a

$$\langle f^1, x_1 \rangle = -\Omega(x_1, x_1) = 0.$$

Comme $f^1 \neq 0$, il existe $x_2 \in V$ tel que $\langle f^1, x_2 \rangle = -1$. Posons $\Omega^b(x_2) = f^2$. On a

$$\langle f^2, x_2 \rangle = -\Omega(x_2, x_2) = 0,$$

et

$$-1 = \langle f^1, x_2 \rangle = -\Omega(x_1, x_2) = \Omega(x_2, x_1) = -\langle f^2, x_1 \rangle. \quad (*)$$

Posons alors

$$\Omega' = \Omega - f^1 \wedge f^2.$$

Le noyau de Ω' contient celui de Ω , et contient aussi x_1 et x_2 , qui n'appartiennent pas au noyau de Ω et sont linéairement indépendants. Donc le rang de Ω' est inférieur ou égal à $2k$, puisque celui de Ω est inférieur ou égal à $2k+2$. On applique à Ω' l'hypothèse de récurrence : son rang est pair, égal à $2q$ ($q \leq k$), et il existe $2q$ éléments de $(\Omega')^b(V)$, notés $f^3, f^4, \dots, f^{2q+1}, f^{2q+2}$, qui forment une base de $(\Omega')^b(V)$, tels que

$$\Omega' = f^3 \wedge f^4 + \dots + f^{2q+1} \wedge f^{2q+2}.$$

On a alors

$$\Omega = f^1 \wedge f^2 + f^3 \wedge f^4 + \dots + f^{2q+1} \wedge f^{2q+2}. \quad (**)$$

D'après les relations (*) ci-dessus, f^1 et f^2 n'appartiennent pas à l'annulateur du noyau de Ω' , qui est égal à l'image de $(\Omega')^b$. Donc $f^1, f^2, f^3, f^4, \dots, f^{2q+1}, f^{2q+2}$ sont linéairement indépendants. L'expression (**) montre qu'ils engendrent l'image de Ω^b , donc ils en forment une base. \square

1.3. Définition. — Une forme bilinéaire antisymétrique Ω sur l'espace vectoriel réel V , de dimension finie, est dite *symplectique* si son rang est égal à la dimension de V . Un espace vectoriel V , muni d'une forme symplectique Ω , est appelé *espace vectoriel symplectique* et noté (V, Ω) .

1.4. Propriétés élémentaires

Soit (V, Ω) un espace vectoriel symplectique.

a) D'après 1.2, le rang de Ω est pair. Comme on le suppose égal à la dimension de V , celle-ci est paire; on la notera $2n$.

b) Dans le cas présent, Ω^b est un isomorphisme de V sur son dual V^* . On notera $\Omega^\#$ son inverse.

c) En ordonnant différemment les éléments de la base de $\Omega^b(V) = V^*$ construite lors de la démonstration de 1.2, on voit qu'il existe une base (pas nécessairement unique) $(\epsilon^1, \dots, \epsilon^{2n})$ de V^* telle que

$$\Omega = \sum_{i=1}^n \epsilon^{n+i} \wedge \epsilon^i.$$

Soit (e_1, \dots, e_{2n}) la base de V duale de la base $(\epsilon^1, \dots, \epsilon^{2n})$ de V^* . On a pour tous $i, j \in \{1, \dots, n\}$:

$$\Omega(e_i, e_j) = 0, \quad \Omega(e_{n+i}, e_{n+j}) = 0, \quad \Omega(e_{n+i}, e_j) = \delta_{ij}.$$

Une base (e_1, \dots, e_{2n}) de l'espace vectoriel symplectique (V, Ω) vérifiant ces relations est dite *canonique*. On a donc montré que tout espace vectoriel symplectique admet une base canonique. Par suite, deux espaces vectoriels symplectiques de même dimension sont isomorphes, par un isomorphisme qui échange leurs formes symplectiques. On remarquera que de ce point de vue, la situation est plus simple que dans le cas des espaces vectoriels euclidiens ou pseudo-euclidiens munis d'une forme bilinéaire symétrique non dégénérée, puisque pour une dimension fixée de l'espace, la signature de cette forme peut prendre diverses valeurs.

d) On définit sur \mathbb{R}^{2n} une structure symplectique en posant, pour $x = (x^1, \dots, x^{2n})$ et $y = (y^1, \dots, y^{2n}) \in \mathbb{R}^{2n}$,

$$\Omega(x, y) = \sum_{i=1}^n (x^{n+i} y^i - x^i y^{n+i}).$$

Pour cette structure, dite *standard*, la base usuelle de \mathbf{R}^{2n} est une base canonique, au sens indiqué ci-dessus.

e) Soit (e_1, \dots, e_{2n}) une base canonique de l'espace vectoriel symplectique (V, Ω) . Posons

$$\Omega_{ij} = \Omega(e_i, e_j).$$

La matrice $J = (\Omega_{ij})$ a pour expression

$$J = \begin{pmatrix} \Omega_{11} & \dots & \Omega_{1\ 2n} \\ \vdots & & \vdots \\ \Omega_{2n\ 1} & \dots & \Omega_{2n\ 2n} \end{pmatrix} = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix},$$

où I_n désigne la matrice $n \times n$ unité.

En convenant d'identifier les éléments x et y de V avec les éléments de \mathbb{R}^{2n} qui les représentent dans la base considérée, on peut écrire, en utilisant les notations matricielles usuelles (${}^t x$ désignant le vecteur-ligne transposé du vecteur-colonne x à $2n$ composantes) :

$$\Omega(x, y) = {}^t x J y.$$

2. Orthogonalité symplectique

2.1. Définition. — Soit (V, Ω) un espace vectoriel symplectique et W un sous-espace vectoriel de V . On appelle *orthogonal symplectique* (ou, lorsqu'il n'y a pas de risque de confusion, *orthogonal*) de W , le sous-espace vectoriel de V :

$$\text{orth } W = \{ x \in V \mid \Omega(x, y) = 0 \text{ pour tout } y \in W \}.$$

2.2. Propriétés élémentaires. — Soit (V, Ω) un espace vectoriel symplectique, W, W_1, W_2 des sous-espaces vectoriels de V . Le lecteur établira aisément les propriétés suivantes.

a) On a $\text{orth}(\text{orth } W) = W$.

b) On a $\dim W + \dim \text{orth } W = \dim V$.

c) On a

$$\Omega^b(\text{orth } W) = W^0 = \{ \eta \in V^* \mid \langle \eta, x \rangle = 0 \text{ pour tout } x \in W \}.$$

d) On a $W_1 \subset W_2$ si et seulement si $\text{orth } W_1 \supset \text{orth } W_2$.

e) On a

$$\text{orth}(W_1 \cap W_2) = \text{orth } W_1 + \text{orth } W_2,$$

$$\text{orth}(W_1 + W_2) = \text{orth } W_1 \cap \text{orth } W_2.$$

2.3. Définitions. — Soit (V, Ω) un espace vectoriel symplectique. Un sous-espace vectoriel W de V est dit

- isotrope si $W \subset \text{orth } W$,
- coïsothrope si $\text{orth } W \subset W$,
- lagrangien si $\text{orth } W = W$,
- symplectique si $W \cap \text{orth } W = \{0\}$.

2.4. Exemples

a) Tout sous-espace de dimension 1 est isotrope; tout sous-espace de codimension 1 est coïsothrope.

b) Un sous-espace W de (V, Ω) est isotrope si et seulement si $\text{orth } W$ est coïsothrope.

c) Soit (V, Ω) un espace vectoriel symplectique, (e_1, \dots, e_{2n}) une base canonique. Les sous-espaces engendrés par (e_1, \dots, e_n) et (e_{n+1}, \dots, e_{2n}) sont lagrangiens, supplémentaires l'un de l'autre. Plus généralement, soit I une partie de $\{1, 2, \dots, n\}$, et J son complémentaire. Le sous-espace engendré par les e_i , $i \in I$, et les e_{n+j} , $j \in J$, est lagrangien, ainsi que celui engendré par les e_{n+i} , $i \in I$, et les e_j , $j \in J$, et ces deux sous-espaces lagrangiens sont supplémentaires. Le sous-espace engendré par les e_i et les e_{n+i} , $i \in I$, est symplectique.

2.5. Proposition. — Soit (V, Ω) un espace vectoriel symplectique et I un sous-espace vectoriel isotrope de V . Il existe un sous-espace lagrangien L tel que $I \subset L \subset \text{orth } I$.

Preuve : On a $I \subset \text{orth } I$. Si $I = \text{orth } I$, I est lagrangien et on peut prendre $L = I$. Dans le cas contraire, il existe un élément x_1 de $\text{orth } I$ qui n'est pas élément de I . Soit $I_1 = I + \langle x_1 \rangle$ (en notant $\langle x_1 \rangle$ le sous-espace vectoriel de dimension 1 engendré par x_1). On vérifie aisément que I_1 est isotrope et que $I \subset I_1 \subset \text{orth } I_1 \subset \text{orth } I$. Si $I_1 = \text{orth } I_1$, on peut prendre $L = I_1$; dans le cas contraire, on répète la même construction en remplaçant I par I_1 . Après un nombre fini de constructions de ce type, on aboutit nécessairement (puisque à chaque opération, la codimension de l'espace isotrope considéré dans son orthogonal décroît de deux unités) à un sous-espace I_n égal à son orthogonal, donc lagrangien, contenant I et contenu dans $\text{orth } I$. \square

2.6. Proposition. — Soit (V, Ω) un espace vectoriel symplectique et W un sous-espace vectoriel de V . Le noyau de la forme bilinéaire Ω_W induite par Ω sur W (c'est-à-dire restriction de Ω à $W \times W$) est

$$\ker \Omega_W = W \cap \text{orth } W .$$

La démonstration, facile, est laissée au lecteur.

2.7. Conséquences

a) Les formes bilinéaires Ω_W et $\Omega_{\text{orth } W}$ induites par Ω sur W et sur son orthogonal ont même noyau. On en déduit :

$$\text{rang}(\Omega_W) - \text{rang}(\Omega_{\text{orth } W}) = 2 \dim W - \dim V .$$

b) Un sous-espace vectoriel W de V est isotrope si et seulement si $\text{rang}(\Omega_W) = 0$, coisotrope si et seulement si $\text{rang}(\Omega_W) = 2 \dim W - \dim V$, symplectique si et seulement si $\text{rang}(\Omega_W) = \dim W$. On voit donc qu'un sous-espace vectoriel symplectique W au sens de la définition 2.3 est bien un sous-espace sur lequel la forme Ω induit une forme symplectique Ω_W .

2.8. Proposition. — Soit (V, Ω) un espace vectoriel symplectique, L un sous-espace lagrangien, (e_1, \dots, e_n) une base de L . Il existe n autres éléments e_{n+1}, \dots, e_{2n} de V tels que (e_1, \dots, e_{2n}) soit une base canonique.

Preuve : L'annulateur L^0 de L (sous-espace du dual V^* de V constitué par les formes linéaires nulles sur L) s'identifie au dual $(V/L)^*$ de l'espace quotient V/L , et on sait que $(\Omega^b(e_1), \dots, \Omega^b(e_n))$ en est une base. Soit $(\dot{f}_1, \dots, \dot{f}_n)$ la base duale de V/L . Pour chaque i , $1 \leq i \leq n$, on choisit un représentant f_i de \dot{f}_i . Par construction :

$$\Omega(e_i, f_j) = -\langle \Omega^b(e_i), f_j \rangle = -\delta_{ij} .$$

Posons

$$e_{n+i} = f_i + 1/2 \sum_{k=1}^n \Omega(f_i, f_k) e_k .$$

On vérifie aisément que (e_1, \dots, e_{2n}) est une base canonique. \square

2.9. Corollaire. — Tout sous-espace lagrangien d'un espace vectoriel symplectique possède un supplémentaire lagrangien.

Preuve : Il suffit en effet de choisir une base (e_1, \dots, e_n) du sous-espace lagrangien donné L , de compléter cette base en une base canonique (e_1, \dots, e_{2n}) de V , et de considérer le sous-espace engendré par (e_{n+1}, \dots, e_{2n}) . \square

2.10. Proposition. — Soient L_1 et L_2 deux sous-espaces vectoriels lagrangiens supplémentaires de l'espace vectoriel symplectique (V, Ω) . L'application de L_2 dans le dual L_1^* de L_1 :

$$y \mapsto -\Omega^b(y) \Big|_{L_1}$$

est un isomorphisme. Si on l'utilise pour identifier V à $L_1 \times L_1^*$, l'expression de Ω devient

$$\Omega((x_1, \alpha_1), (x_2, \alpha_2)) = \langle \alpha_1, x_2 \rangle - \langle \alpha_2, x_1 \rangle.$$

Si (e_1, \dots, e_n) est une base de L_1 et (e_{n+1}, \dots, e_{2n}) la base duale de L_1^* , (e_1, \dots, e_{2n}) est une base canonique de V (identifié à $L_1 \times L_1^*$).

La démonstration ne présente pas de difficulté.

3. Réduction d'un espace vectoriel symplectique

3.1. Proposition. — Soit (V, Ω) un espace vectoriel symplectique et W un sous-espace vectoriel de V . On considère le quotient

$$\dot{W} = W / (W \cap \text{orth } W),$$

et on note $p : W \rightarrow \dot{W}$ la projection canonique.

1. La forme Ω_W induite par Ω sur W est projetable par p sur \dot{W} , et sa projection $\dot{\Omega}$ est une forme symplectique sur \dot{W} .
2. Si I un sous-espace isotrope de V , $\dot{I} = p(I \cap W)$ est un sous-espace isotrope de $(\dot{W}, \dot{\Omega})$. Si I est lagrangien et si, de plus, W est coïsoptrope, \dot{I} est lagrangien.

Preuve :

1. L'espace \dot{W} est précisément le quotient de W par le noyau de Ω_W . On voit alors aisément qu'il existe une forme bilinéaire unique $\dot{\Omega}$ sur \dot{W} dont l'image réciproque par p est Ω_W , et que $\dot{\Omega}$ est symplectique.
2. Si I est isotrope, la forme induite par Ω sur $W \cap I$ est nulle. On en déduit aisément que \dot{I} est isotrope. Si I est lagrangien et W coïsoptrope, on sait déjà que \dot{I} est isotrope, il suffit donc de montrer qu'il est coïsoptrope. Soit $\dot{x} \in \text{orth } \dot{I}$, et $x \in W$ un représentant de \dot{x} . Pour tout $y \in I \cap W$, on a $\Omega(x, y) = 0$, donc $x \in W \cap \text{orth}(I \cap W)$. Mais $\text{orth}(I \cap W) = \text{orth } I + \text{orth } W = I + \text{orth } W$ puisque I est lagrangien. Donc $x = x_1 + x_2$, avec $x_1 \in I$ et $x_2 \in \text{orth } W$. Mais puisque W est coïsoptrope, $\text{orth } W \subset W$, donc $x_2 \in W$. Comme $x \in W$, $x_1 = x - x_2$ est aussi élément de W , tandis que x_2 appartient au noyau de p . Par suite, x_1 est un autre représentant de \dot{x} . Comme $x_1 \in I$, $\dot{x} \in \dot{I}$; on a prouvé que $\text{orth } \dot{I} \subset \dot{I}$. \square

On verra plus tard que la proposition précédente se généralise, et permet de définir la réduction d'une sous-variété d'une variété symplectique. Cette notion a de très importantes applications.

4. Permutations canoniques

Soit (V, Ω) un espace vectoriel symplectique et (e_1, \dots, e_{2n}) une base canonique de V . Il est utile, dans certains cas, de modifier l'ordre des vecteurs éléments de cette base, tout en faisant en sorte que la base reste canonique. Cela permettra par exemple de mettre en évidence un sous-espace vectoriel lagrangien (engendré par les vecteurs de la base qui occupent, pour le nouvel ordre, les n premières places) adapté au problème qu'on a à traiter. Le lemme suivant (dû à Arnol'd) donne un exemple de construction de ce type.

4.1. Lemme. — *Soit (V, Ω) un espace vectoriel symplectique, (e_1, \dots, e_{2n}) une base canonique et X un sous-espace lagrangien de V . Il existe une partie I de $\{1, \dots, n\}$ telle que, si on note J son complémentaire, le sous-espace engendré par les e_i et les e_{n+j} , avec $i \in I$ et $j \in J$, soit lagrangien et supplémentaire de X .*

Preuve : Soit L' le sous-espace engendré par e_1, \dots, e_n et L'' le sous-espace engendré par e_{n+1}, \dots, e_{2n} . Soit $k = \dim(L' \cap X)$.

Si $k = n$, $I = \emptyset$ répond à la question. De même, si $k = 0$, $I = \{1, \dots, n\}$ répond à la question.

Supposons donc $1 < k < n$. $L' \cap X$ étant un sous-espace vectoriel de dimension k de L' , et (e_1, \dots, e_n) étant une base de L' , il existe une famille $(e_{i_1}, \dots, e_{i_{n-k}})$ de $n - k$ vecteurs pris dans cette base qui engendrent un supplémentaire Y' de $L' \cap X$ dans L' . Posons $I = \{i_1, \dots, i_{n-k}\}$, et notons J son complémentaire. Soit Y'' le sous-espace vectoriel engendré par les e_{n+j} , $j \in J$. On sait (exemple 2.4.3) que $Y = Y' \oplus Y''$ est lagrangien. Il reste à prouver qu'il est supplémentaire de X . On a

$$L' = (L' \cap X) \oplus Y', \quad L' \cap X \subset X, \quad Y' = L' \cap Y \subset Y,$$

donc $L' \subset X + Y$. Comme L' , X et Y sont lagrangiens, on a

$$L' = \text{orth } L' \supset \text{orth}(X + Y) = \text{orth } X \cap \text{orth } Y = X \cap Y,$$

donc

$$X \cap Y = X \cap Y \cap L' = (L' \cap X) \cap (L' \cap Y) = \{0\},$$

car $L' \cap X$ et $L' \cap Y$ sont supplémentaires dans L' . Donc X et Y sont supplémentaires dans V . □

4.2. Corollaire. — *Dans les hypothèses du lemme précédent, il existe une permutation σ de $\{1, \dots, 2n\}$ et une application ϵ de $\{1, \dots, 2n\}$ dans $\{-1, 1\}$ telle que $(\epsilon(1)e_{\sigma(1)}, \dots, \epsilon(2n)e_{\sigma(2n)})$ soit une base canonique de V et que le sous-espace lagrangien engendré par les n premiers vecteurs de cette base soit supplémentaire de X .*

Preuve : Il suffit de ranger les éléments de la base canonique donnée dans l'ordre suivant : d'abord les e_i , puis les e_{n+j} , puis les e_{n+i} , et enfin les e_j (avec $i \in I$ et $j \in J$). On remplacera ensuite les e_j par leurs opposés, afin de respecter les conventions de signes figurant dans la définition d'une base canonique. \square

5. Le groupe symplectique

5.1. Définition. — Soient (V, Ω) et (V', Ω') deux espaces vectoriels symplectiques, et $f : V \rightarrow V'$ une application linéaire. On dit que f est symplectique si

$$f^*\Omega' = \Omega.$$

Lorsque V et V' sont de même dimension, une application symplectique $f : V \rightarrow V'$ est appelée isomorphisme symplectique.

5.2. Quelques propriétés

a) On remarque qu'un isomorphisme symplectique $f : V \rightarrow V'$ est un isomorphisme d'espaces vectoriels, car si on désigne par $2n$ la dimension de V (et de V'), Ω^n et $(\Omega')^n$ sont des formes éléments de volume sur V et sur V' , respectivement; on a

$$f^*((\Omega')^n) = \Omega^n,$$

ce qui prouve que f est bien un isomorphisme.

b) La composée de deux applications symplectiques est une application symplectique. L'inverse d'un isomorphisme symplectique est un isomorphisme symplectique.

5.3. Définition. — Soit (V, Ω) un espace vectoriel symplectique. On appelle groupe symplectique de (V, Ω) , et on note $\mathbf{Sp}(V, \Omega)$ l'ensemble des isomorphismes symplectiques de V sur lui-même.

5.4. Quelques propriétés

a) D'après les propriétés 5.2.2, $\mathbf{Sp}(V, \Omega)$ est un sous-groupe du groupe linéaire $\mathbf{GL}(V)$. C'est un sous-groupe fermé puisqu'il est défini par

$$\mathbf{Sp}(V, \Omega) = \{ g \in \mathbf{GL}(V) \mid g^*\Omega = \Omega \}.$$

D'après un théorème de Cartan, $\mathbf{Sp}(V, \Omega)$ est un sous-groupe de Lie de $\mathbf{GL}(V)$.

b) Le déterminant $\det g$ de tout élément g de $\mathbf{Sp}(V, \Omega)$ est égal à 1. En effet, Ω^n est une forme élément de volume sur V et on a, par définition du déterminant :

$$g^*(\Omega^n) = \det g \Omega^n, \quad \text{et} \quad g^*(\Omega^n) = \Omega^n.$$

c) Une application linéaire de l'espace vectoriel symplectique (V, Ω) dans lui-même est élément du groupe symplectique si et seulement si elle transforme une base canonique de V en une autre base canonique de V .

d) Une application élément de $\mathbf{Sp}(V, \Omega)$ transforme tout sous-espace vectoriel isotrope (resp., coïsothrope, resp., lagrangien) de V en un autre sous-espace isotrope (resp., coïsothrope, resp., lagrangien).

e) Le groupe symplectique agit transitivement sur $V \setminus \{0\}$, sur l'ensemble des bases canoniques de (V, Ω) , sur l'ensemble des sous-espaces lagrangiens, sur l'ensemble des couples de sous-espaces lagrangiens transverses, sur l'ensemble des couples de sous-espaces lagrangiens dont l'intersection est de dimension donnée.

5.5. Matrices symplectiques. — Après avoir choisi une base canonique de l'espace vectoriel symplectique (V, Ω) , on peut identifier celui-ci à \mathbb{R}^{2n} . Ainsi qu'on l'a vu en 1.4.e, on peut écrire, en utilisant les notations matricielles,

$$\Omega(x, y) = {}^t x J y, \quad \text{avec} \quad J = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix}, \quad x \text{ et } y \in \mathbb{R}^{2n}.$$

Une application linéaire de V dans V est élément du groupe symplectique si et seulement si la matrice $2n \times 2n$ A qui la représente vérifie

$${}^t x {}^t A J A y = {}^t x J y \quad \text{pour tous } x \text{ et } y \in \mathbb{R}^{2n},$$

c'est-à-dire si et seulement si

$${}^t A J A = J.$$

Les matrices vérifiant cette propriété sont appelées *matrices symplectiques*. Elles forment un groupe, isomorphe au groupe symplectique de (V, Ω) , appelé *groupe symplectique linéaire*, noté $\mathbf{Sp}(2n, \mathbb{R})$.

On remarque que si A est une matrice symplectique, son déterminant est égal à 1, et sa transposée ${}^t A$ est symplectique.

5.6. L'algèbre de Lie du groupe symplectique. — Soit (V, Ω) un espace vectoriel symplectique de dimension $2n$. Puisque $\mathbf{Sp}(V, \Omega)$ est un sous-groupe de Lie du groupe linéaire $\mathbf{GL}(V)$, son algèbre de Lie, notée $\mathfrak{sp}(V, \Omega)$, est une sous-algèbre de Lie de l'algèbre de Lie du groupe linéaire. On sait que celle-ci est l'espace des endomorphismes linéaires de V , avec pour crochet le commutateur. Soit donc $X \in \mathcal{L}(V, V)$ une application linéaire de V dans lui-même. Cherchons à quelle condition X est élément de $\mathfrak{sp}(V, \Omega)$. Pour tout réel $t \in \mathbb{R}$, posons

$$g(t) = \exp(tX).$$

On rappelle que $g(t)$ est la solution maximale de l'équation différentielle

$$\frac{d}{dt} g(t) = X g(t) \quad (*)$$

qui vérifie la donnée de Cauchy

$$g(0) = \text{id}_V .$$

Remarquons d'ailleurs que $g(t)$ commute avec X et qu'on pourrait écrire $g(t)X$ au lieu de $Xg(t)$ au second membre de l'équation différentielle utilisée pour définir $g(t) = \exp(tX)$.

D'après les propriétés connues de l'application exponentielle, l'endomorphisme X est élément de l'algèbre de Lie du groupe symplectique si et seulement si, pour tout $t \in \mathbb{R}$, $g(t) = \exp(tX)$ est élément du groupe symplectique, c'est-à-dire vérifie, pour tous x et $y \in V$,

$$\Omega(g(t)x, g(t)y) = \Omega(x, y) .$$

En dérivant ceci par rapport à t puis en faisant $t = 0$, on voit, compte tenu de (*), qu'une condition nécessaire que doit vérifier X pour appartenir à $\mathfrak{sp}(V, \Omega)$ est

$$\Omega(Xx, y) + \Omega(x, Xy) = 0 \quad \text{pour tous } x \text{ et } y \in V .$$

En utilisant encore (*), on voit aisément que cette condition est aussi suffisante.

Moyennant le choix d'une base canonique, on se ramène, comme dans le paragraphe précédent, au cas où $V = \mathbb{R}^{2n}$. Une matrice $2n \times 2n$ A est élément de l'algèbre de Lie $\mathfrak{sp}(2n, \mathbb{R})$ si et seulement si, pour tous x et $y \in \mathbb{R}^{2n}$,

$${}^t x {}^t A J y + {}^t x J A y = 0 ,$$

c'est-à-dire si et seulement si

$${}^t A J + J A = 0 .$$

En écrivant A sous la forme

$$A = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix} ,$$

où A_1, A_2, A_3 et A_4 sont des matrices $n \times n$, et en portant cette expression dans la condition ci-dessus, on voit que A est élément de $\mathfrak{sp}(2n, \mathbb{R})$ si et seulement si

$${}^t A_1 = -A_4, \quad {}^t A_2 = A_2, \quad {}^t A_3 = A_3 ,$$

c'est-à-dire si et seulement si A est de la forme

$$A = \begin{pmatrix} A_1 & A_2 \\ A_3 & -{}^t A_1 \end{pmatrix} , \quad \text{avec } {}^t A_2 = A_2, \quad {}^t A_3 = A_3 .$$

Il est facile d'en déduire la dimension de $\mathfrak{sp}(2n, \mathbb{R})$ (qui est aussi celle de $\mathbf{Sp}(2n, \mathbb{R})$) :

$$\dim \mathfrak{sp}(2n, \mathbb{R}) = 2n^2 + n .$$

6. Structures complexes adaptées

6.1. Quelques rappels sur les structures complexes. — Soit V un espace vectoriel réel de dimension finie. On appelle *structure complexe*, ou *opérateur complexe* sur V , toute application linéaire \mathcal{I} de V dans lui-même vérifiant $\mathcal{I}^2 = -\text{id}_V$.

On remarque que s'il existe sur V une structure complexe \mathcal{I} , celle-ci vérifie

$$(\det \mathcal{I})^2 = \det(-\text{id}_V) = (-1)^{\dim V}.$$

Ceci n'est possible que si $\dim V$ est paire. Réciproquement, sur un espace vectoriel V de dimension paire $2n$, il existe une structure complexe : il suffit en effet, pour en définir une, de prendre une base (e_1, \dots, e_{2n}) de V et de poser :

$$\mathcal{I}(e_i) = e_{n+i}, \quad \mathcal{I}(e_{n+i}) = -e_i, \quad (1 \leq i \leq n).$$

La donnée d'une structure complexe \mathcal{I} sur un espace vectoriel réel V permet de munir V d'une structure d'espace vectoriel complexe, en posant, pour tout $x \in V$ et tout $a + ib \in \mathbb{C}$ (a et $b \in \mathbb{R}$, $i = \sqrt{-1}$) :

$$(a + ib)x = ax + b\mathcal{I}(x).$$

On notera $V_{\mathbb{C}}$ l'espace vectoriel V muni de cette structure d'espace vectoriel complexe. On remarquera que la dimension (complexe) de $V_{\mathbb{C}}$ est égale à la moitié de la dimension (réelle) de V .

Réciproquement, soit $V_{\mathbb{C}}$ un espace vectoriel complexe de dimension finie. Notons V le même espace, muni de la structure d'espace vectoriel réel sous-jacente. Posons pour tout $x \in V$:

$$\mathcal{I}(x) = ix,$$

en convenant de considérer x , figurant au second membre, comme élément de $V_{\mathbb{C}}$, ce qui permet de le multiplier par $i = \sqrt{-1}$. On voit immédiatement que \mathcal{I} est une structure complexe sur V , et que la structure d'espace vectoriel complexe qui lui correspond est celle initialement donnée.

Une *forme réelle* d'un espace vectoriel réel V muni d'une structure complexe \mathcal{I} est un sous-espace vectoriel réel W de V , tel que

$$V = W \oplus \mathcal{I}(W).$$

Pour trouver une forme réelle de V , il suffit de prendre une base (e_1, \dots, e_n) de $V_{\mathbb{C}}$, et de définir W comme le sous-espace vectoriel réel engendré par (e_1, \dots, e_n) . Soit $V_{\mathbb{C}}$ un espace vectoriel complexe, et $\eta : V_{\mathbb{C}} \times V_{\mathbb{C}} \rightarrow \mathbb{C}$ une forme sesquilinéaire sur $V_{\mathbb{C}}$, c'est-à-dire une application linéaire par rapport à sa première variable, antilinéaire par rapport à la seconde, vérifiant pour tous x et $y \in V_{\mathbb{C}}$:

$$\eta(y, x) = \overline{\eta(x, y)}.$$

On dit que η est une forme *pseudohermitienne* si elle est non dégénérée, c'est-à-dire si pour tout $x \in V_{\mathbb{C}}$, $x \neq 0$, il existe $y \in V_{\mathbb{C}}$ tel que $\eta(x, y) \neq 0$. On dit que η est *hermitienne* si pour tout $x \in V_{\mathbb{C}}$, $x \neq 0$, on a $\eta(x, x) > 0$.

Soit $V_{\mathbb{C}}$ un espace vectoriel complexe de dimension finie et η une forme pseudohermitienne (ou hermitienne) sur cet espace. Pour tous x et $y \in V_{\mathbb{C}}$, notons $G(x, y)$ et $\Omega(x, y)$ la partie réelle et la partie imaginaire de $\eta(x, y)$:

$$\eta(x, y) = G(x, y) + i\Omega(x, y).$$

On vérifie alors que G est un produit scalaire pseudo-euclidien (c'est-à-dire une forme bilinéaire symétrique non dégénérée) et Ω une forme symplectique sur l'espace vectoriel réel V . La forme η sur $V_{\mathbb{C}}$ est hermitienne si et seulement si G est un produit scalaire euclidien sur V , c'est-à-dire une forme bilinéaire symétrique définie positive.

Dans la définition suivante, on considère la construction inverse de celle décrite ci-dessus : une structure symplectique étant donnée sur un espace vectoriel réel, on veut munir cet espace d'une structure complexe et d'une forme pseudo-hermitienne (ou hermitienne) ayant pour partie imaginaire la forme symplectique donnée.

6.2. Définition. — Soit (V, Ω) un espace vectoriel symplectique. Une structure complexe \mathcal{I} sur V est dite *adaptée* à Ω s'il existe sur $V_{\mathbb{C}}$ une forme pseudohermitienne η ayant pour partie imaginaire Ω . On dira que \mathcal{I} est *strictement adaptée* à Ω si elle est adaptée et si la forme η est hermitienne.

6.3. Remarque. — Il est facile de voir que sur tout espace vectoriel symplectique (V, Ω) il existe une structure complexe adaptée (non unique en général). En effet, soit (e_1, \dots, e_{2n}) une base canonique de (V, Ω) . On définit \mathcal{I} en posant

$$\mathcal{I}(e_i) = e_{n+i}, \quad \mathcal{I}(e_{n+i}) = -e_i, \quad (1 \leq i \leq n).$$

On voit alors que (e_1, \dots, e_n) est une base de $V_{\mathbb{C}}$. Soient z et z' deux éléments de $V_{\mathbb{C}}$, ayant pour composantes dans cette base (z_1, \dots, z_n) et (z'_1, \dots, z'_n) , respectivement. On définit la forme hermitienne η en posant

$$\eta(z, z') = \sum_{k=1}^n z_k \overline{z'_k}.$$

On voit aisément que sa partie imaginaire est Ω .

6.4. Proposition. — Soit (V, Ω) un espace vectoriel symplectique et $\mathcal{I} : V \rightarrow V$ une application linéaire. Les propriétés suivantes sont équivalentes.

1. \mathcal{I} est un opérateur complexe adapté à Ω .
2. Ω et \mathcal{I} vérifient, pour tous x et $y \in V$,

$$\Omega(x, y) = \Omega(\mathcal{I}x, \mathcal{I}y),$$

et la forme bilinéaire G définie par

$$G(x, y) = \Omega(\mathcal{I}x, y),$$

est symétrique.

Lorsque ces propriétés équivalentes sont satisfaites, la forme pseudo-hermitienne η dont Ω est la partie imaginaire est donnée par

$$\eta(x, y) = G(x, y) + i\Omega(x, y) = \Omega(\mathcal{I}x, y) + i\Omega(x, y).$$

La forme η est hermitienne si et seulement si G est définie positive.

Preuve : Supposons que \mathcal{I} soit une structure complexe adaptée, et soit η la forme pseudohermitienne ayant Ω pour partie imaginaire. On a

$$\eta(\mathcal{I}x, \mathcal{I}y) = -i^2\eta(x, y) = \eta(x, y),$$

d'où, en distinguant les parties réelle et imaginaire,

$$\Omega(\mathcal{I}x, \mathcal{I}y) = \Omega(x, y), \quad G(\mathcal{I}x, \mathcal{I}y) = G(x, y).$$

D'autre part

$$\eta(\mathcal{I}x, y) = i\eta(x, y) = i(G(x, y) + i\Omega(x, y)) = G(\mathcal{I}x, y) + i\Omega(\mathcal{I}x, y),$$

d'où

$$G(x, y) = \Omega(\mathcal{I}x, y),$$

ce qui prouve que G , définie par la formule indiquée dans l'énoncé, est la partie réelle de η , donc est symétrique.

Réciproquement, supposons les propriétés 2 vérifiées. On a alors

$$\Omega(\mathcal{I}^2x, y) = G(\mathcal{I}x, y) = G(y, \mathcal{I}x) = \Omega(\mathcal{I}y, \mathcal{I}x) = -\Omega(x, y).$$

Ω étant non dégénérée, cela prouve que $\mathcal{I}^2 = -\text{id}_V$, c'est-à-dire que \mathcal{I} est une structure complexe. Il est alors facile de vérifier que $\eta = G + i\Omega$ est une forme pseudohermitienne (hermitienne si et seulement si G est définie positive). \square

6.5. Remarque. — Si \mathcal{I} est une structure complexe adaptée à Ω , il vérifie, pour tous x et $y \in V$:

$$\Omega(\mathcal{I}x, \mathcal{I}y) = \Omega(x, y), \quad G(\mathcal{I}x, \mathcal{I}y) = G(x, y), \quad \eta(\mathcal{I}x, \mathcal{I}y) = \eta(x, y).$$

6.6. Proposition. — Soit (V, Ω) un espace vectoriel symplectique et \mathcal{I} une structure complexe strictement adaptée à Ω . Alors tout sous-espace vectoriel complexe de $V_{\mathbb{C}}$ est un sous-espace symplectique de (V, Ω) ; tout sous-espace lagrangien de (V, Ω) est une forme réelle de $V_{\mathbb{C}}$.

Preuve : Convenons de noter orth_G l'orthogonalité euclidienne (relativement au produit scalaire euclidien G), et orth_{Ω} l'orthogonalité symplectique. Compte tenu

de $G(x, y) = \Omega(\mathcal{I}x, y)$ et de $\mathcal{I}^2 = -\text{id}_V$, on voit que, pour tout sous-espace vectoriel réel W de V

$$\text{orth}_G W = \text{orth}_\Omega(\mathcal{I}W), \quad \text{orth}_G(\mathcal{I}W) = \text{orth}_\Omega W.$$

Or W est un sous-espace vectoriel complexe de $V_{\mathbb{C}}$ si et seulement si $W = \mathcal{I}W$, c'est-à-dire si et seulement si $\text{orth}_\Omega W = \text{orth}_G W$. Mais l'orthogonal euclidien de W étant supplémentaire de W , cela implique que W est symplectique.

De même, W est lagrangien si et seulement si $\text{orth}_\Omega W = W$, c'est-à-dire si et seulement si $\text{orth}_G(\mathcal{I}W) = W$. Mais l'orthogonal euclidien de $\mathcal{I}W$ étant supplémentaire de $\mathcal{I}W$, cela implique que W est une forme réelle de $V_{\mathbb{C}}$. \square

6.7. Exercice. — Soit (V, Ω) un espace vectoriel symplectique et \mathcal{I} un opérateur complexe strictement adapté à Ω . Soit $g \in \mathbf{GL}(V)$. On considère les trois propriétés :

1. g est un isomorphisme symplectique;
2. g est un isomorphisme orthogonal de (V, G) ;
3. g est un isomorphisme complexe de $V_{\mathbb{C}}$.

Montrer que lorsque g vérifie deux des propriétés ci-dessus, il vérifie aussi la troisième, et que c'est un isomorphisme unitaire de $(V_{\mathbb{C}}, \eta)$. En déduire que

$$\begin{aligned} \mathbf{U}(V_{\mathbb{C}}, \eta) &= \mathbf{Sp}(V, \Omega) \cap \mathbf{O}(V, G) \\ &= \mathbf{Sp}(V, \Omega) \cap \mathbf{GL}(V_{\mathbb{C}}) \\ &= \mathbf{O}(V, G) \cap \mathbf{GL}(V_{\mathbb{C}}). \end{aligned}$$

Chapitre IV
Variétés symplectiques

Dans tout ce chapitre, les variétés, applications, formes différentielles considérées seront supposées différentiables de classe C^∞ .

1. Variétés symplectiques

1.1. Définition. — Une forme symplectique sur une variété différentiable M est une 2-forme différentielle Ω fermée (c'est-à-dire vérifiant $d\Omega = 0$) et partout non dégénérée (c'est-à-dire partout de rang égal à la dimension de M). Une variété M munie d'une forme symplectique Ω est appelée variété symplectique et notée (M, Ω) .

1.2. Propriétés élémentaires

a) Soit (M, Ω) une variété symplectique. Pour tout point x de M , l'espace vectoriel tangent en x , $T_x M$, muni de la forme bilinéaire $\Omega(x)$, est un espace vectoriel symplectique. Par suite, d'après 1.4.a, la dimension de M est paire.

b) L'espace \mathbb{R}^{2n} (coordonnées x^1, \dots, x^{2n}), muni de la 2-forme

$$\Omega = \sum_{i=1}^n dx^{n+i} \wedge dx^i$$

est une variété symplectique.

c) Plus généralement, un espace vectoriel symplectique (V, Ω) peut être considéré comme une variété symplectique, Ω étant considérée comme une 2-forme différentielle sur V . En effet, à toute base de V correspond une carte, dans laquelle Ω a des composantes constantes, ce qui prouve que $d\Omega = 0$.

d) Soit (M, Ω) une variété symplectique de dimension $2n$. La $2n$ -forme Ω^n est une forme élément de volume sur M . Par suite, M est orientable.

e) Sur une variété différentiable de dimension 2, toute 2-forme différentielle est automatiquement fermée. Par suite, toute variété différentiable de dimension 2 orientable peut être munie d'une structure symplectique.

f) Soient (M_1, Ω_1) et (M_2, Ω_2) deux variétés symplectiques. On note $p_1 : M_1 \times M_2 \rightarrow M_1$ et $p_2 : M_1 \times M_2 \rightarrow M_2$ les projections de $M_1 \times M_2$ sur ses deux

facteurs. Les 2-formes $p_1^*\Omega_1 + p_2^*\Omega_2$ et $p_1^*\Omega_1 - p_2^*\Omega_2$ sur la variété produit $M_1 \times M_2$, sont toutes deux des formes symplectiques, qu'on notera simplement $\Omega_1 + \Omega_2$ et $\Omega_1 - \Omega_2$, pour alléger l'écriture.

2. La forme de Liouville sur un fibré cotangent

On va voir que le fibré cotangent à une variété différentiable possède une structure symplectique naturelle.

2.1. Quelques notations. — Soit N une variété différentiable, et T^*N son fibré cotangent. On note $q : T^*N \rightarrow N$ la projection canonique, $Tq : T(T^*N) \rightarrow TN$ le prolongement de q aux vecteurs, $p_N : TN \rightarrow N$ et $p_{T^*N} : T(T^*N) \rightarrow T^*N$ les projections canoniques des fibrés tangents, respectivement à N et à T^*N , sur leurs bases. On a le diagramme commutatif

$$\begin{array}{ccc} T(T^*N) & \xrightarrow{Tq} & TN \\ \downarrow p_{T^*N} & & \downarrow p_N \\ T^*N & \xrightarrow{q} & N \end{array} \quad (1)$$

2.2. Définition. — Avec les notations précisées ci-dessus, on appelle *forme de Liouville sur le fibré cotangent* T^*N la 1-forme α sur T^*N définie par la formule, dans laquelle z est un point de T^*N et v un vecteur tangent en z à T^*N :

$$\langle \alpha, v \rangle = \langle p_{T^*N}(v), Tq(v) \rangle = \langle z, Tq(v) \rangle.$$

2.3. Commentaire

a) La formule ci-dessus définit bien une 1-forme différentielle sur T^*N , car d'après la commutativité du diagramme (1), $z = p_{T^*N}(v)$ et $Tq(v)$ sont éléments, respectivement, de l'espace T_x^*N , cotangent en $x = q(z)$ à N , et T_xN , tangent en x à N ; or ces deux espaces vectoriels sont en dualité. On voit ainsi que $\langle \alpha, v \rangle$ est bien défini, et que sur chaque fibre $T_z(T^*N)$, $v \mapsto \langle \alpha, v \rangle$ est linéaire.

b) Considérons une carte de N dont les coordonnées locales sont notées x^1, \dots, x^n ; notons $x^1, \dots, x^n, p_1, \dots, p_n$ les coordonnées locales dans la carte associée de T^*N . La forme de Liouville α a pour expression dans cette carte

$$\alpha = \sum_{i=1}^n p_i dx^i.$$

2.4. Proposition. — La différentielle extérieure $d\alpha$ de la forme de Liouville α est une forme symplectique sur T^*N , dite *forme symplectique canonique*.

Preuve : On a évidemment

$$d(d\alpha) = 0$$

puisque $d \circ d = 0$. De plus, $d\alpha$ est non dégénérée. Son expression locale, dans la carte considérée en 2.3.b, est en effet

$$d\alpha = \sum_{i=1}^n dp_i \wedge dx^i. \quad \square$$

2.5. Proposition. — Soit β une 1-forme différentielle sur la variété N . Puisque β est une application différentiable de N dans son fibré cotangent T^*N , qui vérifie $q \circ \beta = \text{id}_N$ ($q : T^*N \rightarrow N$ désignant la projection canonique), on peut considérer l'image réciproque $\beta^*\alpha$ de la forme de Liouville α sur T^*N . On a :

$$\beta^*\alpha = \beta.$$

Preuve : Soit $x \in N$ et $v \in T_xN$. On a par définition d'une forme image réciproque et d'après la définition de α :

$$\begin{aligned} \langle \beta^*\alpha(x), v \rangle &= \langle \alpha(\beta(x)), T\beta(v) \rangle \\ &= \langle \beta(x), Tq \circ T\beta(v) \rangle \\ &= \langle \beta(x), T(q \circ \beta)(v) \rangle \\ &= \langle \beta(x), v \rangle, \end{aligned}$$

car $q \circ \beta = \text{id}_N$, donc $T(q \circ \beta) = \text{id}_{TN}$. On a donc bien

$$\beta^*\alpha = \beta. \quad \square$$

3. Sous-variétés remarquables d'une variété symplectique

3.1. Définitions. — Soit (M, Ω) une variété symplectique et N une sous-variété de M . On dit que N est une sous-variété isotrope (resp., coïsothrope, resp., lagrangienne, resp., symplectique) si pour tout $x \in N$, T_xN est un sous-espace vectoriel isotrope (resp., coïsothrope, resp., lagrangien, resp., symplectique) de $(T_xM, \Omega(x))$.

3.2. Quelques exemples. — Soit N une variété différentiable, et T^*N son fibré cotangent, muni de la forme symplectique $\Omega = d\alpha$, différentielle extérieure de sa forme de Liouville α .

a) Pour tout $x \in N$, la fibre T_x^*N est une sous-variété lagrangienne de T^*N . Cela résulte en effet immédiatement de l'expression locale de Ω , dans la carte de T^*N , de coordonnées locales $x^1, \dots, x^n, p_1, \dots, p_n$, associée à une carte de N de coordonnées locales x^1, \dots, x^n :

$$\Omega = \sum_{i=1}^n dp_i \wedge dx^i.$$

b) La section nulle de T^*N (image de la forme nulle) est une sous-variété lagrangienne de T^*N . Cela résulte aussi de l'expression locale de Ω indiquée ci-dessus.

c) Plus généralement, soit S une sous-variété de N . On note T_S^*N la restriction à S du fibré cotangent à N , c'est-à-dire le fibré de base S dont la fibre, en chaque point x de S , est l'espace cotangent T_x^*N . On appelle *fibré conormal* à S et on note $(TS)^0$ le sous-fibré de T_S^*N dont la fibre, en chaque point x de S , est l'annulateur de l'espace T_xS tangent en x à S , c'est-à-dire l'ensemble des $\eta \in T_x^*N$ telles que $\langle \eta, v \rangle = 0$ pour tout $v \in T_xS$. On montrera, à titre d'exercice, que $(TS)^0$ est une sous-variété lagrangienne de T^*N . [Indications : On pourra, au voisinage de chaque point de S , choisir une carte de N , de domaine U , adaptée à la sous-variété S , c'est-à-dire une carte dont les coordonnées locales x^1, \dots, x^n sont telles que $S \cap U$ soit l'ensemble des points dont les coordonnées vérifient $x^i = 0$ pour $s+1 \leq i \leq n$, où s désigne la dimension de S . On considèrera alors la carte de T^*N associée à une telle carte de N , et on cherchera comment sont caractérisés, au moyen de leurs coordonnées locales, les points de $(TS)^0$ contenus dans le domaine de cette carte.] On remarquera que lorsque S est réduite à un point x de N , on retrouve le cas a, et que lorsque $S = N$, on retrouve le cas b.

La proposition ci-dessous donne un autre exemple important de sous-variété lagrangienne.

3.3. Proposition. — *Soit N une variété différentiable, T^*N son fibré cotangent, muni de sa 2-forme symplectique canonique $\Omega = d\alpha$. Soit β une forme différentielle sur N . L'image de β est une sous-variété lagrangienne de T^*N si et seulement si β est fermée, c'est-à-dire vérifie $d\beta = 0$.*

Preuve : Soit n la dimension de N . Le fibré cotangent T^*N est donc de dimension $2n$. L'image $\beta(N)$ de β est une sous-variété de T^*N de dimension n . Elle est donc lagrangienne si et seulement si elle est isotrope, c'est-à-dire si et seulement si la 2-forme induite par Ω sur $\beta(N)$ est nulle. Comme β est un difféomorphisme de N sur $\beta(N)$, dont l'inverse est la restriction à $\beta(N)$ de la projection canonique $q : T^*N \rightarrow N$, la 2-forme induite par Ω sur $\beta(N)$ est nulle si et seulement si son image réciproque par β est nulle. Or celle-ci n'est autre que

$$\beta^*(\Omega) = \beta^*(d\alpha) = d(\beta^*\alpha) = d\beta,$$

compte tenu de la proposition 2.5, et du fait que la différentiation extérieure d commute avec la prise d'image réciproque. \square

4. Symplectomorphismes

4.1. Définition. — Soient (M_1, Ω_1) et (M_2, Ω_2) deux variétés symplectiques de même dimension, et $\varphi : M_1 \rightarrow M_2$ une application différentiable.

1. On dit que φ est un *symplectomorphisme local* si

$$\varphi^* \Omega_2 = \Omega_1 .$$

2. On dit que φ est un *symplectomorphisme* si c'est à la fois un symplectomorphisme local et un difféomorphisme.

4.2. Remarques

a) Un symplectomorphisme local est un difféomorphisme local. En effet, si $\varphi : M_1 \rightarrow M_2$ est un symplectomorphisme local, on a

$$\varphi^* \Omega_2 = \Omega_1 , \quad \text{donc aussi} \quad \varphi^*(\Omega_2^n) = \Omega_1^n .$$

Comme Ω_2^n et Ω_1^n sont des formes éléments de volume, respectivement sur M_2 et M_1 , cela prouve que φ est de rang $2n$ en tout point de M_1 , donc est un difféomorphisme local.

b) Si $\varphi : M_1 \rightarrow M_2$ est un symplectomorphisme, son inverse $\varphi^{-1} : M_2 \rightarrow M_1$ est aussi un symplectomorphisme.

4.3. Proposition. — Soient N_1 et N_2 deux variétés de même dimension n , et $\varphi : N_1 \rightarrow N_2$ un difféomorphisme. On note $\widehat{\varphi} : T^*N_1 \rightarrow T^*N_2$ le relèvement de φ aux fibrés cotangents, ainsi défini : pour tout $x \in N_1$, la restriction de $\widehat{\varphi}$ à la fibre $T_x^*N_1$ est la transposée de $T_{\varphi(x)}(\varphi^{-1}) : T_{\varphi(x)}N_2 \rightarrow T_xN_1$. En désignant par α_1 et α_2 les formes de Liouville de T^*N_1 et T^*N_2 , on a

$$\widehat{\varphi}^* \alpha_2 = \alpha_1 , \quad \widehat{\varphi}^*(d\alpha_2) = d\alpha_1 ,$$

donc $\widehat{\varphi}$ est un symplectomorphisme de $(T^*N_1, d\alpha_1)$ sur $(T^*N_2, d\alpha_2)$.

Preuve : On a, pour tout $z \in T^*N_1$ et tout $v \in T_z(T^*N_1)$:

$$\begin{aligned} \langle \widehat{\varphi}^* \alpha_2(z), v \rangle &= \langle \alpha_2(\widehat{\varphi}(z)), T\widehat{\varphi}(v) \rangle \\ &= \langle \widehat{\varphi}(z), Tq_{N_2} \circ T\widehat{\varphi}(v) \rangle \\ &= \langle z, T(\varphi^{-1}) \circ Tq_{N_2} \circ T\widehat{\varphi}(v) \rangle . \end{aligned}$$

Mais $q_{N_2} \circ \widehat{\varphi} = \varphi \circ q_{N_1}$, donc $Tq_{N_2} \circ T\widehat{\varphi} = T\varphi \circ Tq_{N_1}$, ce qui permet d'écrire

$$\begin{aligned} \langle \widehat{\varphi}^* \alpha_2(z), v \rangle &= \langle z, T(\varphi^{-1}) \circ T\varphi \circ Tq_{N_1}(v) \rangle \\ &= \langle z, Tq_{N_1}(v) \rangle \\ &= \langle \alpha_1(z), v \rangle . \end{aligned}$$

On a donc bien $\widehat{\varphi}^* \alpha_2 = \alpha_1$. Comme la différentiation extérieure commute avec la prise d'images réciproques, on en déduit

$$\widehat{\varphi}^*(d\alpha_2) = d\alpha_1,$$

ce qui prouve que $\widehat{\varphi}$ est un symplectomorphisme local. Comme c'est un difféomorphisme, c'est en fait un symplectomorphisme. \square

4.4. Proposition. — Soient (M_1, Ω_1) et (M_2, Ω_2) deux variétés symplectiques de même dimension. Une application différentiable $\varphi : M_1 \rightarrow M_2$ est un symplectomorphisme local si et seulement si son graphe

$$\text{Gr}(\varphi) = \{ (x, y) \in M_1 \times M_2 \mid y = \varphi(x) \}$$

est une sous-variété lagrangienne de $(M_1 \times M_2, \Omega_1 - \Omega_2)$ (avec les notations de 1.2.f).

Preuve : On laisse cette démonstration à faire comme exercice, en donnant seulement quelques indications. Remarquer que le graphe de φ est une sous-variété de dimension $2n$ de $M_1 \times M_2$, de dimension $4n$, et que l'application $x \mapsto (x, \varphi(x))$ est un difféomorphisme de M_1 sur le graphe de φ . Par suite, le graphe de φ est une sous-variété lagrangienne de $(M_1 \times M_2, \Omega_1 - \Omega_2)$ si et seulement si l'image réciproque de $\Omega_1 - \Omega_2$ par l'application $x \mapsto (x, \varphi(x))$ est la forme identiquement nulle sur M_1 . \square

4.5. Remarque. — Soient (M_1, Ω_1) et (M_2, Ω_2) deux variétés symplectiques, pas nécessairement de même dimension. La proposition précédente suggère de considérer les sous-variétés lagrangiennes du produit $M_1 \times M_2$, muni de la forme symplectique $\Omega_1 - \Omega_2$. W. M. Tulczyjew et S. Benenti ont développé l'étude de ces sous-variétés, qu'ils ont appelé *relations symplectiques*, et ont montré qu'elles jouent un rôle important dans de nombreux problèmes de géométrie symplectique et de mécanique.

5. Champs de vecteurs hamiltoniens

5.1. Rappel. — Soit X un champ de vecteurs différentiable sur une variété différentiable M . On appelle *flot réduit* de X l'application Φ , définie sur un ouvert de $\mathbb{R} \times M$ contenant $\{0\} \times M$, à valeurs dans M , telle que pour tout $x \in M$ fixé, $t \mapsto \varphi(t) = \Phi(t, x)$ soit la courbe intégrale maximale de l'équation différentielle

$$\frac{d\varphi(t)}{dt} = X(\varphi(t))$$

vérifiant la donnée de Cauchy

$$\varphi(0) = x.$$

5.2. Proposition. — Soit (M, Ω) une variété symplectique, X un champ de vecteurs différentiable sur M , et Φ son flot réduit. Les trois propriétés suivantes sont équivalentes.

1. Pour tout $t \in \mathbb{R}$, $\Phi_t : x \mapsto \Phi_t(x) = \Phi(t, x)$ est un symplectomorphisme de l'ouvert de (M, Ω) sur lequel il est défini, sur son image.
2. Le champ de vecteurs X vérifie

$$\mathcal{L}(X)\Omega = 0,$$

où $\mathcal{L}(X)$ désigne la dérivée de Lie selon X .

3. La 1-forme différentielle $i(X)\Omega$, produit intérieur de Ω par X , est fermée :

$$di(X)\Omega = 0.$$

Preuve : Le lecteur pourra se reporter, si nécessaire, à l'excellent livre de Claude Godbillon "Géométrie et mécanique" pour approfondir ses connaissances sur la dérivée de Lie. On rappelle la formule

$$\mathcal{L}(X) = i(X)d + di(X),$$

qui nous donne ici

$$\mathcal{L}(X)\Omega = i(X)d\Omega + di(X)\Omega = di(X)\Omega$$

puisque $d\Omega = 0$. Par suite, 2 et 3 sont équivalentes.

D'autre part, on rappelle la formule, valable pour toute forme différentielle η ,

$$\frac{d}{d\theta}(\Phi_\theta^*\eta)|_{\theta=t} = \Phi_t^*(\mathcal{L}(X)\eta).$$

Supposons la propriété 1 vérifiée. Puisque pour tout $\theta \in \mathbb{R}$, $\Phi_\theta^*\Omega = \Omega$, la formule ci-dessus montre que $\mathcal{L}(X)\Omega = 0$. Réciproquement, si $\mathcal{L}(X)\Omega = 0$, cette même formule montre que la dérivée de $\Phi_\theta^*\Omega$ par rapport à θ en $\theta = t$ est nulle, quel que soit t . Par suite, $\Phi_t^*\Omega$ ne dépend pas de t , et comme $\Phi_0^*\Omega = \Omega$, on en déduit que tous les Φ_t sont des symplectomorphismes. \square

5.3. Définitions. — Soit (M, Ω) une variété symplectique. Un champ de vecteurs différentiable X sur M est dit

- localement hamiltonien s'il vérifie les trois propriétés équivalentes de la proposition 5.2,
- globalement hamiltonien, ou s'il n'y a pas de risque de confusion, hamiltonien, s'il existe une fonction différentiable f telle que

$$i(X)\Omega = -df.$$

La fonction f est alors appelée un hamiltonien associé au champ de vecteurs hamiltonien X .

Plus généralement, soit X un champ de vecteurs dépendant du temps sur M , c'est-à-dire une application différentiable X d'un ouvert de $\mathbb{R} \times M$ dans TM telle que, pour tout élément (t, x) de son domaine de définition, $X(t, x)$ soit élément de $T_x M$. On dit que ce champ de vecteurs est

- localement hamiltonien dépendant du temps si, pour chaque $t \in \mathbb{R}$ fixé, le champ de vecteurs X_t , (défini sur un ouvert de M dépendant de t), $x \mapsto X_t(x) = X(t, x)$, est localement hamiltonien,
- globalement hamiltonien dépendant du temps, ou s'il n'y a pas de risque de confusion, hamiltonien dépendant du temps, s'il existe une fonction différentiable f , définie sur le même ouvert de $\mathbb{R} \times M$ que X , telle que pour tout $t \in \mathbb{R}$ on ait, sur l'ouvert de M sur lequel les deux membres de cette égalité sont définis,

$$i(X_t)\Omega = -df_t,$$

où on a noté f_t la fonction, définie sur un ouvert de M dépendant de t , $x \mapsto f_t(x) = f(t, x)$. La fonction f est alors appelée un hamiltonien dépendant du temps associé au champ de vecteurs hamiltonien dépendant du temps X .

On appelle système hamiltonien (resp., système hamiltonien dépendant du temps), un triplet (M, Ω, H) où (M, Ω) est une variété symplectique et H une fonction différentiable définie sur M (resp., sur un ouvert de $\mathbb{R} \times M$). On appelle champ de vecteurs hamiltonien associé à H le champ de vecteurs hamiltonien X_H , éventuellement dépendant du temps, défini par

$$i(X_H)\Omega = -dH$$

dans le cas où H ne dépend pas du temps, et par

$$i(X_{H_t})\Omega = dH_t, \quad t \in \mathbb{R},$$

dans le cas où H dépend de t . On a noté, pour chaque $t \in \mathbb{R}$ fixé, X_{H_t} le champ de vecteurs $x \mapsto X_{H_t}(x) = X_H(t, x)$, et H_t la fonction $x \mapsto H_t(x) = H(t, x)$. On appelle courbes intégrales ou solutions du système hamiltonien (éventuellement dépendant du temps) (M, Ω, H) , les solutions de l'équation différentielle

$$\varphi'(t) = X_H(t, \varphi(t)), \quad \text{ou} \quad \varphi'(t) = X_{H_t}(\varphi(t)),$$

selon que le système hamiltonien considéré est ou n'est pas dépendant du temps.

5.4. Remarques

a) Un champ de vecteurs hamiltonien est automatiquement localement hamiltonien, puisque $i(X)\Omega = -df$ implique $d i(X)\Omega = -d(df) = 0$.

b) Réciproquement, sur une variété simplement connexe, ou plus généralement sur une variété M telle que $H^1(M) = 0$, tout champ de vecteurs localement hamiltonien est hamiltonien; en effet, si la 1-forme différentielle $i(X)\Omega$ est fermée et si $H^1(M) = 0$, il existe une fonction différentiable f telle que $i(X)\Omega = -df$.

5.5. Expression locale. — Soit N une variété différentiable de dimension n , T^*N son fibré cotangent, muni de sa forme symplectique canonique $\Omega = d\alpha$, α étant sa 1-forme de Liouville. Soit $H : T^*N \rightarrow \mathbb{R}$ une fonction différentiable, et X_H le champ de vecteurs hamiltonien associé, défini par

$$i(X_H)\Omega = -dH.$$

On considère une carte de N , dont les coordonnées locales sont notées x^1, \dots, x^n , et la carte associée de T^*N , dont on note $x^1, \dots, x^n, p_1, \dots, p_n$ les coordonnées locales. On rappelle que dans ces coordonnées, la forme de Liouville et sa différentielle extérieure ont pour expressions locales

$$\alpha = \sum_{i=1}^n p_i dx^i, \quad \Omega = d\alpha = \sum_{i=1}^n dp_i \wedge dx^i.$$

La différentielle de H ayant pour expression locale

$$dH = \sum_{i=1}^n \left(\frac{\partial H}{\partial x^i} dx^i + \frac{\partial H}{\partial p_i} dp_i \right),$$

on en déduit, par identification, l'expression locale de X_H :

$$X_H = \sum_{i=1}^n \left(\frac{\partial H}{\partial p_i} \frac{\partial}{\partial x^i} - \frac{\partial H}{\partial x^i} \frac{\partial}{\partial p_i} \right). \quad (*)$$

L'équation différentielle

$$\frac{d\varphi(t)}{dt} = X_H(\varphi(t))$$

s'écrit donc, en coordonnées,

$$\begin{cases} \frac{dx^i}{dt} = \frac{\partial H}{\partial p_i}, \\ \frac{dp_i}{dt} = -\frac{\partial H}{\partial x^i}. \end{cases} \quad (**)$$

Ces équations sont appelées *équations de Hamilton*.

Ces expressions locales ne sont pas particulières au cas où la variété symplectique considérée est un fibré cotangent. On verra en effet plus loin que toute variété symplectique de dimension $2n$ possède un atlas de cartes, dites *cartes canoniques*, dans lesquelles la forme symplectique Ω a pour expression locale, au moyen des coordonnées locales, notées $x^1, \dots, x^n, p_1, \dots, p_n$,

$$\Omega = \sum_{i=1}^n dp_i \wedge dx^i.$$

Dans une telle carte, le champ de vecteurs hamiltonien X_H associé à une fonction différentiable H a pour expression locale (*), et l'équation différentielle de Hamilton a pour expression (**).

5.6. Proposition. — Soient X et Y deux champs de vecteurs localement hamiltoniens sur une variété symplectique (M, Ω) . Leur crochet $[X, Y]$ est globalement hamiltonien, et a pour hamiltonien $i(Y)i(X)\Omega = \Omega(X, Y)$.

Preuve : Rappelons une formule très utile. Soit η une forme différentielle de degré quelconque p , X et Y deux champs de vecteurs, tous définis sur une variété M . On a

$$\mathcal{L}(X)i(Y)\eta = i([X, Y])\eta + i(Y)\mathcal{L}(X)\eta.$$

Appliquons cette formule dans le cas présent, avec $\eta = \Omega$. On obtient

$$i([X, Y])\Omega = \mathcal{L}(X)i(Y)\Omega$$

car $\mathcal{L}(X)\Omega = 0$, le champ X étant localement hamiltonien. D'autre part

$$\mathcal{L}(X)i(Y)\Omega = (i(X)d + di(X))i(Y)\Omega = di(X)i(Y)\Omega = -d(\Omega(X, Y)),$$

car $di(Y)\Omega = 0$. □

5.7. Exercice. — Soit (M, Ω) une variété symplectique. On note $\Omega^b : TM \rightarrow T^*M$ l'application

$$v \mapsto -i(v)\Omega$$

et $\Omega^\sharp : T^*M \rightarrow TM$ l'inverse de Ω^b . On munit TM de la 2-forme

$$\Omega_{TM} = (\Omega^b)^*(d\alpha),$$

image réciproque par Ω^b de la forme $d\alpha$ sur T^*M (α étant la forme de Liouville).

1. Montrer que (TM, Ω_{TM}) est une variété symplectique. Donner l'expression locale de Ω_{TM} , dans une carte de TM associée à une carte canonique de M , c'est-à-dire dans laquelle Ω a pour expression (les coordonnées locales sur M étant notées x^1, \dots, x^{2n}),

$$\Omega = \sum_{i=1}^n dx^{n+i} \wedge dx^i.$$

2. Soit X un champ de vecteurs différentiable sur M . Montrer que X est localement hamiltonien si et seulement si son image $X(M)$ est une sous-variété lagrangienne de (TM, Ω_{TM}) . Ce résultat est dû à W. M. Tulczyjew.

6. Le crochet de Poisson

6.1. Définition. — Soit (M, Ω) une variété symplectique. À tout couple de fonctions différentiables (f, g) sur M , on associe la fonction

$$\{f, g\} = i(X_f)dg,$$

où X_f est le champ de vecteurs hamiltonien associé à f , défini par $i(X_f)\Omega = -df$. On dit que $\{f, g\}$ est le crochet de Poisson des fonctions f et g .

6.2. Proposition. — *Le crochet de Poisson a les propriétés suivantes.*

1. On a, pour tout couple de fonctions différentiables (f, g) sur M , en notant X_f et X_g les champs de vecteurs hamiltoniens associés,

$$\{f, g\} = i(X_f)dg = -i(X_g)df = i(X_g)i(X_f)\Omega = \Omega(X_f, X_g).$$

2. Le crochet de Poisson est une application bilinéaire de $C^\infty(M, \mathbb{R}) \times C^\infty(M, \mathbb{R})$ dans $C^\infty(M, \mathbb{R})$.

3. Le crochet de Poisson est antisymétrique,

$$\{g, f\} = -\{f, g\}.$$

4. Le crochet de Poisson vérifie la formule de Leibniz :

$$\{f, g_1g_2\} = \{f, g_1\}g_2 + g_1\{f, g_2\}, \quad f, g_1, g_2 \in C^\infty(M, \mathbb{R}).$$

5. Le champ hamiltonien $X_{\{f, g\}}$ ayant pour hamiltonien le crochet de Poisson $\{f, g\}$ de deux fonctions f et g , est le crochet des champs hamiltoniens X_f et X_g associés à ces fonctions :

$$X_{\{f, g\}} = [X_f, X_g].$$

6. Le crochet de Poisson vérifie l'identité de Jacobi

$$\{\{f, g\}, h\} + \{\{g, h\}, f\} + \{\{h, f\}, g\} = 0, \quad f, g, h \in C^\infty(M, \mathbb{R}).$$

Preuve : Les propriétés 1, 2, 3 et 4 sont des conséquences immédiates de la définition. La propriété 5 est conséquence de la proposition 5.6. Montrons 6. On a, compte tenu de la propriété 5 (en notant $Z.h = i(Z)dh = \langle dh, Z \rangle$ la dérivée d'une fonction h dans la direction d'un champ de vecteurs Z) :

$$\begin{aligned} \{\{f, g\}, h\} &= [X_f, X_g].h, \\ \{\{g, h\}, f\} &= -\{f, \{g, h\}\} = -X_f.(X_g.h), \\ \{\{h, f\}, g\} &= \{g, \{f, h\}\} = X_g.(X_f.h). \end{aligned}$$

En ajoutant, on obtient

$$\{\{f, g\}, h\} + \{\{g, h\}, f\} + \{\{h, f\}, g\} = 0,$$

en vertu de la formule bien connue applicable à tout couple de champs de vecteurs (Y, Z) et à toute fonction différentiable h :

$$[Y, Z].h = Y.(Z.h) - Z.(Y.h). \quad \square$$

6.3. Expression locale du crochet de Poisson en coordonnées canoniques

Soit (M, Ω) une variété symplectique. Ainsi qu'on l'a déjà annoncé, il existe un atlas de M formé de cartes canoniques, c'est-à-dire telles que la 2-forme Ω ait pour expression locale, au moyen des coordonnées locales x^1, \dots, x^{2n} ,

$$\Omega = \sum_{i=1}^{2n} dx^{n+i} \wedge dx^i.$$

Dans une telle carte, le crochet de Poisson de deux fonctions différentiables a pour expression locale

$$\{f, g\} = \sum_{i=1}^n \left(\frac{\partial f}{\partial x^{n+i}} \frac{\partial g}{\partial x^i} - \frac{\partial f}{\partial x^i} \frac{\partial g}{\partial x^{n+i}} \right).$$

7. Application : intégrales premières d'un système hamiltonien

7.1. Quelques rappels. — Soit X un champ de vecteurs différentiable sur une variété différentiable M . On considère l'équation différentielle qui lui est associée :

$$\frac{d\varphi(t)}{dt} = X(\varphi(t)). \quad (*)$$

On appelle *intégrale première* de cette équation, ou du champ de vecteurs X , toute fonction f , définie sur M , dont la valeur reste constante sur chaque courbe intégrale de l'équation (*), c'est-à-dire telle que, pour toute solution $\varphi : I \rightarrow M$ de l'équation (*), $t \mapsto f(\varphi(t))$ soit constante sur l'intervalle I .

Lorsque f est une fonction différentiable et $\varphi : I \rightarrow M$ une solution de l'équation différentielle (*), l'application composée $f \circ \varphi : t \mapsto f(\varphi(t))$ est dérivable en tout point $t \in I$, et a pour dérivée

$$\frac{d(f \circ \varphi)(t)}{dt} = \left\langle df(\varphi(t)), \frac{d\varphi(t)}{dt} \right\rangle = \langle df(\varphi(t)), X(\varphi(t)) \rangle.$$

Si la fonction différentiable f est une intégrale première de l'équation (1), pour toute solution $\varphi : I \rightarrow M$ de cette équation et tout $t \in I$, on a

$$\langle df(\varphi(t)), X(\varphi(t)) \rangle = 0.$$

Mais d'après le théorème de Cauchy-Lipschitz, pour tout point $x \in M$, il existe une solution φ de l'équation différentielle (*) définie sur un intervalle ouvert contenant l'origine et vérifiant $\varphi(0) = x$. Nous pouvons donc affirmer que si la fonction différentiable f est intégrale première de l'équation différentielle (*), elle vérifie, pour tout point $x \in M$,

$$\langle df(x), X(x) \rangle = 0. \quad (**)$$

Réciproquement, une fonction différentiable f , définie sur M , qui pour tout $x \in M$ vérifie (**), est une intégrale première de l'équation différentielle (*), car chaque

solution de cette équation est définie sur un intervalle, partie connexe de \mathbb{R} ; l'annulation de la dérivée de la fonction $t \mapsto f \circ \varphi(t)$ en tout point t de son intervalle de définition implique la constance de cette fonction.

Soit maintenant (M, Ω, H) un système hamiltonien. Nous allons nous intéresser aux intégrales premières du champ de vecteurs hamiltonien X_H , aussi appelées intégrales premières du système (M, Ω, H) .

7.2. Proposition. — *Soit (M, Ω, H) un système hamiltonien. Une fonction différentiable $f : M \rightarrow \mathbb{R}$ est intégrale première de ce système si et seulement si elle est en involution avec H , c'est-à-dire si et seulement si elle vérifie*

$$\{H, f\} = 0.$$

Preuve : D'après les rappels présentés ci-dessus, la fonction f est intégrale première du système (M, Ω, H) si et seulement si, pour tout $x \in M$,

$$\langle df(x), X_H(x) \rangle = 0.$$

Mais nous avons vu que pour tout $x \in M$,

$$\langle df(x), X_H(x) \rangle = (i(X_H)df)(x) = \{H, f\}(x).$$

Le résultat annoncé en découle immédiatement. □

7.3. Proposition. — *Soient (M, Ω, H) un système hamiltonien, f et g deux fonctions différentiables toutes deux intégrales premières de ce système. Le crochet de Poisson $\{f, g\}$ de ces deux fonctions est aussi intégrale première du système.*

Preuve : D'après l'identité de Jacobi, nous avons

$$\{H, \{f, g\}\} = \{\{H, f\}, g\} + \{f, \{H, g\}\}.$$

Le résultat annoncé en découle immédiatement. □

8. Réduction d'une variété symplectique

8.1. Quelques rappels sur les feuilletages. — Soit N une variété différentiable de dimension n , et k un entier vérifiant $1 \leq k \leq n$.

On appelle *champ de k -directions* (ou parfois aussi *distribution* de rang k) sur N un sous-ensemble \mathcal{F} du fibré tangent TN tel que, pour tout point $x \in N$, $\mathcal{F} \cap T_x N$ soit un sous-espace vectoriel de dimension k de $T_x N$. Ce sous-espace $\mathcal{F} \cap T_x N$ est noté \mathcal{F}_x , et appelé *fibres* de \mathcal{F} au point x .

Un champ de k -directions \mathcal{F} est dit *différentiable* (sous-entendu : de classe C^∞) si tout point x de N possède un voisinage U sur lequel existent k champs de vecteurs différentiables (de classe C^∞) X_i , $1 \leq i \leq k$, dont les valeurs, en tout point $y \in U$, forment une base de la fibre \mathcal{F}_y .

On appelle *intégrale* d'un champ de k -directions différentiable \mathcal{F} , sur la variété différentiable N , une immersion $f : S \rightarrow N$ d'une variété différentiable connexe S dans la variété N telle que, pour tout point $x \in S$, $T_x f(T_x S)$ soit un sous-espace vectoriel de la fibre $\mathcal{F}_{f(x)}$. Lorsque c'est le cas, la dimension de S est nécessairement inférieure ou égale à k , car f étant une immersion, $T_x f$ est injective. On appelle *sous-variété intégrale* de \mathcal{F} une sous-variété de N dont l'injection canonique dans N est une intégrale de \mathcal{F} . Bien souvent, on n'exige pas que S soit une sous-variété de N au sens strict, c'est-à-dire une sous-variété plongée (dont la topologie en tant que variété est identique à la topologie induite par celle de N). On demande seulement que S soit une "sous-variété immergée" de N , c'est-à-dire un sous-ensemble de N possédant une structure de variété différentiable telle que son injection canonique dans N soit une immersion. Une sous-variété intégrale de \mathcal{F} est dite *maximale* si toute sous-variété intégrale de \mathcal{F} qui la contient lui est égale.

Un champ de k directions différentiable \mathcal{F} sur la variété N est dit *complètement intégrable* si, pour tout point x de N , il existe une sous-variété intégrale de \mathcal{F} de dimension k passant par le point x . Lorsque c'est le cas, on montre que par tout point de N passe une unique sous-variété intégrale maximale de \mathcal{F} . On dit alors que \mathcal{F} définit un *feuilletage* de N . Les sous-variétés intégrales maximales de \mathcal{F} sont appelées *feuilles* de ce feuilletage. L'appartenance à une même feuille est une relation d'équivalence sur N , dont les classes sont précisément les feuilles. Celles-ci forment donc une partition de N . De plus, une intégrale quelconque $f : S \rightarrow N$ de \mathcal{F} (la dimension de S pouvant être inférieure à k) a nécessairement son image $f(S)$ entièrement contenue dans une feuille, et f , considérée comme application de S dans cette feuille, est une immersion (dans le cas où la dimension de S est k , c'est un difféomorphisme local).

La structure locale d'une variété feuilletée (c'est-à-dire d'une variété différentiable N munie d'un champ de k -directions différentiable et complètement intégrable) est très simple : tout point de N possède un voisinage U qui est le domaine d'une carte dont les coordonnées locales x^1, \dots, x^n sont telles que \mathcal{F} soit localement engendré, au dessus de l'ouvert U , par les k champs de vecteurs $\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^k}$. Les composantes connexes des intersections de U avec les feuilles du feuilletage défini par \mathcal{F} sont les sous-variétés définies par des équations de la forme $x^{k+1} = c^{k+1}, \dots, x^n = c^n$, où c^{k+1}, \dots, c^n sont des constantes.

Un théorème dû à Frobenius donne des conditions nécessaires et suffisantes pour qu'un champ de k -directions différentiable \mathcal{F} soit complètement intégrable. Une des formes les plus souvent utilisées de ces conditions est la suivante : \mathcal{F} est complètement intégrable si et seulement si, pour tout couple (X, Y) de sections locales différentiables de \mathcal{F} , le crochet $[X, Y]$ est une section locale de \mathcal{F} . On

rappelle qu'une section locale différentiable de \mathcal{F} est un champ de vecteurs différentiable X , défini sur un ouvert U de N , et tel que, pour tout $x \in U$, $X(x)$ soit élément de la fibre \mathcal{F}_x . Pour plus de détails sur les diverses formes du théorème de Frobenius, le lecteur pourra consulter le livre de Claude Godbillon déjà cité.

Les notions de complète intégrabilité et de feuilletage s'étendent aux champs de directions différentiables dont le rang k n'est pas nécessairement constant. Il existe des généralisations du théorème de Frobenius (théorèmes de Stefan et Sussmann) adaptées à ce cas, qui se révéleront utiles dans la suite du cours (notamment lors de l'étude des variétés de Poisson). Mais pour le moment nous n'avons à considérer que des feuilletages au sens usuel, dont toutes les feuilles sont de même dimension.

8.2. Proposition. — *Soit (M, Ω) une variété symplectique et N une sous-variété de M telle que la forme $\Omega_N = i_N^* \Omega$, induite par Ω sur N , soit de rang constant. Alors $TN \cap \text{orth}(TN)$ est un sous-fibré vectoriel complètement intégrable de TN .*

Preuve : Pour tout point $x \in N$, on a

$$\ker \Omega_N(x) = T_x N \cap \text{orth}(T_x N).$$

La dimension de $T_x N \cap \text{orth}(T_x N)$ ne dépend donc pas du choix du point x de N . Par suite, $\ker \Omega_N = TN \cap \text{orth}(TN)$ est un sous-fibré vectoriel de TN . Pour montrer qu'il est complètement intégrable il suffit, d'après le théorème de Frobenius, de montrer que si X et Y sont deux sections différentiables de $TN \cap \text{orth}(TN)$, c'est-à-dire deux champs de vecteurs différentiables sur N tels que, pour tout $x \in N$, $X(x)$ et $Y(x)$ soient éléments de $T_x N \cap \text{orth}(T_x N)$, alors $[X, Y]$ est aussi une section de $TN \cap \text{orth}(TN)$. On a

$$\begin{aligned} i([X, Y])\Omega_N &= \mathcal{L}(X)i(Y)\Omega_N - i(Y)\mathcal{L}(X)\Omega_N \\ &= -i(Y)\mathcal{L}(X)\Omega_N \\ &= -i(Y)(i(X)d\Omega_N + di(X)\Omega_N) \\ &= 0, \end{aligned}$$

compte tenu de $d\Omega_N = 0$, $i(Y)\Omega_N = 0$, $i(X)\Omega_N = 0$. Ceci montre que $[X, Y]$ est une section de $\ker \Omega_N = TN \cap \text{orth}(TN)$. \square

8.3. Théorème. — *On fait les mêmes hypothèses que dans la proposition 7.2. On suppose de plus que le feuilletage de N défini par le sous-fibré complètement intégrable $TN \cap \text{orth}(TN)$ de TN est simple, c'est-à-dire que l'ensemble P des feuilles de ce feuilletage est une variété différentiable et que la projection canonique $\pi : N \rightarrow P$ est une submersion. Alors il existe sur P une 2-forme symplectique Ω_P unique telle que*

$$\pi^* \Omega_P = i_N^* \Omega = \Omega_N.$$

Preuve : Notons \mathcal{F} le feuilletage de N défini par $TN \cap \text{orth}(TN)$. S'il existe une 2-forme Ω_P sur P telle que $\pi^*\Omega_P = \Omega_N$, on doit nécessairement avoir, pour tout point $\dot{x} \in P$ et tout couple de vecteurs \dot{v} et \dot{w} éléments de $T_{\dot{x}}P$,

$$\Omega_P(\dot{x})(\dot{v}, \dot{w}) = \Omega_N(x)(v, w),$$

où x est un point quelconque de $\pi^{-1}(\dot{x})$, v et w deux vecteurs éléments de T_xN tels que

$$T\pi(v) = \dot{v}, \quad T\pi(w) = \dot{w}.$$

Cela montre que si elle existe, la 2-forme Ω_P est unique. Pour prouver son existence, il suffit de vérifier que $\Omega_N(x)(v, w)$ ne dépend pas du choix de x dans $\pi^{-1}(\dot{x})$, ni des choix de v et w dans $(T_x\pi)^{-1}(\dot{v})$ et $(T_x\pi)^{-1}(\dot{w})$, respectivement.

Si v et v' appartiennent tous deux à $(T_x\pi)^{-1}(\dot{v})$, $v - v'$ est élément de $T_x(\pi^{-1}(\dot{x})) = \ker \Omega_N(x)$, donc $\Omega_N(x)(v, w) = \Omega_N(x)(v', w)$. Le même raisonnement vaut aussi pour w .

Soient x et x' deux éléments de $\pi^{-1}(\dot{x})$. Chaque feuille du feuilletage \mathcal{F} étant connexe, il existe un champ de vecteurs Z , défini sur un ouvert de N contenant x et x' , à valeurs dans $TN \cap \text{orth}(TN)$, tel que x et x' soient sur la même courbe intégrale de Z . Notons Φ le flot de Z . La courbe intégrale de Z passant par x pour $t = 0$ est $t \mapsto \Phi_t(x) = \Phi(t, x)$. Puisque x' est sur cette courbe intégrale, il existe un réel τ tel que $\Phi_\tau(x) = x'$. Soient v et w deux vecteurs éléments de T_xN , vérifiant

$$T\pi(v) = \dot{v}, \quad T\pi(w) = \dot{w}.$$

Posons

$$v' = T\Phi_\tau(v), \quad w' = T\Phi_\tau(w).$$

Puisque Z est une section de $TN \cap \text{orth}(TN)$, chaque courbe intégrale de ce champ de vecteurs est entièrement contenue dans une feuille du feuilletage \mathcal{F} . Par suite, $\pi \circ \Phi_\tau = \pi$, donc

$$\begin{aligned} T\pi(v') &= T\pi \circ T\Phi_\tau(v) = T\pi(v) = \dot{v}, \\ T\pi(w') &= T\pi \circ T\Phi_\tau(w) = T\pi(w) = \dot{w}. \end{aligned}$$

On a

$$\Omega_N(x')(v', w') = \Omega_N(\Phi_\tau(x))(T\Phi_\tau(v), T\Phi_\tau(w)) = (\Phi_\tau^*(\Omega_N))(x)(v, w).$$

Mais d'après une formule bien connue concernant la dérivée de Lie,

$$\frac{d}{dt}(\Phi_t^*\Omega_N) = \Phi_t^*(\mathcal{L}(Z)\Omega_N) = 0,$$

car

$$\mathcal{L}(Z)\Omega_N = i(Z)d\Omega_N + di(Z)\Omega_N = 0,$$

puisque $d\Omega_N = 0$ et $i(Z)\Omega_N = 0$. On a donc

$$\Phi_\tau^*\Omega_N = \Omega_N,$$

et par suite

$$\Omega_N(x')(v', w') = \Omega_N(x)(v, w).$$

On a donc prouvé l'existence et l'unicité de Ω_P .

L'espace vectoriel $T_{\hat{x}}P$ s'identifie au quotient de T_xN par son sous-espace $T_xN \cap \text{orth}(T_xN)$. On reconnaît, dans la définition de $\Omega_P(\hat{x})$, la construction de la structure d'espace vectoriel symplectique réduit sur ce quotient, étudiée au chapitre I. Ceci prouve que Ω_P est partout de rang égal à la dimension de P .

Enfin on a

$$\pi^*(d\Omega_P) = d(\pi^*\Omega_P) = d\Omega_N = 0,$$

ce qui implique $d\Omega_P = 0$ car π étant une submersion, π^* est injective. \square

8.4. Théorème. — *Dans les hypothèses du théorème 7.3, soit H une fonction différentiable définie sur M , telle que le champ de vecteurs hamiltonien correspondant X_H soit tangent à la sous-variété N (en chacun de ses points). Il existe alors une fonction différentiable unique \hat{H} définie sur P telle que*

$$H|_N = \hat{H} \circ \pi.$$

De plus, soit $X_{\hat{H}}$ le champ de vecteurs hamiltonien sur P associé à la fonction \hat{H} (relativement à la structure symplectique Ω_P). La restriction à N du champ de vecteurs X_H est projectable par π sur P , et a pour projection $X_{\hat{H}}$.

Preuve : On note $\sharp : T^*M \rightarrow TM$ le morphisme de fibrés vectoriels inverse de $\flat : TM \rightarrow T^*M$, $\flat v = -i(v)\Omega$. Par hypothèse, pour tout point x de N , $X_H(x) = \sharp(dH(x))$ est élément de T_xN ; par suite, $dH(x)$ est élément de l'annulateur $(\text{orth } T_xN)^0$ de l'orthogonal de T_xN . Par suite, H est constante sur chaque feuille du feuilletage \mathcal{F} , puisque celui-ci est défini par le sous-fibré $TN \cap \text{orth}(TN)$. Il existe donc bien une fonction $\hat{H} : P \rightarrow \mathbb{R}$ unique telle que $H|_N = \hat{H} \circ \pi$. La fonction \hat{H} est différentiable car la structure de variété de P est la structure quotient de celle de N . On a

$$\Omega_N = \pi^*\Omega_P, \quad H|_N = \pi^*\hat{H}, \quad \text{donc } d(H|_N) = \pi^*(d\hat{H}),$$

d'où l'on déduit

$$\begin{aligned} \pi^* \left(i(T\pi(X_H|_N))\Omega_P \right) &= -i(X_H|_N)\Omega_N \\ &= -d(H|_N) \\ &= -\pi^*(d\hat{H}) \\ &= \pi^*(i(X_{\hat{H}})\Omega_P). \end{aligned}$$

Mais π étant une submersion, et Ω_P étant non dégénérée, on en déduit que pour tout point $x \in N$:

$$T\pi(X_H(x)) = X_{\hat{H}}(\pi(x)). \quad \square$$

8.5. Remarques

a) Le théorème 7.4 facilite la recherche des courbes intégrales du champ de vecteurs X_H contenues dans N , car on peut chercher d'abord leurs projections sur P , qui sont les courbes intégrales de $X_{\widehat{H}}$. Mais dans certaines applications, la construction est utilisée dans l'autre sens : un système hamiltonien sur une variété symplectique (P, Ω_P) étant donné, on cherche à déterminer un autre système hamiltonien, sur une variété symplectique de dimension plus grande (M, Ω) , dont le système donné se déduit par réduction. Il arrive en effet que le système sur (M, Ω) soit plus simple que le système initialement donné, par exemple en raison de symétries plus visibles, bien que la dimension de la variété sur laquelle il est défini soit plus grande. Cette méthode a été utilisée, notamment par Kazhdan, Kostant et Sternberg (Comm. Pure and Appl. Math., 31, 1978, 481–508) et par J. P. Francoise et O. Ragnisco (Ann. Inst. Henri Poincaré, Physique théorique, 49, 3, 1989, 369–375), pour prouver la complète intégrabilité de certains systèmes hamiltoniens.

b) Le théorème 7.4 a été généralisé par A. Lichnerowicz (Journal of Differential Geometry, 12, 1977, 253–300) au cas où le champ de vecteurs hamiltonien X_H n'est pas nécessairement tangent à N en chacun de ces points, mais où la restriction de ce champ de vecteurs à N est une section du fibré vectoriel $TN + \text{orth}(TN)$.

8.6. Exemples de réduction

a) *Réduction d'une variété de niveau du hamiltonien.* — Soit (M, Ω) une variété symplectique et $H : M \rightarrow \mathbf{R}$ une fonction différentiable. Soit $h \in \mathbf{R}$ une valeur régulière de H , c'est-à-dire une valeur telle que $dH(x) \neq 0$ pour tout $x \in H^{-1}(h)$. On sait alors que $N = H^{-1}(h)$ est une sous-variété de M , de codimension 1. Puisque H garde une valeur constante le long de chaque courbe intégrale de X_H , ce champ de vecteurs est tangent à N en chacun de ses points. De plus, X_H ne s'annule en aucun point de N , car dH ne s'annule pas sur N . Pour tout point x de N , l'orthogonal de $T_x N$ est le sous-espace vectoriel de dimension 1 de $T_x M$ engendré par $X_H(x)$. Il est contenu dans $T_x N$. Le feuilletage \mathcal{F} est donc, dans le cas présent, le feuilletage dont les feuilles sont les trajectoires de X_H contenues dans N . Si ce feuilletage est simple, on voit que l'ensemble P de ces trajectoires est une variété munie d'une 2-forme symplectique Ω_P , telle que $\pi^* \Omega_P = \Omega_N$. Le hamiltonien réduit $\widehat{H} : P \rightarrow \mathbf{R}$ est constant, égal à h , et le champ hamiltonien réduit $X_{\widehat{H}}$ est identiquement nul.

b) *L'oscillateur harmonique.* — On prend pour variété symplectique \mathbb{R}^{2n} (coordonnées $x^1, \dots, x^n, y^1, \dots, y^n$), muni de la forme symplectique

$$\Omega = \sum_{k=1}^n dy^k \wedge dx^k.$$

On prend pour hamiltonien

$$H = \frac{1}{2} \sum_{k=1}^n ((x^k)^2 + (y^k)^2).$$

Le champ hamiltonien correspondant est

$$X_H = \sum_{k=1}^n \left(y^k \frac{\partial}{\partial x^k} - x^k \frac{\partial}{\partial y^k} \right).$$

Les équations de Hamilton s'écrivent donc

$$\begin{cases} \frac{dx^k}{dt} = y^k, \\ \frac{dy^k}{dt} = -x^k. \end{cases}$$

On peut écrire ce qui précède en utilisant des coordonnées complexes $z^k = x^k + iy^k$. On note $z = (z^1, \dots, z^n) \in \mathbb{C}^n$, et on munit \mathbb{C}^n de la norme

$$\|z\| = \left(\sum_{k=1}^n |z^k|^2 \right)^{1/2}.$$

Le hamiltonien du système s'écrit alors

$$H = \frac{1}{2} \|z\|^2,$$

et les équations de Hamilton ont pour expression

$$\frac{dz}{dt} = -iz.$$

Leur solution, prenant la valeur $z_0 \in \mathbb{C}^n$ pour $t = 0$, est

$$z(t) = z_0 \exp(it).$$

Prenons par exemple $h = 1/2$. C'est une valeur régulière de H , et $N = H^{-1}(1/2)$ est la sphère S^{2n-1} de \mathbb{C}^n , centrée à l'origine, de rayon 1. Sa dimension (réelle) est $2n - 1$. Les courbes intégrales de X_H situées sur cette sphère sont des grands cercles, orbites de l'action de S^1 (cercle trigonométrique) sur S^{2n-1} :

$$(\exp(i\alpha), z) \mapsto \exp(i\alpha)z.$$

La variété symplectique réduite s'identifie à l'espace projectif complexe $\mathbb{C}\mathbb{P}^{n-1}$, quotient de $\mathbb{C}^n \setminus \{0\}$ par la relation d'équivalence : z équivalent à z' s'il existe $\lambda \in \mathbb{C}$, $\lambda \neq 0$, tel que $z' = \lambda z$. La projection $\pi : S^{2n-1} \rightarrow \mathbb{C}\mathbb{P}^{n-1}$ est appelée *fibration de Hopf*.

8.7. Exercice. — Dans l'exemple ci-dessus, déterminer l'expression de la forme symplectique de la variété symplectique réduite $\mathbb{C}\mathbb{P}^{n-1}$, dans la carte ayant pour domaine la projection de l'ouvert de $\mathbb{C}^n \setminus \{0\}$ constitué par les $z = (z^1, \dots, z^n)$ tels que $z^1 \neq 0$.

9. Les théorèmes de Darboux et de Jacobi-Lie-Carathéodory

Les deux théorèmes exposés dans ce paragraphe ont de très importantes conséquences. Le premier (théorème de Darboux) montre que la structure locale d'une variété symplectique est identique à celle d'un espace vectoriel symplectique. Le second (théorème de Jacobi-Lie-Carathéodory) est à rapprocher de la proposition 2.8 du chapitre III. Il montre que sur une variété symplectique de dimension $2n$ on peut trouver, au voisinage de chaque point, un système de coordonnées canoniques dont les n premières sont imposées d'avance, pourvu qu'elles soient deux à deux en involution et que leurs différentielles soient linéairement indépendantes.

9.1. Théorème de Darboux. — Soit (M, Ω) une variété symplectique de dimension $2n$. Tout point de M possède un voisinage qui est le domaine d'une carte canonique. Cela signifie que dans les coordonnées locales x^1, \dots, x^{2n} associées à cette carte, la forme symplectique Ω a pour expression locale

$$\Omega = \sum_{i=1}^n dx^{n+i} \wedge dx^i.$$

Preuve : Pour tout entier $k \geq 1$, appelons H_k l'assertion suivante :

Dans une variété symplectique de dimension $\leq 2k$, tout point possède un voisinage qui est le domaine d'une carte canonique.

Montrons que l'assertion H_1 est vraie. Soit donc (M, Ω) une variété symplectique de dimension 2 et a un point de M . Soient x et y les coordonnées locales dans une carte de M dont le domaine contient a . Dans ces coordonnées locales, Ω a pour expression

$$\Omega = f(x, y) dx \wedge dy,$$

où f est une fonction différentiable ne s'annulant en aucun point de son domaine de définition. Pour (x, y) assez voisin de (x_a, y_a) , coordonnées locales du point a , posons

$$z(x, y) = \int_{x_a}^x f(\xi, y) d\xi.$$

La fonction $(x, y) \mapsto z(x, y)$ est différentiable et vérifie

$$dz = \frac{\partial z(x, y)}{\partial x} dx + \frac{\partial z(x, y)}{\partial y} dy = f(x, y) dx + \frac{\partial z(x, y)}{\partial y} dy,$$

donc

$$dz \wedge dy = f(x, y) dx \wedge dy = \Omega.$$

Cela prouve que (y, z) est un système de coordonnées locales associé à une carte canonique dont le domaine contient le point a . Nous avons prouvé que l'assertion H_1 est vraie.

Pour un entier $k - 1 \geq 1$, supposons l'assertion H_{k-1} vraie, et montrons qu'alors l'assertion H_k est vraie aussi. Soit donc (M, Ω) une variété symplectique de dimension $2k$, et a un point de M . Soit f une fonction différentiable définie sur un voisinage du point a vérifiant $df(a) \neq 0$. Le champ de vecteurs hamiltonien X_f , défini par $i(X_f)\Omega = -df$, vérifie $X_f(a) \neq 0$. D'après le théorème de redressement local d'un champ de vecteurs, il existe une carte de M , dont le domaine contient le point a , dont les coordonnées locales z^1, \dots, z^{2k} sont telles que X_f ait pour expression

$$X_f = \frac{\partial}{\partial z^1}.$$

Posons $g = z^1$. Nous voyons alors que f et g sont deux fonctions différentiables, définies sur un voisinage du point a , qui vérifient

$$\{f, g\} = 1, \quad (*)$$

ce qui prouve, au passage, que df et dg sont linéairement indépendantes sur leur domaine de définition commun. Les deux champs de vecteurs hamiltoniens X_f et X_g sont donc linéairement indépendants sur leur domaine de définition commun et vérifient

$$[X_f, X_g] = X_{\{f, g\}} = 0.$$

D'après le théorème de Frobenius (rappelé paragraphe 7.1) ces champs de vecteurs engendrent un champ de 2-directions complètement intégrable, donc un feuilletage \mathcal{F} dont les feuilles, de dimension 2 sont tangentes en chacun de leurs points aux champs de vecteurs X_f et X_g . Il existe donc $2k - 2$ fonctions différentiables h_i , $1 \leq i \leq 2k - 2$, définies sur un voisinage du point a , dont les différentielles sont linéairement indépendantes, telles que les feuilles du feuilletage \mathcal{F} soient localement déterminées par des équations de la forme

$$h_i = c_i, \quad 1 \leq i \leq 2k - 2,$$

les c_i étant des constantes. Nous avons, pour tout i ($1 \leq i \leq 2k - 2$),

$$\{f, h_i\} = i(X_f)dh_i = 0, \quad \{g, h_i\} = i(X_g)dh_i = 0, \quad (**)$$

puisque X_f et X_g sont tangentes aux feuilles de \mathcal{F} et que sur chacune de ces feuilles les fonctions h_i sont constantes.

Montrons que df , dg et les dh_i , pour $1 \leq i \leq 2k - 2$, sont linéairement indépendantes. Supposons qu'il existe des coefficients réels λ , μ et ξ_i , avec $1 \leq$

$i \leq 2k - 2$, tels que

$$\lambda df + \mu dg + \sum_{i=1}^{2k-2} \xi_i dh_i$$

s'annule en un point. En faisant le crochet de Poisson de $\lambda f + \mu g + \sum_{i=1}^{2k-2} \xi_i h_i$ successivement avec f et avec g , nous voyons que λ et μ sont nécessairement nuls. Mais alors les ξ_i sont tous nuls aussi, puisque par hypothèse les dh_i , avec $1 \leq i \leq 2k - 2$, sont linéairement indépendantes.

Ainsi les fonctions f , g et h_i , avec $1 \leq i \leq 2k - 2$, forment un système de coordonnées locales au voisinage du point a . En tenant compte de (*) et (**), nous voyons que dans ces coordonnées locales, Ω a pour expression

$$\Omega = df \wedge dg + \sum_{1 \leq i < j \leq 2k-2} \varphi_{ij} dh_i \wedge dh_j,$$

où les φ_{ij} sont des fonctions différentiables des coordonnées locales f , g et h_i , $1 \leq i \leq 2k - 2$. En exprimant que $d\Omega = 0$, nous obtenons

$$\frac{\partial \varphi_{ij}}{\partial f} = 0, \quad \frac{\partial \varphi_{ij}}{\partial g} = 0.$$

Par suite, au voisinage du point a , les φ_{ij} ne dépendent que des coordonnées h_1, \dots, h_{2k-2} . Mais alors la 2-forme fermée

$$\Omega - df \wedge dg = \sum_{1 \leq i < j \leq 2k-2} \varphi_{ij} dh_i \wedge dh_j,$$

qui est de rang $2k - 2$, peut être considérée comme une forme symplectique définie sur un ouvert de \mathbb{R}^{2k-2} . Puisque nous avons supposé l'assertion H_{k-1} vraie, il existe sur cet ouvert un changement de coordonnées

$$(h_1, \dots, h_{2k-2}) \mapsto (x_2, \dots, x_k, x_{k+2}, \dots, x_{2k})$$

tel que, dans ces nouvelles coordonnées, on ait

$$\Omega - df \wedge dg = \sum_{i=2}^k dx_{k+i} \wedge dx_i.$$

En posant $g = x_1$ et $f = x_{k+1}$, nous pouvons écrire

$$\Omega = \sum_{i=1}^k dx_{k+i} \wedge dx_i.$$

Nous voyons ainsi qu'au voisinage du point a , (x_1, \dots, x_{2k}) est un système de coordonnées canoniques. L'assertion H_k est donc vraie.

En conclusion nous avons prouvé, par récurrence sur la dimension de la variété considérée, le théorème de Darboux. \square

9.2. Théorème de Jacobi-Lie-Carathéodory. — Soit (M, Ω) une variété symplectique de dimension $2n$. On suppose qu'il existe, sur un ouvert U de M , n fonctions différentiables f_1, \dots, f_n , deux à deux en involution,

$$\{f_i, f_j\} = 0, \quad 1 \leq i, j \leq n,$$

dont les différentielles sont linéairement indépendantes sur U . Au voisinage de chaque point de U , il existe n autres fonctions différentiables g_1, \dots, g_n , telles que l'expression locale de la forme symplectique Ω soit

$$\Omega = \sum_{i=1}^n df_i \wedge dg_i.$$

La détermination locale des fonctions g_1, \dots, g_n se fait par des calculs utilisant exclusivement des quadratures, des dérivations partielles et des éliminations.

Preuve : Supposons connu, au voisinage du point considéré, un système de coordonnées locales canoniques de M , noté $(x^1, \dots, x^n, y_1, \dots, y_n)$. On a donc

$$\Omega = \sum_{i=1}^n dx^i \wedge dy_i.$$

Supposons également qu'au voisinage du point considéré, les différentielles des fonctions $f_1, \dots, f_n, x^1, \dots, x^n$ soient linéairement indépendantes. Un lemme d'algèbre symplectique dû à Arnold (lemme 4.1 du chapitre III) montre d'ailleurs qu'on peut toujours, par une permutation canonique des coordonnées locales $x^1, \dots, x^n, y_1, \dots, y_n$, se ramener à ce cas. On peut alors prendre pour coordonnées locales $f_1, \dots, f_n, x^1, \dots, x^n$, et exprimer localement Ω sous la forme

$$\Omega = \sum_{i=1}^n \eta_i \wedge df_i,$$

où η_1, \dots, η_n sont des 1-formes appartenant à l'idéal engendré par dx^1, \dots, dx^n . Cela résulte en effet du fait que les fonctions f_i ($1 \leq i \leq n$) sont deux à deux en involution. On a donc

$$\eta_i = \sum_{k=1}^n \alpha_{ik}(x, f) dx^k.$$

Moyennant l'introduction de constantes, on peut aussi supposer qu'au point considéré, les coordonnées locales x^1, \dots, x^n sont nulles.

On note Ψ_t la famille à un paramètre $t \in [0, 1]$ d'applications

$$\Psi_t : (x^1, \dots, x^n, f_1, \dots, f_n) \mapsto ((1-t)x^1, \dots, (1-t)x^n, f_1, \dots, f_n).$$

On a

$$\Psi_0 = \text{id}, \quad \Psi_1(x^1, \dots, x^n, f_1, \dots, f_n) = (0, \dots, 0, f_1, \dots, f_n).$$

L'image réciproque de la forme symplectique Ω par l'application Ψ_t a pour expression

$$\Psi_t^* \Omega = \sum_{i,k} (1-t) \alpha_{ik}((1-t)x, f) dx^k \wedge df_i.$$

En dérivant par rapport à t , on obtient

$$\frac{d}{dt} \Psi_t^* \Omega = - \sum_{i,k} \alpha_{ik}((1-t)x, f) dx^k \wedge df_i - \sum_{i,k,l} (1-t) x^l \frac{\partial \alpha_{ik}}{\partial x^l}((1-t)x, f) dx^k \wedge df_i.$$

En exprimant que Ω est fermée, on obtient

$$0 = d\Omega = - \sum_{i,k,l} \frac{\partial \alpha_{ik}(x, f)}{\partial x^l} dx^l \wedge dx^k \wedge df_i + \sum_{i,k,l} \frac{\partial \alpha_{ik}(x, f)}{\partial f_l} df_l \wedge dx^k \wedge df_i,$$

d'où

$$\frac{\partial \alpha_{ik}}{\partial x^l} - \frac{\partial \alpha_{il}}{\partial x^k} = 0, \quad \frac{\partial \alpha_{ik}}{\partial f_l} - \frac{\partial \alpha_{lk}}{\partial f_i} = 0. \quad (*)$$

On peut donc écrire

$$\frac{d}{dt} \Psi_t^* \Omega = - \sum_{i,k} \alpha_{ik}((1-t)x, f) dx^k \wedge df_i - \sum_{i,k,l} (1-t) x^l \frac{\partial \alpha_{il}}{\partial x^k}((1-t)x, f) dx^k \wedge df_i.$$

Compte tenu de $\Psi_0^* \Omega = \Omega$ et de $\Psi_1^* \Omega = 0$, on peut écrire

$$\begin{aligned} \Omega &= - \int_0^1 \left(\frac{d}{dt} \Psi_t^* \Omega \right) dt \\ &= \sum_{i,k} \left[\int_0^1 \alpha_{ik}((1-t)x, f) dt \right] dx^k \wedge df_i \\ &\quad + \sum_{i,k,l} \left[\int_0^1 (1-t) x^l \frac{\partial \alpha_{il}}{\partial x^k}((1-t)x, f) dt \right] dx^k \wedge df_i. \end{aligned}$$

Posons

$$g_i = - \sum_l \left[\int_0^1 x^l \alpha_{il}((1-t)x, f) dt \right].$$

On a alors

$$\begin{aligned} dg_i &= - \sum_l \left[\int_0^1 \alpha_{il}((1-t)x, f) dt \right] dx^l \\ &\quad - \sum_{k,l} \left[\int_0^1 (1-t) x^l \frac{\partial \alpha_{il}}{\partial x^k}((1-t)x, f) dt \right] dx^k \\ &\quad - \sum_{k,l} \left[\int_0^1 x^l \frac{\partial \alpha_{il}}{\partial f_k}((1-t)x, f) dt \right] df_k. \end{aligned}$$

On en déduit

$$\begin{aligned} \sum_i df_i \wedge dg_i &= \sum_{i,l} \left[\int_0^1 \alpha_{il}((1-t)x, f) dt \right] dx^l \wedge df_i \\ &+ \sum_{i,k,l} \left[\int_0^1 (1-t)x^l \frac{\partial \alpha_{il}}{\partial x^k}((1-t)x, f) dt \right] dx^k \wedge df_i \\ &+ \sum_{i,k,l} \left[\int_0^1 x^l \frac{\partial \alpha_{il}}{\partial f_k}((1-t)x, f) dt \right] df_k \wedge df_i. \end{aligned}$$

Mais le troisième terme est nul, compte tenu de la seconde relation (*) ci-dessus.

On a donc bien

$$\Omega = \sum_{i=1}^n df_i \wedge dg_i. \quad \square$$

9.3. Remarque. — Plaçons-nous dans les hypothèses du théorème de Jacobi-Lie-Carathéodory. La famille de fonctions $(g_1, \dots, g_n, f_1, \dots, f_n)$ est alors un système de coordonnées locales sur M . On peut appeler les f_i *pré-actions* et les g_i *pré-angles*. Supposons qu'on ait de plus une fonction différentiable H en involution avec chacune des fonctions f_i , c'est-à-dire vérifiant, pour chaque i ($1 \leq i \leq n$),

$$\{H, f_i\} = 0.$$

En d'autres termes, (M, Ω, H) est un système hamiltonien dont les fonctions f_i ($1 \leq i \leq n$) sont des intégrales premières deux à deux en involution. En chaque point x de l'ouvert U , les champs de vecteurs hamiltoniens X_H et X_{f_i} engendrent un sous-espace vectoriel isotrope de $T_x M$. Comme les X_{f_i} sont linéairement indépendants, leurs valeurs au point x engendrent un sous-espace vectoriel de dimension n de $T_x M$. Un sous-espace vectoriel isotrope de $T_x M$ étant nécessairement de dimension inférieure ou égale à n , $X_H(x)$ est combinaison linéaire des $X_{f_i}(x)$, $1 \leq i \leq n$. Par produit intérieur avec Ω , nous en déduisons que $dH(x)$ est combinaison linéaire des $df_i(x)$. Nous pouvons donc écrire

$$dH(x) = \sum_{i=1}^n \omega_i(x) df_i(x),$$

les ω_i étant des fonctions différentiables. Cette expression montre que les dérivées partielles de H , exprimées au moyen des coordonnées locales $f_1, \dots, f_n, g_1, \dots, g_n$, vérifient

$$\frac{\partial H}{\partial f_i} = \omega_i, \quad \frac{\partial H}{\partial g_i} = 0, \quad (1 \leq i \leq n).$$

Autrement dit localement, au voisinage de chaque point, H ne dépend que des n premières coordonnées locales f_1, \dots, f_n . Il en est de même des fonctions ω_i , puisque ce sont des dérivées partielles de la fonction H .

Dans le système de coordonnées locales $(f_1, \dots, f_n, g_1, \dots, g_n)$, les équations de Hamilton déterminées par le champ de vecteurs hamiltonien X_H s'écrivent

$$\begin{cases} \frac{dg_i}{dt} = \omega_i = \frac{\partial H}{\partial f_i}, \\ \frac{df_i}{dt} = 0. \end{cases}$$

Ces équations s'intègrent immédiatement, et conduisent à

$$\begin{cases} g_i(t) = g_i(t_0) + (t - t_0) \frac{\partial H}{\partial f_i}(f(t_0)), \\ f_i(t) = f_i(t_0). \end{cases}$$

Localement, on a intégré le système hamiltonien (M, Ω, H) , en le linéarisant. La détermination explicite des solutions locales n'a utilisé que des quadratures, des dérivations partielles et des éliminations. C'est pour cette raison qu'un système hamiltonien (M, Ω, H) , sur une variété symplectique (M, Ω) de dimension $2n$, qui possède n intégrales premières f_1, \dots, f_n , deux à deux en involution et dont les différentielles sont linéairement indépendantes sur un ouvert dense de M , est dit *complètement intégrable*. L'étude des systèmes complètement intégrables est poursuivie dans le chapitre suivant.

Les théorèmes de Darboux et de Jacobi-Lie-Carathéodory indiqués ci-dessus indiquent des propriétés locales, au voisinage d'un point. Il existe des généralisations de ces théorèmes qui indiquent des propriétés semi-globales, valables au voisinage d'une sous-variété de la variété symplectique considérée. Le théorème des actions-angles (qui sera exposé dans le prochain chapitre) peut être considéré comme une version semi-globale du théorème de Jacobi-Lie-Carathéodory. J.J Duistermaat en a d'ailleurs donné d'autres versions encore plus globales, notamment dans l'article *On global action-angle variables*, Comm. on Pure and Appl. Math. **33** (1980) 687–706. Quant au théorème de Darboux, A. Weinstein en a donné plusieurs généralisations semi-globales. Nous en indiquons ci-dessous une, dont nous aurons besoin plus loin pour prouver le théorème des actions-angles.

9.4. Théorème de Darboux-Weinstein. — *Soit (M, Ω) une variété symplectique et N une sous-variété lagrangienne de M . Il existe un difféomorphisme symplectique d'un ouvert de M contenant N sur un ouvert du fibré cotangent T^*N (muni de sa forme symplectique canonique) contenant l'image de la section nulle, dont la restriction à N est la section nulle de T^*N , c'est-à-dire l'application associant à chaque point de N le covecteur nul attaché à ce point.*

Le lecteur intéressé pourra étudier la démonstration dans l'article d'A. Weinstein, *Symplectic manifolds and their Lagrangian submanifolds*, Advances in Math. **6** (1971) 329–346, ou dans le livre de P. Libermann et C.-M. Marle, *Symplectic geometry and analytical mechanics*, Reidel, Dordrecht, 1987.

Chapitre V

Les systèmes intégrables et leurs perturbations

1. La complète intégrabilité

1.1. Définition. — Soit (M, Ω, H) un système hamiltonien sur une variété symplectique (M, Ω) de dimension $2n$. On dit que ce système est *complètement intégrable* (ou, en abrégé, lorsqu'il n'y a pas de risque de confusion, *intégrable*) s'il existe n fonctions différentiables f_1, \dots, f_n , intégrales premières de ce système, deux à deux en involution, dont les différentielles sont linéairement indépendantes sur un ouvert dense de M .

1.2. Commentaire. — Ainsi qu'on l'a vu plus haut (chapitre IV, paragraphe 7), avec les notations de la définition ci-dessus, les fonctions f_i sont des intégrales premières du système hamiltonien (M, Ω, H) si et seulement si elles vérifient, pour tout i ($1 \leq i \leq n$),

$$\{H, f_i\} = 0.$$

Elles sont deux à deux en involution si et seulement si, pour tous i et j ($1 \leq i, j \leq n$),

$$\{f_i, f_j\} = 0.$$

En utilisant le théorème de Jacobi-Lie-Carathéodory nous avons montré (chapitre IV, remarque 9.3) que dans les hypothèses de la définition ci-dessus, localement, au voisinage de chaque point de l'ouvert U sur lequel les différentielles des fonctions f_i sont linéairement indépendantes, la valeur en un point du hamiltonien H ne dépend que des valeurs, en ce point, des n fonctions f_i . Comme de plus l'application $f = (f_1, \dots, f_n) : U \rightarrow \mathbb{R}^n$ est une submersion, pour toute composante connexe U_k de l'ouvert U , $V_k = f(U_k)$ est un ouvert de \mathbb{R}^n sur lequel il existe une fonction différentiable \widehat{H}_k telle que

$$H|_{U_k} = \widehat{H}_k \circ f.$$

Nous voyons donc que sur chaque composante connexe U_k de l'ouvert U , le hamiltonien H est composé de $f = (f_1, \dots, f_n)$ et d'une fonction différentiable \widehat{H}_k définie sur un ouvert de \mathbb{R}^n . Dans de nombreux exemples, le hamiltonien H est en fait tout simplement l'une des fonctions f_i ($1 \leq i \leq n$).

Nous avons vu aussi (chapitre IV, remarque 9.3) que les courbes intégrales d'un système hamiltonien complètement intégrable peuvent être localement explicitement déterminées en n'employant que des dérivations partielles, des quadratures et des éliminations. C'est cette propriété qui justifie qu'on dise de tels systèmes qu'ils sont complètement intégrables.

Cependant, l'intérêt principal des systèmes complètement intégrables ne tient pas au fait que leurs courbes intégrales sont calculables par des opérations plus élémentaires que la résolution d'une équation différentielle. Il résulte plutôt du fait que l'évolution de ces systèmes au cours du temps est en général très simple et ne donne lieu à aucun comportement chaotique. C'est en effet une conséquence des deux théorèmes (de Liouville et des actions-angles) ci-dessous.

1.3. Théorème de Liouville. — Soit (M, Ω, H) un système hamiltonien complètement intégrable sur une variété symplectique (M, Ω) de dimension $2n$. Soit $f = (f_1, \dots, f_n) : M \rightarrow \mathbb{R}^n$ une famille de n intégrales premières de ce système, deux à deux en involution, dont les différentielles sont linéairement indépendantes sur un ouvert dense U de M . Soit $c = (c_1, \dots, c_n) \in \mathbb{R}^n$, et N une composante connexe de $U \cap f^{-1}(c)$. On suppose N compacte. Alors N est une sous-variété lagrangienne de M difféomorphe à un tore T^n . De plus, le champ de vecteurs hamiltonien X_H est tangent à N en chacun de ses points, et N est invariant par le flot de X_H . Il existe sur N un système de coordonnées angulaires $(\alpha_1, \dots, \alpha_n)$ dans lequel la restriction à N du flot de X_H a pour expression

$$\alpha_i(t) = \alpha_i(0) + \omega_i t, \quad (1 \leq i \leq n),$$

où les ω_i sont des constantes.

Preuve : La restriction de f à U étant une submersion, $U \cap f^{-1}(c)$ est une sous-variété de M , de dimension n . Pour chaque point x de cette sous-variété, $T_x(U \cap f^{-1}(c))$ est l'ensemble des vecteurs $v \in T_x M$ qui vérifient, pour tout i ($1 \leq i \leq n$), $\langle df_i(x), v \rangle = 0$, ou encore $\Omega(X_{f_i}(x), v) = 0$, où on a noté X_{f_i} le champ de vecteurs hamiltonien associé à f_i . L'espace $T_x(U \cap f^{-1}(c))$ est donc l'orthogonal symplectique du sous-espace vectoriel de $T_x M$ engendré par les vecteurs $X_{f_i}(x)$, $1 \leq i \leq n$. Mais ceux-ci sont au nombre de n , linéairement indépendants, et ils vérifient, pour tous i et j , $\Omega(x)(X_{f_i}(x), X_{f_j}(x)) = \{f_i, f_j\}(x) = 0$. Le sous-espace vectoriel qu'ils engendrent est donc lagrangien, égal à son orthogonal. Nous avons ainsi prouvé que pour tout $x \in N$, $T_x(U \cap f^{-1}(c))$ est un sous-espace vectoriel lagrangien de $T_x M$. La sous-variété $U \cap f^{-1}(c)$ de M est donc lagrangienne. Il en est de même de chacune de ses composantes connexes, en particulier de N . Au cours de la démonstration nous avons vu aussi que pour tout point x de N , les vecteurs $X_{f_i}(x)$ forment une base de $T_x N$.

Comme les champs de vecteurs X_{f_i} sont tangents à N en chacun de ses points, cette sous-variété est invariante par le flot de chacun de ces champs de vecteurs. Comme N est compacte, pour chaque i ($1 \leq i \leq n$) le flot réduit Φ_i de la restriction à N de X_{f_i} est défini pour toute valeur du temps t . De plus, pour tous i et j ($1 \leq i, j \leq n$), nous avons

$$[X_{f_i}, X_{f_j}] = X_{\{f_i, f_j\}} = 0.$$

On montre que cela implique, pour tous t_i et $t_j \in \mathbb{R}$ et tout $x \in N$,

$$\Phi_i(t_i, \Phi_j(t_j, x)) = \Phi_j(t_j, \Phi_i(t_i, x)). \quad (*)$$

Posons alors, pour tout $t = (t_1, \dots, t_n) \in \mathbb{R}^n$ et tout $x \in N$,

$$\Psi(t, x) = \Phi_1\left(t_1, \Phi_2\left(t_2, \dots, \Phi_n(t_n, x)\right)\right).$$

En utilisant (*) et les propriétés de composition du flot réduit d'un champ de vecteurs, on montre aisément que pour tous $t = (t_1, \dots, t_n)$ et $s = (s_1, \dots, s_n) \in \mathbb{R}^n$ et tout $x \in N$,

$$\Psi(t, \Psi(s, x)) = \Psi(t + s, x).$$

On vérifie aussi que pour tout $x \in N$,

$$\Psi(0, x) = x.$$

L'application $\Psi : \mathbb{R}^n \times N \rightarrow N$ est donc une action du groupe additif \mathbb{R}^n sur la variété N . En remarquant que les champs de vecteurs X_{f_i} dont les flots engendrent cette action forment, en tout point de N , une base de l'espace vectoriel tangent à N en ce point, on voit que l'action Ψ est localement libre et localement transitive. Comme N est connexe et compacte, cette action est globalement transitive. Soit a un point particulier de N , et G_a le stabilisateur de ce point, relativement à l'action Ψ , c'est-à-dire

$$G_a = \{ t \in \mathbb{R}^n \mid \Psi(t, a) = a \}.$$

On vérifie aisément que G_a est un sous-groupe fermé du groupe additif \mathbb{R}^n et que pour tout $t \in \mathbb{R}^n$, $\Psi(t, a)$ ne dépend que de la classe d'équivalence de t modulo le sous-groupe G_a . Il existe donc une application $\psi : \mathbb{R}^n / G_a \rightarrow N$ telle que pour tout $t \in \mathbb{R}^n$, on ait $\Psi(t, a) = \psi(\hat{t})$, où on a noté \hat{t} l'élément de \mathbb{R}^n / G_a , classe d'équivalence de t modulo G_a . La théorie des actions différentiables de groupes de Lie permet alors de montrer que ψ est un difféomorphisme de l'espace homogène \mathbb{R}^n / G_a sur N . La dimension de \mathbb{R}^n / G_a est $\dim \mathbb{R}^n - \dim G_a = n - \dim G_a$. La dimension de N , sous-variété lagrangienne de M , est n . Par suite, la dimension de G_a est nulle; en d'autres termes, G_a est un sous-groupe discret du groupe additif \mathbb{R}^n . Or les sous-groupes discrets du groupe additif \mathbb{R}^n sont bien connus : ce sont les réseaux. Un réseau de rang p (où p est un entier $\leq n$) dans \mathbb{R}^n est, par définition, un sous-ensemble de \mathbb{R}^n de la forme

$$\left\{ \sum_{i=1}^p k_i v_i \mid k = (k_1, \dots, k_p) \in \mathbb{Z}^p \right\},$$

où (v_1, \dots, v_p) est une famille de p éléments linéairement indépendants de \mathbb{R}^n , appelée *système de générateurs* du réseau. On montre que le quotient de \mathbb{R}^n par un réseau de rang p est isomorphe au groupe $\mathbb{R}^{n-p} \times T^p$, produit de \mathbb{R}^{n-p} et du tore T^p de dimension p . Dans le cas qui nous intéresse ici, \mathbb{R}^n/G_a est diffeomorphe à N , qui par hypothèse est compact; donc le rang du réseau G_a est n , et \mathbb{R}^n/G_a est isomorphe au tore T^n , c'est-à-dire au produit de n cercles trigonométriques T , tous isomorphes à $\mathbb{R}/2\pi\mathbb{Z}$. Nous avons donc prouvé que N est diffeomorphe au tore T^n .

Soit (v_1, \dots, v_n) un système de générateurs du réseau G_a . L'application de \mathbb{R}^n sur N ,

$$\alpha = (\alpha_1, \dots, \alpha_n) \mapsto x = \Psi \left(\sum_{i=1}^n \frac{\alpha_i v_i}{2\pi}, a \right)$$

est différentiable, partout de rang n ; l'image réciproque du point a par cette application est le réseau $(2\pi\mathbb{Z})^n$ de \mathbb{R}^n , et l'image réciproque d'un point quelconque de N est un translaté de ce réseau. On dit que cette application est un *revêtement* de N par \mathbb{R}^n . En considérant les α_i comme définis modulo 2π , on peut inverser cette application. Son inverse, $x \mapsto \alpha = (\alpha_1, \dots, \alpha_n)$, est un diffeomorphisme de N sur le tore $\mathbb{R}^n/(2\pi\mathbb{Z})^n$, c'est-à-dire un système de coordonnées angulaires sur le tore N .

D'autre part, le raisonnement présenté dans la remarque 9.3 du chapitre IV a montré qu'au voisinage de chaque point, le hamiltonien H n'est fonction que de $f = (f_1, \dots, f_n)$. Nous pouvons donc écrire

$$dH = \sum_{i=1}^n \frac{\partial H(f_1, \dots, f_n)}{\partial f_i} df_i,$$

donc aussi, pour les champs de vecteurs hamiltoniens,

$$X_H = \sum_{i=1}^n \frac{\partial H(f_1, \dots, f_n)}{\partial f_i} X_{f_i}.$$

La sous-variété N étant localement déterminée par les équations $f_i = c_i$, où les c_i sont des constantes, nous avons, sur cette sous-variété,

$$X_H \Big|_N = \sum_{i=1}^n \lambda_i X_{f_i} \Big|_N, \quad (**)$$

où les $\lambda_i = \frac{\partial H(f_1, \dots, f_n)}{\partial f_i} \Big|_{f=c}$ sont des constantes sur N .

Soit (e_1, \dots, e_n) la base canonique de \mathbb{R}^n ; avec les notations matricielles habituelles, l'élément e_i de cette base est le vecteur-colonne dont la seule composante non nulle est la i -ème, égale à 1. Les vecteurs $\frac{v_i}{2\pi}$, $1 \leq i \leq n$, forment une autre base

de \mathbb{R}^n , et nous pouvons écrire, pour chaque j ($1 \leq j \leq n$),

$$e_j = \sum_{i=1}^n A_{ij} \frac{v_i}{2\pi}.$$

Les A_{ij} , coefficients de la matrice de changement de base, sont des constantes.

Pour tout $t = (t_1, \dots, t_n) \in \mathbb{R}^n$, nous pouvons écrire

$$t = \sum_{j=1}^n t_j e_j = \sum_{i=1}^n \left(\sum_{j=1}^n A_{ij} t_j \right) \frac{v_i}{2\pi}. \quad (***)$$

Mais pour chaque point x de N et chaque i ($1 \leq i \leq n$), le vecteur $\frac{\partial}{\partial \alpha_i}(x)$ de la base de $T_x N$ associée aux coordonnées angulaires $\alpha = (\alpha_1, \dots, \alpha_n)$, est

$$\frac{\partial}{\partial \alpha_i}(x) = \frac{\partial}{\partial \alpha_i} \Psi \left(\sum_{k=1}^n \frac{\alpha_k v_k}{2\pi}, x \right) \Big|_{\alpha=0}.$$

De même, la valeur au point x du champ de vecteurs X_{f_j} est

$$X_{f_j}(x) = \frac{\partial}{\partial t_j} \Psi(t, x) \Big|_{t=0}.$$

En remplaçant t par son expression (***) donnée ci-dessus et en effectuant cette dérivation partielle, nous obtenons

$$X_{f_j}(x) = \sum_{i=1}^n A_{ij} \frac{\partial}{\partial \alpha_j}(x).$$

Compte tenu de (**), nous en déduisons l'expression de X_H :

$$X_H(x) = \sum_{i=1}^n \omega_i \frac{\partial}{\partial \alpha_i}(x), \quad \text{avec} \quad \omega_i = \sum_{j=1}^n A_{ij} \lambda_j.$$

L'équation différentielle associée au champ de vecteurs $X_H|_N$ s'écrit donc, dans les coordonnées angulaires $(\alpha_1, \dots, \alpha_n)$,

$$\frac{d\alpha_i(t)}{dt} = \omega_i, \quad 1 \leq i \leq n;$$

les ω_i étant des constantes sur N , son flot est donné par les formules très simples

$$\alpha_i(t) = \alpha_i(0) + \omega_i t, \quad 1 \leq i \leq n. \quad \square$$

1.4. Remarque. — Le théorème de Liouville donne une description précise des courbes intégrales du champ de vecteurs X_H contenues dans la sous-variété lagrangienne N , composante connexe de $U \cap f^{-1}(c)$, pour une valeur fixée de $c = (c_1, \dots, c_n)$. Il est naturel de se demander comment ces courbes se modifient lorsqu'on modifie légèrement la valeur de c . Le théorème des actions-angles, essentiellement dû à Arnol'd, répond à cette question. Il a été publié pour la première fois en français dans le livre de V. Arnol'd et A. Avez, *Problèmes ergodiques de la mécanique classique*, Gauthier-Villars, Paris, 1967. Aussi nous l'appellerons *théorème de Liouville-Arnol'd-Avez*.

1.5. Théorème (Liouville-Arnol'd-Avez). — *Les hypothèses et notations sont les mêmes que celles du théorème de Liouville. Il existe, sur un voisinage ouvert W de la sous-variété lagrangienne N de M , $W \subset U$, n fonctions différentiables I_1, \dots, I_n , chaque I_i n'étant fonction que de (f_1, \dots, f_n) , dont les différentielles dI_1, \dots, dI_n sont linéairement indépendantes, telles que les flots des champs de vecteurs $\sharp dI_i$ soient périodiques, de période 2π . Il existe également sur W des variables angulaires $\gamma_1, \dots, \gamma_n$, définies modulo 2π , telles que*

$$\Omega = \sum_{i=1}^n dI_i \wedge d\gamma_i.$$

L'ouvert W de M s'identifie à un voisinage de la section nulle du fibré cotangent au tore de dimension n , $\gamma_1, \dots, \gamma_n$ étant les variables angulaires sur ce tore, I_1, \dots, I_n les coordonnées conjuguées sur les fibres cotangentes. On dit que $(I_1, \dots, I_n, \gamma_1, \dots, \gamma_n)$ sont des variables actions-angles. La restriction à W du hamiltonien H n'est fonction que de $I = (I_1, \dots, I_n)$, non des angles $\gamma = (\gamma_1, \gamma_n)$. Sur l'ouvert W , le flot du champ de vecteurs X_H est donné par les formules

$$\begin{cases} \gamma_i(t) = \gamma_i(0) + \frac{\partial H(I(t_0))}{\partial I_i} t, \\ I_i(t) = I_i(t_0), \end{cases} \quad (1 \leq i \leq n).$$

Preuve : Puisque N est connexe et compact, c'est un tore lagrangien. D'après le théorème de Darboux-Weinstein (IV.9.4), il existe un voisinage W de N dans M difféomorphe, par un difféomorphisme symplectique, à un voisinage de la section nulle de T^*N . On peut donc définir une projection $q : W \rightarrow N$ correspondant à la projection canonique de T^*N sur N . En restreignant éventuellement W , on peut faire en sorte que (q, f) soit un difféomorphisme de W sur le produit de N et d'un voisinage ouvert $f(W)$ de a dans \mathbb{R}^n . Pour tout $z \in f(W)$, $f^{-1}(z)$ est une sous-variété lagrangienne qui rencontre chaque fibre de $q : W \rightarrow N$ en un point unique. On peut donc identifier cette sous-variété au graphe d'une 1-forme fermée η_z , paramétrée par z . Soit $\rho : \mathbb{R}^n \rightarrow N$ le revêtement universel du tore N . Pour tout $z \in f(W)$, il existe une fonction différentiable S_z , définie sur \mathbb{R}^n , telle que

$$\rho^* \eta_z = dS_z.$$

Chaque fonction S_z est d'ailleurs déterminée à une constante additive arbitraire près. En choisissant convenablement ces constantes, on peut faire en sorte que la fonction $S : \mathbb{R}^n \times f(W) \rightarrow \mathbb{R}$, définie par

$$S(\beta, z) = S_z(\beta), \quad \beta \in \mathbb{R}^n, \quad z \in f(W),$$

soit différentiable. On a

$$\rho^* \eta_z = \sum_{i=1}^n \frac{\partial S}{\partial \beta^i}(\beta, z) d\beta^i.$$

De plus, les dérivées partielles $\frac{\partial S}{\partial \beta^i}$ sont périodiques de période 2π en chacune des variables β^1, \dots, β^n . Posons, pour tout $z \in f(W)$,

$$I_i(z) = \frac{1}{2\pi} \int_0^{2\pi} \frac{\partial S}{\partial \beta^i}(\beta^1, \dots, \beta^n, z) d\beta^i.$$

Les fonctions I_i ainsi définies ne dépendent que de z , non de β , et vérifient $I_i(a) = 0$, puisque $f^{-1}(a)$ s'identifie à la section nulle de T^*N .

Lorsqu'on considère les β^i , $1 \leq i \leq n$, comme des coordonnées angulaires, définies modulo 2π , sur le tore N , et qu'on identifie W à un ouvert de T^*N avec, sur les fibres, les coordonnées p_1, \dots, p_n , conjuguées des coordonnées angulaires β_1, \dots, β_n , le difféomorphisme $(q, f)^{-1} : N \times f(W) \rightarrow W$ s'exprime par

$$(\beta, z) \mapsto \left(\beta, p_1 = \frac{\partial S(\beta, z)}{\partial \beta^1}, \dots, p_n = \frac{\partial S(\beta, z)}{\partial \beta^n} \right).$$

Le jacobien

$$\det \left(\frac{\partial^2 S(\beta, z)}{\partial \beta^i \partial z_j} \right)$$

ne s'annule donc en aucun point. Un lemme élémentaire ([15], page 175) permet d'en déduire que le déterminant

$$\det \left(\frac{\partial I_i}{\partial z_j} \right)$$

ne s'annule en aucun point z de $f(W)$. En restreignant éventuellement W , on peut donc faire en sorte que (β, I) , avec $\beta = (\beta^1, \dots, \beta^n)$ et $I = (I_1, \dots, I_n)$, soit un système de coordonnées sur W , les β^i étant des coordonnées angulaires, définies modulo 2π . On notera $\widehat{S}(\beta, I)$ l'expression de la fonction S au moyen de ces nouvelles coordonnées. On a

$$\frac{\partial \widehat{S}(\beta, I)}{\partial \beta^i} = \frac{\partial S(\beta, z)}{\partial \beta^i}, \quad 1 \leq i \leq n,$$

car les nouvelles coordonnées I ne sont fonctions que des coordonnées z , non des β . La 1-forme de Liouville α de T^*N , considérée grâce aux identifications faites comme une 1-forme sur W , a donc pour expression

$$\alpha = \sum_{i=1}^n p_i d\beta^i = \sum_{i=1}^n \frac{\partial \widehat{S}(\beta, I)}{\partial \beta^i} d\beta^i.$$

Par suite, compte tenu de la symétrie des dérivées partielles secondes de \widehat{S} par rapport aux variables β^i et β^j ,

$$\Omega = d\alpha = \sum_{(i,j)} \frac{\partial^2 \widehat{S}(\beta, I)}{\partial \beta^i \partial I_j} dI_j \wedge d\beta^i.$$

Posons

$$\gamma^j(\beta, I) = \frac{\partial \widehat{S}(\beta, I)}{\partial I_j}.$$

On peut alors écrire

$$d\widehat{S} = \sum_{i=1}^n \left(\frac{\partial \widehat{S}}{\partial \beta^i} d\beta^i + \frac{\partial \widehat{S}}{\partial I_i} dI_i \right) = \alpha + \sum_{i=1}^n \gamma^i dI_i,$$

d'où

$$\Omega = d\alpha = \sum_{i=1}^n dI_i \wedge d\gamma^i.$$

Par suite, l'application $(\beta, I) \mapsto (\gamma, I)$ est un difféomorphisme local. D'autre part, les fonctions $\gamma^j(\beta, I)$ vérifient

$$\gamma^j(\beta^1, \dots, \beta^k + 2\pi, \dots, \beta^n, I) - \gamma^j(\beta^1, \dots, \beta^n, I) = 2\pi \frac{\partial I_k}{\partial I_j} = 2\pi \delta_k^j.$$

Elles sont donc de la forme

$$\gamma^j(\beta, I) = \beta^j + \widetilde{\gamma}^j(\beta, I),$$

où les fonctions $\widetilde{\gamma}^j$ sont périodiques de période 2π en chacune des variables β^1, \dots, β^n . Par suite, la classe modulo 2π de chaque $\gamma^j(\beta, I)$ ne dépend que de I et des classes modulo 2π de chacune des variables β^1, \dots, β^n . En d'autres termes, lorsqu'on les considère comme des fonctions à valeurs angulaires (définies modulo 2π), les γ^j peuvent être considérées comme définies sur l'ouvert W de M . Pour chaque valeur fixée de I , l'application $\beta \mapsto \gamma(\beta, I)$ est une application étale du tore de dimension n sur lui-même, donc une application de revêtement. Mais

$$(\beta, s) \mapsto \beta + s\widetilde{\gamma}(\beta, I), \quad 0 \leq s \leq 1,$$

est une homotopie différentiable de cette application de revêtement sur l'application identique du tore de dimension n . Donc pour chaque valeur fixée de I , $\beta \mapsto \gamma(\beta, I)$ est un difféomorphisme du tore sur lui-même. Par suite, (I, γ) est un système de coordonnées sur l'ouvert W de M , les n coordonnées $\gamma = (\gamma^1, \dots, \gamma^n)$ étant à valeurs angulaires (définies modulo 2π). C'est le système de coordonnées actions-angles cherché. \square

1.6. Remarque. — La détermination des coordonnées actions-angles $I_1, \dots, I_n, \gamma^1, \dots, \gamma^n$ utilise seulement des quadratures, des éliminations et des dérivations partielles. On a cependant, au cours de la démonstration, utilisé l'identification de l'ouvert W de M à un voisinage de la section nulle du fibré cotangent au tore et supposé connues la 1-forme de Liouville α , la projection canonique $q : W \rightarrow N$, et un système de coordonnées angulaires $\beta = (\beta_1, \dots, \beta_n)$ sur N . Ainsi qu'on va le voir, la détermination effective de coordonnées actions-angles est possible, par quadratures, éliminations et dérivations partielles, même lorsqu'on ne connaît pas

explicitement α , q et β . Il suffit de connaître, sur un voisinage de N , une 1-forme η telle que

$$d\eta = \Omega$$

et, pour chaque élément z d'un voisinage de a dans \mathbf{R}^n , une famille de n courbes fermées $c_1(z), \dots, c_n(z)$, contenues dans $f^{-1}(z)$ et dépendant différentiablement de z , engendrant l'homologie du tore $f^{-1}(z)$. Les coordonnées d'action I_i peuvent alors être définies par les formules

$$I_i = \frac{1}{2\pi} \int_{c_i} \eta.$$

En utilisant la formule de Stokes et en remarquant que les sous-variétés $f^{-1}(z)$ sont lagrangiennes, on vérifie en effet que les I_i ainsi définies ne sont fonctions que de z et ne diffèrent des I_i introduites dans la démonstration que par l'addition éventuelle de constantes. En procédant comme dans la démonstration de 1.2, on peut alors, sur un voisinage dans M de chaque point de N , déterminer les coordonnées angulaires γ_i , à des constantes additives arbitraires près. Compte tenu de la compacité de N , on peut enfin choisir ces constantes arbitraires afin d'avoir des coordonnées angulaires globales.

2. Quelques exemples

2.1. Les actions-angles du pendule sphérique. — On considère un point matériel de masse m , mobile sur une sphère de rayon R . On note \vec{x} le vecteur représentant la position de ce point, \vec{v} sa vitesse et $\vec{p} = m\vec{v}$ son impulsion. Les vecteurs \vec{x} et \vec{p} doivent vérifier

$$\|\vec{x}\|^2 = R^2, \quad \vec{x} \cdot \vec{p} = 0.$$

L'espace des phases, formé par les couples de vecteurs (\vec{x}, \vec{p}) vérifiant les conditions ci-dessus, est de dimension 4 : il s'identifie au fibré cotangent à la sphère S^2 . On a sur cet espace deux intégrales premières du système, en involution,

– le hamiltonien H , qui a pour expression

$$H = \frac{1}{2\|\vec{p}\|^2} - m\vec{g} \cdot \vec{x},$$

où \vec{g} désigne le vecteur accélération de la pesanteur ;

– la composante verticale K du moment cinétique, qui a pour expression

$$K = (\vec{e}_3, \vec{x}, \vec{p}),$$

produit mixte de \vec{e}_3 (vecteur unitaire vertical dirigé vers le haut), \vec{x} et \vec{p} .

On pose

$$\vec{g} \cdot \vec{e}_3 = -g.$$

Les valeurs critiques de (H, K) sont :

- celles correspondant aux positions d'équilibre du système,

$$(-mgR, 0) \quad \text{et} \quad (+mgR, 0) ;$$

- celles correspondant aux mouvements périodiques du système étudiés par Huygens, dans lesquels le point matériel parcourt, à vitesse constante, un cercle intersection de la sphère de rayon R avec un plan horizontal passant au dessous du centre de la sphère; ces valeurs critiques $(h_c(\lambda), k_c(\lambda))$ de (H, K) sont données, en fonction du paramètre λ vérifiant $0 < \lambda \leq 1$, par

$$h_c(\lambda) = mgR \left(\frac{(1 - \lambda^2)(1 + 3\lambda^2)}{2\lambda^2} - 1 \right) ,$$

$$k_c(\lambda) = \pm mg^{1/2} R^{3/2} \frac{1 - \lambda^4}{\lambda} .$$

L'image de l'application (H, K) est l'ensemble des $(h, k) \in \mathbf{R}^2$ vérifiant

$$h \geq h_c(\lambda), \quad k = k_c(\lambda), \quad \text{avec} \quad 0 < \lambda \leq 1 .$$

Pour toute valeur régulière (h, k) de (H, K) appartenant à l'image de cette application, $(H, K)^{-1}(h, k)$ est connexe, donc est un tore de dimension 2. On obtient un premier générateur c_1 de l'homologie de ce tore en considérant une orbite de l'action du groupe des rotations autour de l'axe vertical passant par le centre de la sphère. L'action correspondante I_1 est donnée par

$$I_1(h, k) = \frac{1}{2\pi} \int_{c_1} \alpha ,$$

la 1-forme de Liouville α ayant pour expression

$$\alpha = p_1 dx_1 + p_2 dx_2 + p_3 dx_3 .$$

On a noté x_1, x_2, x_3 et p_1, p_2, p_3 les composantes des vecteurs \vec{x} et \vec{p} dans un repère orthonormé $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$, avec \vec{e}_3 vertical dirigé vers le haut. Mais lorsqu'on parcourt le cycle c_1 , paramétré au moyen d'un angle φ variant de 0 à 2π , les différentielles de x_1, x_2 et x_3 vérifient

$$dx_1 = -x_2 d\varphi, \quad dx_2 = x_1 d\varphi, \quad dx_3 = 0 .$$

On obtient donc

$$I_1(h, k) = \frac{1}{2\pi} \int_0^{2\pi} (-p_1 x_2 + p_2 x_1) d\varphi = \frac{1}{2\pi} \int_0^{2\pi} k d\varphi = k .$$

On voit donc que la première variable d'action I_1 est simplement

$$I_1 = K .$$

Un second générateur de l'homologie du tore $(H, K)^{-1}(h, k)$ est obtenu en fixant $x_2 = 0$, et en faisant varier x_3 entre ses deux valeurs extrêmes possibles. On est amené à traiter séparément les cas où $k \neq 0$, et où $k = 0$.

Lorsque $k \neq 0$, x_3 reste compris dans un intervalle $[x_{3 \min}, x_{3 \max}]$, avec $-R < x_{3 \min} < x_{3 \max} < +R$. Puisqu'on a fixé $x_2 = 0$, et que la valeur k de K est imposée, les autres variables x_1, p_1, p_2, p_3 sont liées à x_3 par les relations

$$x_1^2 = R^2 - x_3^2, \quad p_1 x_1 + p_3 x_3 = 0, \quad p_2 x_1 = k,$$

ce qui montre que x_1 ne s'annule pas, et conduit à l'expression de la forme de Liouville

$$\alpha = p_1 dx_1 + p_2 dx_2 + p_3 dx_3 = \frac{R^2}{R^2 - x_3^2} p_3 dx_3.$$

Mais en exprimant que la valeur h de H est fixée, et en utilisant les relations précédentes, on voit que p_3 doit vérifier

$$p_3^2 = \frac{2m}{R^2} (R^2 - x_3^2) (h - mgx_3) - \frac{k^2}{R^2}.$$

L'intervalle $[x_{3 \min}, x_{3 \max}]$ des valeurs possibles de x_3 est déterminé en exprimant que le membre de droite de l'égalité ci-dessus doit être ≥ 0 . On voit que pour chaque valeur de x_3 appartenant à cet intervalle, il y a deux valeurs possibles de p_3 , opposées l'une de l'autre. On parcourt le générateur c_2 de l'homologie du tore en faisant d'abord croître x_3 de $x_{3 \min}$ à $x_{3 \max}$ et en prenant pour p_3 la valeur possible ≥ 0 , puis en faisant décroître x_3 de $x_{3 \max}$ à $x_{3 \min}$, et en prenant pour p_3 la valeur possible ≤ 0 . Afin de simplifier l'écriture, on introduit les variables réduites

$$u = \frac{x_3}{R}, \quad h^* = \frac{h}{mgR}, \quad k^* = \frac{k}{\sqrt{R^3 m^2 g}}.$$

On obtient

$$I_2(h, k) = \frac{\sqrt{2}}{\pi} m R^{3/2} g^{1/2} \int_{u_{\min}}^{u_{\max}} \frac{1}{1-u^2} \sqrt{(h^* - u)(1-u^2) - k^{*2}/2} du.$$

Lorsque $k = 0$, x_3 peut varier dans l'intervalle $[-R, R \inf(h^*, 1)]$. Un calcul semblable au précédent conduit à

$$I_2(h, 0) = \frac{\sqrt{2}}{\pi} m R^{3/2} g^{1/2} \int_{-1}^{\inf(h^*, 1)} \frac{1}{1-u^2} \sqrt{(h^* - u)(1-u^2)} du.$$

2.2. Autres exemples. — Le lecteur trouvera d'autres exemples de calculs explicites d'actions-angles, notamment pour le problème de Kovalevskaja, dans les publications de J.-P. Francoise, *Sur les actions-angles de la toupie de Kowalevski*, C. R. Acad. Sc. Paris **300** I (1985) 427–430, et *Calculs explicites d'actions-angles*, Séminaire de Mathématiques supérieures de l'université de Montréal (éd. G. Sabidussi) **102** (1986) 101–120. Le livre de R.H. Cushman et L.M. Bates *Global aspects of classical integrable systems*, Birkhäuser, Basel, 1997, présente une étude très détaillée de nombreux systèmes complètement intégrables. Il traite, notamment, de l'oscillateur harmonique, du système de Kepler, du pendule

sphérique et de deux des trois cas classiques où le mouvement d'un corps solide ayant un point fixe est complètement intégrables (Euler-Poinsot et Euler-Lagrange; le troisième cas étant celui de S. Kovalevskaja)(*)

3. Perturbation d'un système intégrable

Les systèmes hamiltoniens complètement intégrables ont un caractère relativement exceptionnel : étant donné un système hamiltonien complètement intégrable (M, Ω, H) , il suffit en général de modifier le hamiltonien H en lui ajoutant un terme arbitrairement petit pour que le nouveau système ne soit plus complètement intégrable. Or il existe de nombreux systèmes intéressants en pratique voisins de systèmes complètement intégrables. Considérons par exemple le système solaire, représenté par une famille finie de masses ponctuelles (schématisant le soleil et les planètes) évoluant dans l'espace physique de dimension 3, s'attirant mutuellement selon la loi de Newton. Il se trouve que la masse du soleil est 1000 fois plus grande que celle de tous les autres astres du système solaire réunis. Il est donc raisonnable, en première approximation, de négliger les forces que les planètes exercent les unes sur les autres, à l'exception de celles exercées par le soleil sur chacun des autres astres. Dans cette approximation, le système apparaît comme le produit d'un nombre fini de sous-systèmes évoluant séparément, chacun de ces sous-systèmes décrivant l'évolution d'un point matériel soumis à l'attraction gravitationnelle du soleil, considéré comme fixe dans l'espace. Le problème de l'évolution de tels sous-systèmes, appelé *problème de Kepler*, est complètement intégrable. Il en est donc de même du produit d'un nombre fini de ces sous-systèmes, indépendants les uns des autres.

Pour étudier l'évolution d'un système hamiltonien voisin d'un système complètement intégrable, tel que le système solaire, les astronomes ont développé des méthodes de calcul basées sur la théorie des perturbations. Avec l'aide des ordinateurs actuels, ces méthodes permettent de déterminer, de manière assez satisfaisante, l'évolution du système solaire pendant des durées de l'ordre du millier d'années. Cependant ces méthodes ne permettent pas de prévoir l'évolution du système solaire pendant des durées très grandes (de l'ordre du million d'années), ni de prédire son comportement asymptotique lorsque le temps t tend vers $\pm\infty$.

(*) Sophie (ou Sonya) Kovalevskaya, mathématicienne russe, a publié son célèbre mémoire sur le mouvement d'un corps solide en français, dans *Acta Mathematica*, 1888, sous le nom de Kowalevski. Peut-être était-ce pour avoir plus de chances d'être prise au sérieux, la communauté scientifique de l'époque laissant peu de place aux dames. Son vrai nom, selon le dictionnaire Larousse, est bien Kovalevskaja.

Cela tient au fait que les séries obtenues par application de la théorie des perturbations sont divergentes, en raison de *petits dénominateurs*. On appelle ainsi des expressions, figurant au dénominateur des termes successifs des séries qui expriment la solution, qui sont des combinaisons linéaires à coefficients entiers des fréquences qui apparaissent dans le mouvement du système complètement intégrable dont le système considéré est une perturbation.

Soit (M, Ω, H) un système hamiltonien complètement intégrable sur une variété symplectique de dimension $2n$. Comme nous l'avons vu dans le paragraphe 1, sous une hypothèse de compacité, chaque courbe intégrale du système reste, pour toutes les valeurs du temps, sur un tore de dimension n , et a sur ce tore une expression très simple : chaque coordonnée angulaire varie linéairement en fonction du temps ; en particulier (puisque les angles sont définis modulo 2π) chaque coordonnée angulaire est une fonction périodique du temps. On exprime cette propriété en disant que ces mouvements sont *quasi-périodiques*. En 1954, A.N. Komogorov a annoncé un résultat tout à fait remarquable, obtenu par des moyens mathématiques indépendants de la théorie des perturbation : sous une hypothèse de non dégénérescence (précisée ci-dessous), si on perturbe légèrement le hamiltonien d'un système complètement intégrable, de nombreux mouvements du système perturbé restent quasi-périodiques ; les courbes intégrales correspondantes restent, pour toutes les valeurs du temps, sur des tores de dimension n fixes, et chacune des variables angulaires sur ces tores est une fonction périodique du temps. Dans toute partie bornée de M , la mesure de l'ensemble des données de Cauchy pour lesquelles le mouvement n'est pas quasi-périodique tend vers zéro lorsque la perturbation du hamiltonien tend elle-même vers zéro.

La démonstration complète de ce théorème a été donnée par V.I. Arnol'd en 1963 pour les hamiltoniens analytiques et étendue aux hamiltoniens différentiables par J. Moser en 1966. Le théorème résumant ces résultats est connu sous le nom de *théorème de Kolmogorov-Arnol'd-Moser* (en abrégé *théorème KAM*).

Avant de donner l'énoncé précis de ce théorème, fixons les notations et introduisons quelques notions nouvelles.

3.1. Hypothèses et notations. — Puisqu'on s'intéresse à une perturbation d'un système complètement intégrable, et qu'on sait que les systèmes complètement intégrables admettent des variables actions-angles, on se place d'emblée dans les variables actions-angles du système non perturbé. L'espace des phases est le fibré cotangent à un tore de dimension n muni de sa forme symplectique canonique ; et on utilise sur ce fibré des variables actions-angles $I_1, \dots, I_n, \gamma_1, \gamma_n$. Les γ_i sont les coordonnées angulaires sur le tore et les actions I_i les coordonnées dans les fibres, relativement aux bases de ces fibres naturellement associées aux angles. La forme

symplectique a pour expression

$$\Omega = \sum_{i=1}^n dI_i \wedge d\gamma_i.$$

Attention à une erreur à ne pas commettre : les 1-forme $d\gamma_i$, malgré la notation utilisée pour les désigner, sont fermées mais non exactes, car les angles γ_i sont des fonctions à valeurs dans le cercle trigonométrique $\mathbb{R}/2\pi\mathbb{Z}$, non des fonctions à valeurs réelles.

Le hamiltonien du système est de la forme

$$H = H_0 + \epsilon H_1,$$

où H_0 est un hamiltonien complètement intégrable, fonction des seules variables d'action, $I = (I_1, \dots, I_n)$. Quant à H_1 , c'est une fonction de toutes les variables, actions et angles, et ϵ est un paramètre positif petit, qu'on fera tendre vers 0.

Dans la suite on appellera *système non perturbé* le système ayant H_0 pour hamiltonien, et *système perturbé* le système ayant $H = H_0 + \epsilon H_1$ pour hamiltonien.

3.2. Tores résonants et non résonants du système non perturbé. — Les équations de Hamilton pour le système non perturbé s'écrivent

$$\begin{cases} \frac{d\gamma_i}{dt} = \omega_i = \frac{\partial H_0}{\partial I_i}, \\ \frac{dI_i}{dt} = 0, \end{cases} \quad 1 \leq i \leq n.$$

Les ω_i sont des fonctions des variables d'action $I = (I_1, \dots, I_n)$. Les équations de Hamilton ont pour solutions

$$\begin{cases} \gamma_i(t) = \gamma_i(0) + \omega_i(I(0))t, \\ I_i(t) = I_i(0), \end{cases} \quad 1 \leq i \leq n.$$

Pour chaque $c = (c_1, \dots, c_n) \in \mathbb{R}^n$, la sous-variété d'équations $I = c$ est un tore de dimension n , sur lequel $\gamma_1, \dots, \gamma_n$ sont des coordonnées angulaires. Une courbe intégrale des équations de Hamilton issue d'un point de ce tore reste sur ce tore pour toutes les valeurs du temps, et chaque angle γ_i est une fonction périodique du temps de période $\frac{2\pi}{\omega_i(c)}$. Les $\omega_i(c)$ sont appelées *fréquences angulaires*. Sur chaque tore d'équations $I = c$, ce sont des constantes.

Le tore d'équations $I = c$ est dit *non résonant* si les fréquences angulaires $\omega_i(c)$ sont rationnellement indépendantes, c'est-à-dire si l'équation

$$\sum_{i=1}^n k_i \omega_i(c) = 0, \quad (*)$$

dans laquelle l'inconnue $k = (k_1, \dots, k_n)$ est élément de \mathbb{Z}^n , n'a pas d'autre solution que la solution nulle.

On montre que chaque courbe intégrale située sur un tore non résonant est partout dense sur ce tore; autrement dit, l'adhérence de cette courbe est le tore entier.

On dit que le tore d'équations $I = c$ a une *résonance d'ordre k* (où k est un entier vérifiant $1 \leq k \leq n-1$) si l'équation (*) ci-dessus admet k solutions indépendantes (c'est-à-dire dont aucune combinaison linéaire à coefficients entiers non tous nuls n'est identiquement nulle), et n'admet pas $k+1$ solutions indépendantes. Le tore est dit *complètement résonant* s'il admet une résonance d'ordre $n-1$. Dans ce cas, les ω_i ($1 \leq i \leq n$) sont tous multiples entiers d'un même réel ω , et toutes les courbes intégrales situées sur le tore sont périodiques de période $\frac{2\pi}{\omega}$.

On montre que sur un tore ayant une résonance d'ordre k , l'adhérence de chaque courbe intégrale du système hamiltonien non perturbé est un sous-tore de dimension $n-k$. En particulier, en faisant $k = n-1$, on retrouve le fait que l'adhérence de chaque courbe intégrale est un tore de dimension 1, c'est-à-dire un cercle, ce qui est en accord avec le fait que chacune de ces courbes est périodique.

3.3. Définition. — *Avec les hypothèses et notations précisées ci-dessus, on dit que le système hamiltonien non perturbé est non dégénéré si*

$$\det \left(\frac{\partial \omega_i}{\partial I_j} \right) = \det \left(\frac{\partial^2 H_0}{\partial I_i \partial I_j} \right) \neq 0.$$

La propriété importante des systèmes non dégénérés est la suivante : pour un tel système, dans un voisinage arbitrairement petit d'un tore d'équations $I = c$ quelconque, il existe des tores non résonants (et d'ailleurs aussi des tores résonants). Cela résulte du fait que les irrationnels (ainsi d'ailleurs que les rationnels) forment une partie dense de \mathbb{R} . Lorsque le système hamiltonien non perturbé considéré est non dégénéré on peut, au voisinage de chaque tore, utiliser pour variables les fréquences angulaires ω_i et les angles γ_i , au lieu d'utiliser les actions I_i et les angles γ_i .

3.4. Théorème de Kolmogorov-Arnol'd-Moser. — *On considère un système hamiltonien, sur le fibré cotangent au tore de dimension n muni de sa forme symplectique canonique, de hamiltonien*

$$H = H_0 + \epsilon H_1,$$

où H_0 ne dépend que des actions. On suppose le système non perturbé (de hamiltonien H_0) non dégénéré. Pour ϵ assez petit, beaucoup de courbes intégrales du système perturbé restent, pour toutes les valeurs du temps, sur des tores de dimension n voisins des tores invariants du système non perturbé (c'est-à-dire déduits des tores invariants du système non perturbé par de petites déformations). Sur chacun des tores invariants du système perturbé, chaque courbe intégrale est

dense et chaque mouvement est quasi-périodique, avec n fréquences angulaires rationnellement indépendantes. Dans toute partie bornée de l'espace des phases, la mesure de l'ensemble des points qui ne sont pas sur un tore invariant tend vers 0 lorsque ϵ tend vers 0.

Pour des indications sur la démonstration et des compléments, le lecteur pourra consulter le livre de V.I. Arnol'd et A. Avez *Problèmes ergodiques de la mécanique classique*, Gauthier-Villars, Paris, 1967, le livre de V.I. Arnol'd *Méthodes mathématiques de la mécanique classique*, éditions Mir, Moscou, 1974, ou sa traduction anglaise enrichie *Mathematical methods of classical mechanics*, second edition, Springer-Verlag, New York, 1978.

Une autre approche de la stabilité des courbes intégrales des systèmes hamiltoniens qui sont des perturbations de systèmes complètement intégrables, de hamiltonien de la forme $H = H_0 + \epsilon H_1$, avec H_0 complètement intégrable, a été initiée par N.N. Nekhoroshev en 1971. Elle est basée sur une étude fine permettant de majorer les variations des actions I_i au cours du mouvement du système perturbé. Cette approche a permis d'établir des résultats de stabilité pendant un temps fini de la forme suivante : pour une variation maximale admise des actions, il existe un seuil $\epsilon_0 > 0$ tel que, lorsque $\epsilon < \epsilon_0$, la variation des actions au cours du mouvement du système perturbé reste inférieure à la variation maximale admise pendant un temps fini, d'autant plus grand que ϵ est plus petit. C'est un domaine où des recherches sont toujours très activement menées (travaux de P. Lochak, L. Niederman, J.-P. Marco, D. Sauzin, J. Cresson, ...).