TRADUIT DU RUSSE

ANALYSE MATHÉMATIQUE

G. CHILOV

FONCTIONS DE PLUSIEURS VARIABLES RÉELLES

1ère ET 2e PARTIES

ÉDITIONS DE MOSCOU

Georges Chilov, docteur ès sciences, professeur à l'Université Lomonossov de Moscou, était un savant soviétique éminent qui se spécialisait dans l'analyse mathématique.

Ses travaux dans cette matière constituent un apport important au développement des mathématiques.

Vingt monographies et manuels, ainsi qu'un grand nombre d'articles publiés dans des revues mathématiques appartiennent à sa plume.

Dans les publications du prof. Chilov, la profondeur de l'investigation et le sentiment de l'esprit moderne se trouvent heureusement associés à l'accessibilité de l'exposé.

Georges Chilov enseignait avec éclat à l'Université de Moscou. Il a ainsi formé une pléiade de mathématiciens soviétiques dont quelques savants de renom.

Les livres écrits par le prof. Chilov sont très connus et ont été traduits en plusieurs langues européennes.

E. GRANIRER


ИЗДАТЕЛЬСТВО
МИР

Г. Е. ШИЛОВ

МАТЕМАТИЧЕСКИЙ АНАЛИЗ

ФУНКЦИИ НЕСКОЛЬКИХ ВЕЩЕСТВЕННЫХ ПЕРЕМЕННЫХ

части 1-2

ИЗДАТЕЛЬСТВО «НАУКА» МОСКВА

G. CHILOV

Analyse mathématique

FONCTIONS DE PLUSIEURS VARIABLES RÉELLES

1ère et 2e parties

Editions Mir • Moscou

Traduit du russe par
VITALI KHARINE

На французском языке

© Traduction française • Editions Mir · 1975

Avant-propos

Le présent livre peut être considéré comme suite de l'ouvrage du même auteur « Analyse mathématique (fonctions d'une variable) » paru précédemment dans les Editions Mir (Moscou, 1973). Les principes de base restent les mêmes: conformément à l'esprit moderne, l'analyse mathématique se présente comme un système hautement organisé des structures, des abstractions de différents degrés, qui sont étroitement liées entre elles et avec des applications. Dans les ouvrages scientifiques, tels que « Eléments de mathématiques » de N. Bourbaki, la réalisation successive de ce principe conduit à un exposé strictement déductif de la théorie; dans les livres pédagogiques, l'exposé inductif est souvent préféré, puisque permet au lecteur de suivre la formation des conceptions de plus en plus abstraites et de voir la nécessité des généralisations. C'est ainsi qu'est conçu notre livre. Formellement, les chapitres « L'intégration et la dérivation » et « Géométrie différentielle classique » ne sont pas indispensables pour notre cours — les résultats principaux de ces chapitres s'obtiennent comme conséquences spéciales des théories plus générales (c'est ce qu'on voit par la suite dans le livre); nous avons tout de même décidé d'en faire précéder les théories plus abstraites, afin que le lecteur soit amené à la nécessité de certaines notions générales, telles que, par exemple, espace riemannien ou forme différentielle sur une variété, et qu'il soit prêt à s'en servir dans les applications. La composition de ces chapitres vise le même but: par exemple, la théorie des surfaces développables qui est un des sujets favoris dans les cours de la géométrie différentielle devient pour nous un épisode de passage, tandis que les rôles principaux sont joués par les coefficients de connexion de Riemann-Christoffel.

De même que les « Fonctions d'une variable », le présent livre se compose de trois parties. Les deux premières, « Calcul différentiel et intégral » et « Des espaces vectoriels aux variétés », sont devant le lecteur; la troisième, « Analyse sur les variétés », paraîtra plus tard.

Le chapitre 1 expose la théorie de la dérivation des fonctions d'un nombre fini ou infini de variables. L'importance du calcul différentiel des fonctions de plusieurs variables est évidente ; mais il s'avère que le calcul différentiel des fonctions d'une infinité de variables (plus exactement, du point d'un espace normé) peut lui aussi très bien servir aux besoins de l'analyse classique, par exemple à l'étude des solutions des équations différentielles ordinaires (et l'étude des extrémums des fonctions de ce genre débouche sur les problèmes du calcul des variations).

Le chapitre 2 est consacré aux dérivées d'ordre supérieur. De même qu'en cas de fonctions d'une variable, les dérivées d'ordre supérieur permettent de décrire d'une façon plus exacte (que la dérivée première) le comportement de la fonction au voisinage d'un point donné. Il en existe également maintes autres applications. Il se trouve, par exemple, que le problème classique de compatibilité d'un système complet d'équations à dérivées partielles devient bien transparent après l'avoir interprété comme problème de résolubilité d'une équation différentielle ordinaire (mais avec une variable indépendante multidimensionnelle !), qui ne demande que la symétrie de la dérivée deuxième de la solution exprimée par le second membre de l'équation.

C'est la théorie de l'intégration que l'on construit dans le chapitre 3. La partie abstraite de la théorie représente une généralisation de l'intégrale de Riemann au cas de l'« espace chargé », i.e. d'un espace métrique muni d'une mesure additive. Comme dans le livre « Fonctions d'une variable », nous ne faisons pas intervenir l'intégrale de Lebesgue, puisque nous n'avons à intégrer que les fonctions continues ou les fonctions à un « petit » ensemble de points de discontinuité ; nous nous passons donc d'une mesure σ-additive. En tant qu'applications, nous envisageons la théorie de la mesure des volumes dans l'espace multidimensionnel et celle de la mesure des surfaces (dans ses différentes conceptions) ; une certaine attention est prêtée aux intégrales impropres de volume et de surface.

L'analyse vectorielle, dans le chapitre 4, est utilisée comme langage dans lequel on exprime les liens qui existent entre les opérations intégrales et différentielles sur les fonctions de plusieurs variables. Les opérations différentielles principales (gradient, divergence, rotationnel) sont traitées d'un point de vue commun, comme densités de certaines fonctions additives du domaine. La théorie est poussée jusqu'au problème inverse (restitution d'un champ d'après sa divergence et son rotationnel). Une partie essentielle du chapitre se déroule dans l'espace tridimensionnel, vu le caractère spécifique de la définition du rotationnel.

La deuxième partie du livre est ouverte par le chapitre 5 « Géométrie différentielle classique » dont l'objet principal est la connexion engendrée sur une surface par la métrique de l'espace eucli-

dien qui la contient, de même qu'aux lignes géodésiques et à la translation, ces dernières étant des attributs de la connexion. En développant ces constructions on voit peu à peu qu'il n'est pas nécessaire d'emprunter la métrique d'une surface à l'espace euclidien qui la contient ; ainsi, la courbure constante se réalise sur le plan, la sphère et l'hyperboloïde, la métrique de ce dernier étant induite non pas par l'espace euclidien qui le contient mais par une forme quadratique ayant un carré négatif. Ceci ouvre la voie directe vers les espaces riemanniens (chapitre 6). Il s'avère que, dans le cas général d'un espace riemannien quelconque, la courbure constante est toujours réalisée sur le plan (multidimensionnel), la sphère (multidimensionnelle) et l'hyperboloïde (multidimensionnel), la métrique de ce dernier étant induite par une forme quadratique avec un carré négatif, qui est positive sur l'hyperboloïde même.

Dans le chapitre 7, la place centrale appartient aux formes différentielles du degré arbitraire ; elles servent à formuler les théorèmes intégraux du type Stokes, aussi bien qu'à poser correctement le problème inverse généralisé de l'analyse vectorielle : précisément, ce problème consiste à retrouver une forme d'après sa différentielle et sa codifférentielle.

La numération reste la même que dans les livres précédents ; par exemple, le symbole *6.23b* signifie « chapitre 6, paragraphe 2, point 3, sous-point *b* ». Ces symboles sont mis dans les titres de colonnes de chaque page, ce qui permet de trouver rapidement le passage voulu. On rencontre souvent dans ce livre les références aux livres de l'auteur « Analyse mathématique (fonctions d'une variable) » (tomes I, II, Moscou, 1973) et « Математический анализ (конечномерные линейные пространства) » (Moscou, 1969) ; ces références sont désignées par les symboles analogues précédés par la lettre *U* (pour le premier des livres mentionnés) et *L* (pour le deuxième).

L'auteur

PREMIÈRE
PARTIE

Calcul différentiel et intégral

Dérivées du premier ordre

Dans le présent chapitre, nous avons pour objet le calcul différentiel des fonctions dont la variable indépendante est à plusieurs dimensions; seules les dérivées du premier ordre sont, pour le moment, considérées. L'idée principale est celle de linéarisation, qui consiste à extraire de l'accroissement d'une fonction sa partie linéaire principale, ce qui rend élémentaire l'étude locale de la fonction à des infiniment petits d'ordre supérieur près. Ce qu'on appelle fonctions dérivables sont juste les fonctions admettant un tel procédé. L'étude de simples propriétés des fonctions dérivables par la méthode de linéarisation se fait d'une même manière pour la fonction d'une ou de plusieurs variables réelles et même pour celle d'une infinité de variables (plus exactement, pour une fonction d'un point d'un espace normé). Certains écarts du cas d'une variable n'apparaissent que dans l'énoncé du théorème de la moyenne (§ 1.4). Plus tard, un théorème manifestement nouveau, qui n'avait pas lieu dans la théorie des fonctions d'une variable, commence à jouer un rôle primordial; il s'agit du théorème sur la fonction implicite (§ 1.5). Sans surestimer sa valeur, on peut dire que c'est le théorème chef du calcul différentiel des fonctions de plusieurs variables — tant ses applications sont importantes, que ce soit le cas de dimension finie ou le cas général (de dimension infinie) : dépendance de la solution d'une équation différentielle ordinaire d'un paramètre (§ 1.8), structure locale d'une fonction dérivable (§ 1.6), théorie d'extrémum lié (§ 1.7) et bien d'autres que nous verrons dans les chapitres suivants.

§ 1. 1. Fonctions continues

Avant de développer le calcul différentiel des fonctions de plusieurs variables réelles, il est bon de rappeler certaines définitions de base se rapportant à la théorie des fonctions continues.

1.11. Soit $y = f(x)$ une fonction définie sur un ensemble X et à valeurs dans un ensemble Y. Par la suite, on utilisera l'une des manières suivantes de noter ce fait :

$$y = f(x), \quad y : X \to Y,$$
$$y = f(x) : \quad X \to Y,$$
$$y = f(x) \qquad (X \to Y),$$
$$x \to y = f(x),$$
$$x \to f(x),$$

les deux dernières lignes réservées au cas où X et Y sont définis dans le texte. S'il faut indiquer que la fonction $f(x)$ est définie sur un sous-ensemble $E \subset X$ et prend ses valeurs dans un sous-ensemble $F \subset Y$, nous noterons également:

$$y = f(x): \quad (E \subset X) \to (F \subset Y),$$

soit, en omettant les parenthèses:

$$y = f(x): \quad E \subset X \to F \subset Y.$$

La fonction $f(x)$ est dite *numérique*, lorsque $Y \subset R_1$ (droite réelle). Une telle fonction est également dite *réelle* (plus exactement, *à valeurs réelles*). Si $Y \not\subset R_1$, on prend pour règle d'appeler $f(x)$ *application*. Lorsque Y est un espace vectoriel (*U12.11*), par exemple $Y = R_n$ (espace réel n-dimensionnel), la fonction $f(x): X \to Y$ est dite *vectorielle*. Si X est une partie de R_1, alors $f(x)$ s'appelle *fonction d'une variable réelle*. Si X est un domaine dans R_n, $f(x)$ s'appelle *fonction de n variables réelles*; normalement, on choisit pour telles les coordonnées $x_1, \ldots, x_n$ d'un point x par rapport à une certaine base dans R_n.

Pour se faire une idée géométrique sur telle ou telle fonction numérique d'une variable réelle, nous avons l'habitude de tracer son graphique en portant, dans le sens de l'axe des y, sa valeur $f(x)$ pour tout point x de son domaine de définition. En cas de fonction numérique $f(x_1, x_2)$ de deux variables réelles, on peut, quel que soit un point (x_1, x_2) de l'ensemble X sur le plan R_2, porter la valeur $f(x_1, x_2)$ dans le sens du troisième axe, celui des y. Le graphique d'une fonction d'une variable est, normalement, une courbe; celui d'une fonction numérique de deux variables représente, au moins pour les fonctions assez simples, une surface que l'on peut dessiner en estimant les règles de la perspective.

Exemples. **a.** Une fonction du premier degré $y = k_1x_1 + k_2x_2 + b$ a pour graphique un *plan*. Les nombres k_1 et k_2 s'appellent coefficients angulaires du plan, leur interprétation géométrique est claire de la fig. 1.1-1.

b. Le graphique de la fonction du second degré $y = x_1^2 + x_2^2$ est un paraboloïde de révolution (pour $y = a_1x_1^2 + a_2x_2^2$, $a_1 > 0$, $a_2 > 0$ ce sera un *paraboloïde elliptique*) montré fig. 1.1-2.

c. La fonction du second degré $y = x_2^2 - x_1^2$ a pour graphique une surface en forme de selle appelée *paraboloïde hyperbolique*; si l'axe des y est dirigé vers le haut, alors les sections verticales de la surface sont des paraboles et les sections horizontales des hyperboles (fig. 1.1-3).

En cas de fonction de trois variables ou plus on peut, bien entendu, appeler « graphique » de la fonction l'ensemble des points $(x_1, \ldots, x_n, f(x_1, \ldots, x_n))$, $x \in X$ de l'espace $(n+1)$-dimen-

sionnel, ce qui est conforme à la définition générale du graphique donnée dans *U2.83*; cependant, il est un peu difficile de s'en faire une idée géométrique; dans des cas pareils, l'évidence géométrique

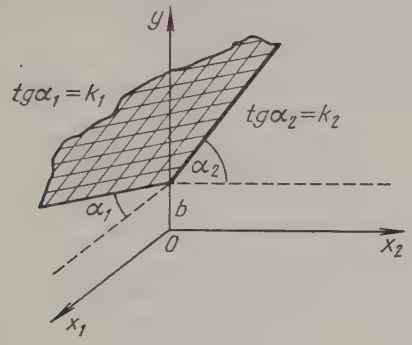


Fig. 1.1-1

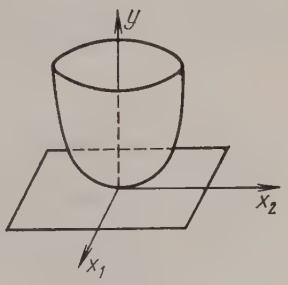


Fig. 1.1-2

doit céder sa place à la logique. Les graphiques ne sont d'ailleurs pas indispensables dans le cas de deux variables ni même d'une seule, bien que personne n'en nie l'utilité.

d. On peut parfois se figurer géométriquement une fonction en considérant ses lignes (ou surfaces) de niveau. Une *ligne* (*surface*) *de niveau d'une fonction* $y = f(x)$ *est le lieu géométrique des points où la fonction prend une même valeur* $y = y_0$. Ainsi, les lignes de niveau de la fonction $f(x) = \rho(x, a) = |x - a|$ $(R_2 \to R_1)$ sont (fig. 1.1-4) les circonférences centrées au point a (de même que le point a lui-même où la fonction prend la valeur nulle). Les lignes de niveau de la fonction $f(x) = \rho(x, a) + \rho(x, b)$ $(R_2 \to R_1)$ représentent (fig. 1.1-5) les ellipses de foyers aux points a et b (et le segment joignant ces points); celles de la fonction $f(x) = \rho(x, a) - \rho(x, b)$ (fig. 1.1-6) sont les hyperboles de foyers a et b ainsi que leur axe de symétrie et deux demi-droites; celles de la fonction $f(x) = \rho(x, a) \cdot \rho(x, b)$ $(R_2 \to R_1)$ sont une famille (fig. 1.1-7) d'ovales de Cassini dont une lemniscate de Bernoulli (voir exercice 6); enfin, celles de la fonction $f(x) = \rho(x, a)/\rho(x, b)$ $(R_2 \to R_1)$ forment une famille de circonférences de centres sur la droite ab complétée de l'axe de symétrie des points a et b (fig. 1.1-8). Les surfaces de niveau des mêmes fonctions définies dans R_3 sont engendrées par la rotation des courbes correspondantes autour de l'axe ab.

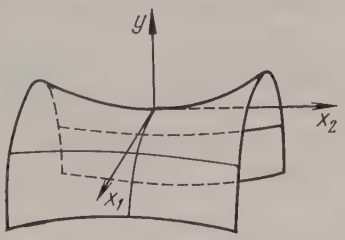


Fig. 1.1-3

c. Nous avons déjà dit dans *U2.83* que, dans le cas général, le graphique d'une fonction $y = f(x) : E \subset X \to Y$ est l'ensemble des points $\{x, f(x)\}$ dans le produit cartésien de X et Y. Mais il est

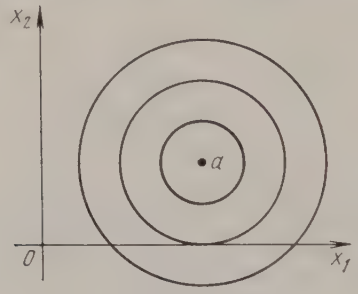


Fig. 1.1-4

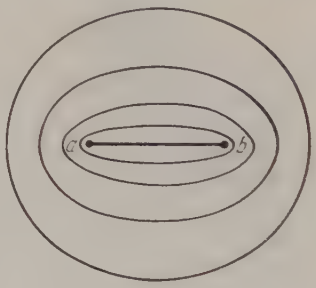


Fig. 1.1-5

difficile de s'en faire une idée géométrique. Si $Y = X$, il s'avère parfois utile de lier à la fonction $y = f(x)$ un *champ vectoriel* : chaque point $x \in E$ est l'origine d'un vecteur (d'une flèche) dont l'extrémité

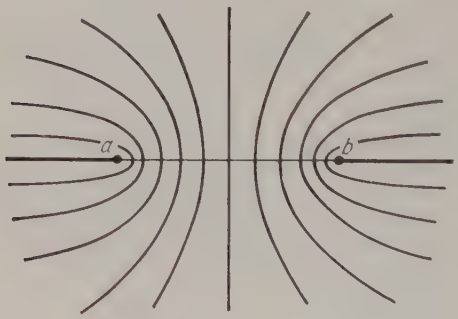


Fig. 1.1-6

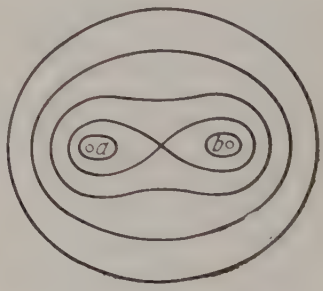


Fig. 1.1-7

est le point $x + f(x)$. Par exemple, la figure 1.1-9 montre le champ vectoriel correspondant à la fonction

$$f(z) = \frac{i}{2} z : R_2 \to R_2.$$

1.12. Rappelons maintenant la notion de continuité (*U5.11*).

a. Pour que l'on puisse discuter la continuité d'une fonction $f(x)$ $(X \to Y)$, il faut supposer que l'ensemble X où la fonction est définie, de même que l'ensemble Y où elle prend ses valeurs, soient des espaces métriques (*U3.11*). Ceci fait, la fonction $y = f(x)$ est dite *continue pour* $x = a$ si, quel que soit $\varepsilon > 0$, il existe $\delta > 0$

tel que, une fois $\rho_X(x, a) < \delta$, $x \in X$, on a $\rho_Y(f(x), f(a)) < \varepsilon$. Ici ρ_X désigne la métrique de l'espace X, ρ_Y celle de Y. (En général, nous désignons la métrique, i.e. la distance entre les points d'un espace métrique M, par ρ, et s'il faut indiquer l'espace, par ρ_M.) Il y a une autre définition équivalente : la fonction $f(x)$ est *continue pour* $x = a$, lorsque la relation $f(x_m) \to f(a)$ a lieu pour toute suite $x_1, \ldots, x_m, \ldots$ de points de l'ensemble X qui converge

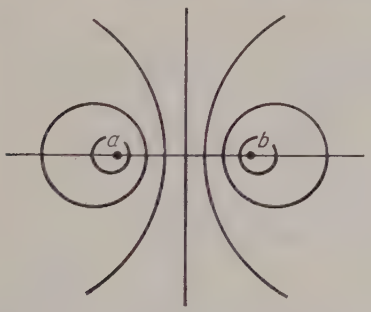


Fig. 1.1-8

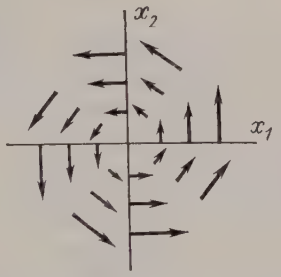


Fig. 1.1-9

vers le point a. La démonstration de l'équivalence des définitions formulées est donnée dans *U5.11*, et nous n'y insistons plus.

Une fonction $f(x)$ $(X \to Y)$ continue à chaque point de l'ensemble X est dite *continue sur l'ensemble* X.

b. Le plus simple exemple d'une fonction continue est fourni par la fonction *constante* $f(x)$ $(X \to Y)$ qui, pour tout $x \in X$, prend une même valeur $y_0 \in Y$.

c. Un autre exemple : la fonction numérique $f(x) = \rho(x, a)$ $(X \to R_1)$ où a est un point fixe de l'espace X. La continuité résulte de l'inégalité du triangle (cf. *U5.12b*).

d. La fonction $f(x) \equiv x$ $(X \to X)$, qui à tout élément x de l'espace métrique X fait correspondre le même élément x, est le plus simple exemple d'une fonction continue non constante.

e. Fonctions de plusieurs variables. Soient $X_1, \ldots, X_n$ des ensembles ; alors l'ensemble X de toutes les n-tuples de la forme

$$x = \{x_1, \ldots, x_n\}, \quad x_1 \in X_1, \ldots, x_n \in X_n$$

est appelé *produit cartésien* des ensembles $X_1, \ldots, X_n$ et noté $X_1 \times X_2 \times \ldots \times X_n$. N'importe quelle fonction $y = f(x) : X \to Y$ peut être traitée comme fonction de n variables $x_1, \ldots, x_n$.

Soient $X_1, \ldots, X_n$ et Y des espaces métriques. Il est alors possible, pour une fonction $y = f(x) : X \to Y$, de définir sa « continuité pour $x = a = \{a_1, \ldots, a_n\}$ par rapport à l'ensemble des variables » ; cela veut dire que, quel que soit $\varepsilon > 0$, il existe $\delta > 0$

tel que les inégalités $\rho(x_1, a_1) < \delta, \ldots, \rho(x_n, a_n) < \delta$ impliquent l'inégalité $\rho(f(x), f(a)) < \varepsilon$. La continuité de la fonction $f(x)$ par rapport à l'ensemble $x_1, \ldots, x_n$ sur tout l'espace X est définie de la façon naturelle. La continuité par rapport à l'ensemble pour $x = a$ peut être interprétée comme continuité ordinaire, à condition que l'espace X soit muni d'une métrique convenable; on peut, par exemple, se servir de la définition:

$$\rho(\{x_1, \ldots, x_n\}, \{y_1, \ldots, y_n\}) = \\ = \max\{\rho(x_1, y_1), \ldots, \rho(x_n, y_n)\}. \quad (1)$$

Parfois on utilise d'autres métriques sur X, par exemple

$$\rho(\{x_1, \ldots, x_n\}, \{y_1, \ldots, y_n\}) = \sum_{i=1}^{n} \rho(x_i, y_i) \quad (2)$$

ou bien

$$\rho(\{x_1, \ldots, x_n\}, \{y_1, \ldots, y_n\}) = \sqrt{\sum_{i=1}^{n} \rho^2(x_i, y_i)}. \quad (3)$$

Dans tous ces cas, il est aisé de voir que la formule en question ((1), (2) ou (3)) définit effectivement une métrique sur X, et que la continuité de $f(x)$ par rapport à l'ensemble $x_1, \ldots, x_n$ (cette continuité étant définie indépendamment de la métrique de X) équivaut à la continuité de $f(x)$ par rapport à la métrique proposée.

1.13. Propriétés des fonctions continues.

a. Voici un théorème important:

Théorème sur la continuité d'une fonction composée (*U5.15*): *Soit* $y = f(x)$ $(X \to Y)$ *une fonction définie sur un espace métrique* X *à valeurs dans un espace métrique* Y *et continue en un point* $x = a$; *soit* $g(y)$ $(Y \to Z)$ *une fonction définie sur l'espace métrique* Y, *à valeurs dans un espace métrique* Z *et continue pour* $y = f(a)$. *Alors la fonction composée* $\Phi(x) = g[f(x)]$ $(X \to Z)$ (qui est nécessairement définie dans un voisinage du point a) *est continue pour* $x = a$.

La fonction $g[f(x)]$ s'appelle *composée* des fonctions $f(x)$ et $g(y)$.

Comme la fonction numérique $\rho(x, a)$ $(X \to R_1)$ est continue sur l'espace métrique X (on l'a vu dans *1.12c*), toute fonction de la forme $g[\rho(x, a)]$, où $g(y)$ est une fonction continue pour $y \geqslant 0$ de la variable réelle y, l'est aussi d'après le théorème précédent.

b. Si les fonctions $f_1(x), \ldots, f_m(x), \ldots$ sont définies sur un même ensemble X, prennent leurs valeurs dans un espace métrique Y et forment une suite convergente sur X, alors on peut définir sur le même ensemble X la fonction limite $f(x) = \lim_{m \to \infty} f_m(x)$: $X \to Y$. Si X est un espace métrique, la convergence de $f_m(x)$ vers $f(x)$ est uniforme et les fonctions $f_m(x)$ $(m = 1, 2, \ldots)$ sont continues, alors $f(x)$ est continue elle aussi (*U5.96*).

c. Une fonction $y = f(x)$ $(X \to Y)$ à valeurs dans un espace métrique Y est dite *bornée* (sur X) si les nombres $\rho_Y(f(x'), f(x''))$ forment un ensemble borné lorsque x' et x'' parcourent X. L'ensemble $Y(X)$ de toutes les fonctions bornées continues, qui sont définies dans un espace métrique X et prennent les valeurs dans un espace métrique Y, devient lui-même espace métrique si l'on pose

$$\rho_{Y(X)}(f, g) = \sup_{x \in X} \rho_Y(f(x), g(x)).$$

Si Y est complet, $Y(X)$ l'est aussi (*U12.23f*).

d. Soient X et Y des espaces métriques; une fonction $y = f(x)$ $(X \to Y)$ est dite *uniformément continue* (sur X) si, pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que la relation $\rho_X(x', x'') < \delta$ implique $\rho_Y(f(x'), f(x'')) < \varepsilon$ (*U5.17b*). Sur un compact (espace métrique compact) X (*U3.91a*), toute fonction continue est uniformément continue (*U5.17b*). L'ensemble de toutes les valeurs d'une fonction continue sur le compact X est lui-même compact dans Y (*U5.16a*), donc borné.

Si X n'est pas compact, il existe des fonctions continues mais non bornées, ainsi que des fonctions continues et bornées mais non uniformément continues sur X (exercices 29-30).

e. Voici une méthode de construction d'applications continues que nous utiliserons par la suite. Soient X, Y, Z des espaces métriques, $z = \Phi(x, y)$: $X \times Y \to Z$ une fonction bornée et uniformément continue sur $X \times Y$, $f(x)$ $(X \to Y)$ une fonction continue; alors la fonction $\Phi(x, f(x))$ $(X \to Z)$ est bornée et, en vertu de *a*, continue. Il est donc défini l'application Ff, $(Ff)(x) = \Phi(x, f(x))$, de l'espace $Y(X)$ dans l'espace $Z(X)$. Montrons qu'elle est continue. En effet, quel que soit $\bar{f} = \bar{f}(x) \in Y(X)$, nous avons

$$\rho_{Z(X)}(F\bar{f}, Ff) = \sup_{x \in X} \rho_Z(\Phi(x, \bar{f}(x)), \Phi(x, f(x))). \qquad (1)$$

Pour $\varepsilon > 0$ donné, par l'hypothèse de la continuité uniforme de la fonction $\Phi(x, y)$, on peut trouver $\delta > 0$ tel que, une fois $\rho_{Y(X)}(\bar{f}, f) < \delta$, i.e. $\sup_{x \in X} \rho_Y(\bar{f}(x), f(x)) < \delta$, le second membre de (1) est inférieur à ε, ce qu'il nous faut.

1.14. Fonctions linéaires.

a. Envisageons l'espace vectoriel n-dimensionnel R_n métrisé à l'aide d'une norme (*U12.31*) et y choisissons une base. En faisant correspondre à tout point $x \in R_n$ sa coordonnée x_k d'indice k fixe $(k = 1, \ldots, n)$, nous aboutissons à une fonction numérique continue et, qui plus est, linéaire (*U12.15*) sur R_n. La plus générale forme d'une fonction numérique linéaire sur R_n est $f(x) = \sum_{k=1}^{n} c_k x_k$, où c_k $(k = 1, \ldots, n)$ sont des nombres fixes.

b. Comme nous l'avons déjà vu (*U12.71*), les plus simples fonctions numériques continues sur un espace vectoriel normé, à part les constantes, sont les fonctionnelles linéaires.

Nous en connaissons certains exemples. Ainsi, dans l'espace $R^s(a, b)$ des fonctions réelles continues sur l'intervalle $[a, b]$, il est défini la fonctionnelle linéaire $x(t) \to F(x)$ de la forme

$$F(x) = \int_a^b D(t)\, x(t)\, dt,$$

$D(t)$ étant une fonction continue donnée (*U12.71i*). Dans un espace hilbertien réel H (*U12.41*), le produit scalaire

$$F(x) = (f, x),$$

où f est un vecteur fixe de l'espace H, sert d'exemple de fonctionnelle linéaire.

c. D'une façon générale, parmi toutes les fonctions continues agissant d'un espace normé X dans un autre espace normé Y, les plus simples, les constantes mises à part, sont les opérateurs linéaires continus (*U12.71*), i.e. les fonctions continues $A(x): X \to Y$ vérifiant la condition

$$A(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 A x_1 + \alpha_2 A x_2,$$

quels que soient x_1, x_2 de X et les constantes α_1, α_2.

Quels que soient deux opérateurs linéaires continus $A: X \to Y$, $B: X \to Y$ et deux nombres α, β, la combinaison linéaire $\alpha A + \beta B$ est définie en tant qu'opérateur linéaire continu de X dans Y qui agit d'après la formule

$$(\alpha A + \beta B)\, x = \alpha A(x) + \beta B(x).$$

L'espace vectoriel de *tous* les opérateurs linéaires continus $A: X \to Y$ muni de la norme $\| A \| = \sup_{|x| \leqslant 1} | A(x) |$ est un espace *de Banach* (i.e. normé et complet). On le désigne par $L(X, Y)$.

Si $X = R_1$, alors $L(X, Y) = L(R_1, Y)$ s'identifie naturellement avec Y: à tout $a \in Y$, il correspond l'opérateur $A \in L(R_1, Y)$ défini par la formule $At = ta$ $(t \in R_1)$, et tout opérateur $A \in L(R_1, Y)$ agit d'après la formule $At = At \cdot 1 = t \cdot A(1) = ta$, où $a = A(1)$. En abrégé, l'espace $L(X, X)$ est noté $L(X)$.

d. Signalons quelques opérateurs linéaires en rapport avec la décomposition en somme directe. Rappelons que, par définition, un espace vectoriel X est la somme directe de ses sous-espaces $X_1, \ldots, X_n$ si, pour tout $x \in X$, on a la décomposition

$$x = x_1 + \ldots + x_n, \qquad x_1 \in X_1, \ldots, x_n \in X_n \qquad (1)$$

et que cette décomposition est unique, c'est-à-dire que la relation

$$x = x'_1 + \ldots + x'_n, \qquad x'_1 \in X_1, \ldots, x'_n \in X_n$$

implique $x'_1 = x_1, \ldots, x'_n = x_n$.

Il suffit de prouver l'unicité de la décomposition (1) pour seul le vecteur nul: si la décomposition

$$0 = h_1 + \ldots + h_n, \qquad h_1 \in X_1, \ldots, h_n \in X_n$$

implique $h_1 = \ldots = h_n = 0$, alors la décomposition (1) est unique pour tout $x \in X$.

Soit X un espace vectoriel somme directe de sous-espaces X_1 et X_2. Alors les composantes x_1, x_2 de n'importe quel vecteur $x \in X$ sont les fonctions de x que l'on peut noter

$$x_1 = P_1(x), \quad x_2 = P_2(x).$$

On vérifie aisément que les fonctions $P_1(x)$ et $P_2(x)$ sont linéaires; elles s'appellent *projecteurs* (ou *opérateurs de projection*) sur les sous-espaces X_1 et X_2 respectivement. *Lorsque X est un espace vectoriel normé complet, et que les sous-espaces X_1 et X_2 sont fermés, les opérateurs P_1 et P_2 sont continus* (*U12.72d*). Naturellement, ce résultat reste valable quel que soit le nombre (fini) de termes dans la somme.

e. Inversement, d'après n'importe quels espaces vectoriels $X_1, \ldots, X_n$ on peut construire un espace vectoriel X qui sera leur somme directe. Pour le faire, on considère le produit cartésien (*1.12e*) des espaces $X_1, \ldots, X_n$ (l'ensemble de toutes les n-tuples $x = \{x_1, \ldots, x_n\}$, $x_1 \in X_1, \ldots, x_n \in X_n$) et l'on y introduit les opérations linéaires d'après les règles:

$$\{x_1, \ldots, x_n\} + \{y_1, \ldots, y_n\} = \{x_1 + y_1, \ldots, x_n + y_n\},$$
$$\alpha \{x_1, \ldots, x_n\} = \{\alpha x_1, \ldots, \alpha x_n\}.$$

L'espace X_k s'identifie de façon naturelle avec le sous-espace $X'_k \subset X$ composé des éléments de la forme

$$\{0, \ldots, 0, x_k, 0, \ldots, 0\}, \quad x_k \in X_k. \tag{2}$$

Si $X_1, \ldots, X_n$ sont des espaces normés, on peut doter X_n, lui aussi, d'une norme en posant, par exemple:

$$\| x \| = \max (\| x_1 \|, \ldots, \| x_n \|). \tag{3}$$

Ceci fait, lorsque $X_1, \ldots, X_n$ sont des espaces complets, X l'est également, les sous-espaces X'_k (2) s'avérant des sous-espaces fermés dans X.

f. Soit X une somme directe de ses sous-espaces (fermés), $X = X_1 + \ldots + X_n$, de sorte que, pour tout $x \in X$, on ait $x = P_1 x + \ldots + P_n x$ où $P_1, \ldots, P_n$ sont les projecteurs correspondants; soit, d'une façon analogue, Y une somme directe de ses sous-espaces (fermés), $Y = Y_1 + \ldots + Y_m$, de sorte que, pour

tout $y \in Y$, on ait $y = Q_1 y + \ldots + Q_m y$ où $Q_1, \ldots, Q_m$ sont les projecteurs correspondants. Posons

$$A_{ij} = Q_i A P_j \quad (i = 1, \ldots, m;\ j = 1, \ldots, n). \tag{4}$$

Les opérateurs A_{ij} agissent de X dans Y_i, mais il est naturel de considérer l'opérateur A_{ij} sur X_j seulement. Ces opérateurs composent la $(n \times m)$-matrice opératorielle

$$\| A_{ij} \| \equiv \begin{Vmatrix} A_{11} \ldots A_{1n} \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ A_{m1} \ldots A_{mn} \end{Vmatrix}. \tag{5}$$

Pour tout $x = \sum_{j=1}^{n} P_j x \equiv \sum_{j=1}^{n} x_j \equiv \sum_{j=1}^{n} P_j x_j$, nous avons

$$y_i \equiv Q_i y = Q_i A x = \sum_{j=1}^{n} Q_i A P_j x = \sum_{j=1}^{n} Q_i A P_j x_j = \sum_{j=1}^{n} A_{ij} x_j, \tag{6}$$

ce qui rappelle la notation matricielle ordinaire d'un opérateur linéaire de R_n dans R_m. La formule (6) montre que l'opérateur A est complètement déterminé par la matrice $\| A_{ij} \|$. De plus, à tout ensemble donné d'opérateurs linéaires continus $A_{ij} : X_j \to Y_i$ $(i = 1, \ldots, m;\ j = 1, \ldots, n)$ il correspond un opérateur linéaire continu $A : X \to Y$ défini par la formule

$$Ax = \sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij} x_j$$

ou, ce qui revient au même,

$$y_i \equiv (Ax)_i = \sum_{j=1}^{n} A_{ij} x_j ;$$

il est aisé de voir que la $(n \times m)$-matrice opératorielle, construite d'après la règle ci-dessus, coïncide avec la matrice $\| A_{ij} \|$. Ainsi, les opérateurs $A : X \to Y$ sont en correspondance biunivoque naturelle avec les $(n \times m)$-matrices opératorielles $\| A_{ij} \|$.

Notons ensuite que, étant donnés deux opérateurs linéaires $A : X \to Y$ et $B : X \to Y$, avec les matrices $\| A_{ij} \|$ et $\| B_{ij} \|$ respectivement, à toute combinaison linéaire $\alpha A + \beta B$ il correspond évidemment la matrice $\| \alpha A_{ij} + \beta B_{ij} \|$. Par conséquent, la correspondance biunivoque signalée est un isomorphisme de l'espace vectoriel $L(X, Y)$ et de l'espace vectoriel de toutes les $(n \times m)$-matrices opératorielles $\| A_{ij} \|$ (avec les opérations linéaires ordinaires).

g. En particulier, dans l'hypothèse de f, l'espace $L(X, Y)$ se trouve isomorphe à la somme directe de mn espaces $L(X_j, Y_i)$.

En posant $n = 1$ nous voyons que $L(X, Y)$ est isomorphe à la somme directe de m espaces $L(X, Y_i)$, et si $m = 1$, alors $L(X, Y)$ est isomorphe à la somme directe de n espaces $L(X_j, Y)$.

h. Supposons qu'en plus des espaces X et Y considérés dans f il y ait un troisième espace Z somme directe des sous-espaces (fermés) $Z_1 + \ldots + Z_p$. Soit $B: Z \to X$ un opérateur linéaire continu; d'après les décompositions $X = X_1 + \ldots + X_n$ et $Z = Z_1 + \ldots + \ldots + Z_p$, on peut lui faire correspondre la matrice $\| B_{ik} \|$, $i = 1, \ldots, n$; $k = 1, \ldots, p$. L'opérateur AB agit de Z dans Y; on a $ABz = A \sum_{j,k} B_{jk} z_k = \sum_{i,j,k} A_{ij} B_{jk} z_k$, de sorte que $(ABz)_i = \sum_{j,k} A_{ij} B_{jk} z_k$. Par conséquent, $(AB)_{ik} = \sum_{j=1}^{n} A_{ij} B_{jk}$; cela correspond à la règle usuelle de la multiplication des matrices numériques.

i. La représentation des opérateurs linéaires par les matrices opératorielles s'applique parfois en démontrant l'inversibilité de tel ou tel opérateur. Soient, par exemple, $A: X_1 \to Y_1$ et $B: X_2 \to Y_2$ des opérateurs inversibles, $C: X_2 \to Y_1$ un opérateur quelconque; montrons que *l'opérateur* $S: X_1 + X_2 \to Y_1 + Y_2$ *défini par la matrice*

$$\begin{Vmatrix} A & C \\ 0 & B \end{Vmatrix} \tag{7}$$

est un opérateur inversible. Par analogie avec l'inversion d'une matrice bidimensionnelle triangulaire, cherchons l'opérateur inverse $S^{-1}: Y_1 + Y_2 \to X_1 + X_2$ comme matrice opératorielle triangulaire

$$\begin{Vmatrix} P & R \\ 0 & Q \end{Vmatrix}.$$

En appliquant à l'égalité exigée

$$\begin{Vmatrix} A & C \\ 0 & B \end{Vmatrix} \cdot \begin{Vmatrix} P & R \\ 0 & Q \end{Vmatrix} = E \equiv \begin{Vmatrix} E_1 & 0 \\ 0 & E_2 \end{Vmatrix}$$

la règle de multiplication des matrices opératorielles, nous arrivons aux relations

$$AP = E_1, \quad BQ = E_2, \quad AR + CQ = 0.$$

Nous tirons de la première $P = A^{-1}$, de la seconde, $Q = B^{-1}$. En multipliant la troisième par A^{-1} trouvons $R = -A^{-1}CB^{-1}$. Ainsi, l'opérateur donné par la matrice

$$\begin{Vmatrix} A^{-1} & -A^{-1}CB^{-1} \\ 0 & B^{-1} \end{Vmatrix}$$

est l'inverse à droite de l'opérateur S donné par la matrice (7). Il est facile de prouver que le même opérateur est inverse à gauche de S.

1.15. Opérations sur les fonctions continues.

a. Si $f(x)$ et $g(x)$ sont deux fonctions définies sur un même ensemble X et à valeurs dans un espace vectoriel Y, on peut définir sur X encore une fonction $y(x) = f(x) + g(x)$ $(X \to Y)$ appelée *somme* des fonctions $f(x)$ et $g(x)$. Lorsque X est un espace métrique et Y un espace vectoriel normé et que les fonctions $f(x)$ et $g(x)$ sont continues, il en est de même de $y(x)$ (*U12.32b*).

b. Si une fonction $y(x)$ est définie sur un ensemble X et prend ses valeurs dans un espace vectoriel Y et si C est un opérateur linéaire de Y dans un espace vectoriel Z, alors on peut définir sur X la fonction $z(x) = Cy(x)$. Supposons que X soit un espace métrique, Y et Z des espaces normés, et que $y(x)$ et C soient continus. Alors la fonction $z(x)$ est continue elle aussi (*1.13a*). Par exemple, les composantes $x_1 = P_1x, \ldots, x_n = P_nx$ d'un vecteur x dans une décomposition directe $X_1 + \ldots + X_n$ d'un espace normé complet X sont des fonctions continues de x (*1.14d*).

c. Soit Y un espace vectoriel somme directe de ses sous-espaces $Y_1, \ldots, Y_n$. Alors, pour toute fonction $y(x): X \to Y$, on définit, en se servant des projections respectives $P_1, \ldots, P_n$, encore n fonctions

$$y_1(x) = P_1y(x)\ (X \to Y_1), \ \ldots, \ y_n(x) = P_ny(x)\ (X \to Y_n).$$

Si, de plus, Y est un espace normé complet et les sous-espaces $Y_1, \ldots, Y_n$ sont fermés, la continuité de la fonction $y(x)$ implique, en vertu de b et de *1.14d*, celle des fonctions $y_k(x)$ $(k = 1, \ldots, n)$. La réciproque est immédiate: étant données n fonctions continues $y_1(x): X \to Y_1, \ldots, y_n(x): X \to Y_n$, la fonction $y(x) \equiv \{y_1(x), \ldots, y_n(x)\} = y_1(x) + \ldots + y_n(x)$ $(X \to Y)$ est aussi continue d'après a.

d. On peut considérer des produits de divers genres de deux fonctions $\lambda(x)$ et $f(x)$ définies sur un même ensemble X. Un tel produit est, par exemple, défini si $f(x)$ applique X dans un espace vectoriel Y, $\lambda(x)$ étant une fonction numérique. On a pour résultat de la multiplication une fonction $(\lambda f)(x) \equiv \lambda(x) f(x)$ $(X \to Y)$. Un autre exemple correspond au cas où $\lambda(x)$ et $f(x)$ sont à valeurs dans une algèbre Y; les valeurs du produit appartiennent alors à cette même algèbre. Un produit est également défini pour $f(x)$: $X \to Y$ et $\lambda(x): X \to L(Y, Z)$; la fonction $\lambda(x) f(x)$ applique X dans Z (ici Y et Z sont des espaces vectoriels). La linéarité du produit par rapport à chaque facteur est une propriété commune à tous ces exemples.

Donnons la définition générale suivante. Supposons que, sur la somme directe de deux espaces normés Λ et F, soit défini un opérateur continu $B(\lambda, f)$ $(\lambda \in \Lambda,\ f \in F)$ linéaire par rapport à chacune des variables λ et f, à valeurs dans un espace normé B; nous l'appellerons *produit généralisé* des éléments λ et f et noterons $b = \langle \lambda, f \rangle$. Soient, de plus, deux fonctions $\lambda(x): X \to \Lambda$ et $f(x): X \to F$; alors il est défini la fonction $\langle \lambda(x), f(x) \rangle$ $(X \to B)$ appelée *produit généralisé* des fonctions $\lambda(x)$ et $f(x)$. Tous les trois exemples entrent bien dans ce schéma si l'on pose respectivement:

$\langle \lambda(x), f(x) \rangle = \lambda(x) f(x)$ au sens ordinaire de la multiplication de $f(x)$ par le nombre $\lambda(x)$;

$\langle \lambda(x), f(x) \rangle = \lambda(x) f(x)$ au sens de la multiplication des éléments de l'algèbre;

$\langle \lambda(x), f(x) \rangle = \lambda(x) f(x)$ au sens de l'image du vecteur $f(x)$ par l'opérateur $\lambda(x)$.

Nous affirmons que *le produit généralisé* $b(x) = \langle \lambda(x), f(x) \rangle$ *est une fonction continue de* x *dès que* $\lambda(x)$ *et* $f(x)$ *le sont.* En effet, $b(x)$ est la composée de deux fonctions:

$$x \to \{\lambda(x), f(x)\} \in \Lambda + F \quad \text{et} \quad \{\lambda, f\} \to \langle \lambda, f \rangle \in B.$$

La première est continue par hypothèse de continuité de $\lambda(x)$ et $f(x)$ et en vertu de *c*; la deuxième par hypothèse de continuité de la fonction $\langle \lambda, f \rangle: \Lambda + F \to B$. Le produit $b(x) = \langle \lambda(x), f(x) \rangle$ est continu d'après le théorème sur la continuité d'une composée.

En particulier, chacun des produits vus plus haut est continu à condition de continuité des facteurs.

e. Les résultats *a* et *d* sont valables, en particulier, pour les fonctions numériques $X \to R_1$ continues sur un espace métrique X.

Par exemple, les coordonnées d'un point $x = \{x_1, \ldots, x_n\}$ de l'espace $X = R_n$ étant des fonctions continues de x, et la fonction $1/t$ $(R_1 \to R_1)$ étant continue pour $t \neq 0$, on voit que, d'après *a*, *d* et *1.13a*, les polynômes et les fonctions rationnelles des coordonnées sont continus (les dernières en dehors des zéros du dénominateur). En vertu de *1.13b*, la limite d'une suite de polynômes qui converge uniformément sur un ensemble $G \subset R_n$ est une fonction $(G \to R_1)$ continue sur G. Pour certains ensembles $G \subset R_n$, la réciproque est également valable: toute fonction continue $f(x)$ $(G \to R_1)$ peut être représentée comme limite d'une suite uniformément convergente de polynômes; ceci est vrai, par exemple, pour les compacts (*U12.54*).

1.16. Le plus simple type de discontinuité d'une fonction numérique d'une variable réelle $(R_1 \to R_1)$ est le *point de discontinuité de la première espèce* où les limites à droite $f(x + 0)$ et à gauche $f(x - 0)$ d'une fonction $f(x)$ existent mais ne sont pas égales. Pour une fonction numérique de n variables réelles, on pourrait appeler point de discontinuité de la première espèce un point a où la fonction $f(x)$ admet une limite suivant chaque demi-droite aboutissant

au point a, sans que ces limites soient toutes égales. En exemple on peut signaler la fonction de deux variables réelles $f(x_1, x_2) = \frac{x_1^2}{x_1^2 + x_2^2}$ $(R_2 \to R_1)$ qui, sur toute demi-droite $x_1 = t \cos \alpha$, $x_2 = t \sin \alpha$, $0 < t < \infty$, prend une valeur constante $\cos^2 \alpha$. Pourtant, nous heurtons ici la complication suivante.

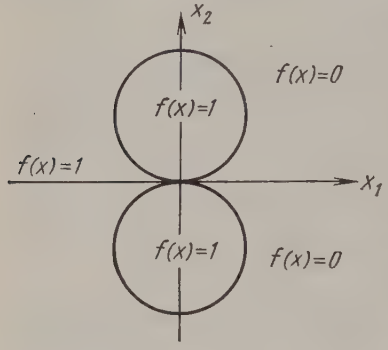


Fig. 1.1-10

Toute fonction d'une variable dont les limites à gauche et à droite à un point $x = a$ se confondent et coïncident avec sa valeur au point a est continue en a (*U5.24*). Or, une fonction de n variables dont les limites suivant tous les rayons menant à un point a existent et se confondent n'est pas forcément continue en a. La fonction $f(x_1, x_2)$ $(R_2 \to R_1)$ qui vaut 1 sur l'axe des x_1 ainsi que dans deux cercles disposés dans les demi-plans supérieur et inférieur et tangents à cet axe au point (0, 0), et qui est nulle partout ailleurs dans le plan (fig. 1.1-10), en sert d'exemple. Sur toute demi-droite aboutissant au point (0, 0), cette fonction a 1 pour limite, mais elle n'est pas continue au point (0, 0). Pour cette raison, la notion de « point de discontinuité de la première espèce » ne s'utilise pas en cas de fonctions de plusieurs variables.

Le même exemple montre que, pour les fonctions de plusieurs variables, les points de discontinuité peuvent bien ne pas être isolés ; même pour des fonctions très simples, ils peuvent remplir des lignes entières ou, s'il s'agit de plus de deux variables, des surfaces entières ; ainsi, la fonction $f(x)$: $R_n \to R_1$ qui vaut 1 dans la boule $\{x \in R_n : \sum_{j=1}^{n} x_j^2 \leqslant 1\}$ et qui est nulle partout ailleurs, est discontinue sur toute la sphère $\{x \in R_n : \sum_{j=1}^{n} x_j^2 = 1\}$.

§ 1. 2. Fonctions dérivables

L'idée principale du calcul différentiel est de remplacer une fonction donnée au voisinage d'un point par une fonction du premier degré, de sorte que l'erreur qui s'en produit soit un infiniment petit d'ordre supérieur par rapport à l'accroissement de la variable indépendante. Les fonctions numériques d'une variable x, pour lesquelles une telle approximation est possible, forment la classe des fonctions dérivables de x. Or, la possibilité d'approximation locale par une fonction du premier degré n'exige point que l'argument soit unidimensionnel. Nous commençons par donner les définitions nécessaires pour une fonction numérique de plusieurs variables, puis formulerons la définition générale qui sera valable au cas où le domaine de définition et celui des valeurs de la fonction sont des espaces vectoriels normés.

1.21. Rappelons avant tout la définition d'une fonction numérique dérivable d'une variable réelle (*U7.11*).

Une fonction numérique $f(x)$ d'une variable réelle x, $a \leqslant x \leqslant b$, est dite *dérivable* en un point $x = c$ s'il existe la limite

$$\lim_{h \to 0} \frac{f(c+h) - f(c)}{h} = f'(c). \tag{1}$$

Dans ce cas, on peut dégager la partie principale de l'accroissement de la fonction $f(x)$ correspondant au passage de la valeur $x = c$ de l'argument à sa valeur $x = c + h$, cette partie étant la fonction linéaire de h:

$$f(c+h) - f(c) = f'(c)\, h + o(h), \quad \lim_{h \to 0} \frac{o(h)}{h} = 0.$$

Inversement, si l'accroissement d'une fonction $f(x)$ correspondant au passage de $x = c$ à $x = c + h$ admet une partie principale linéaire en h, c'est-à-dire que, pour une constante D, on a la relation

$$f(c+h) - f(c) = Dh + o(h), \quad \lim_{h \to 0} \frac{o(h)}{h} = 0, \tag{2}$$

alors la limite $\lim\limits_{h \to 0} \dfrac{f(c+h) - f(c)}{h}$ existe et vaut D.

Ainsi, dans le cas d'une fonction numérique d'une variable, les relations (1) et (2) représentent deux définitions équivalentes de la dérivabilité au point $x = c$.

1.22. a. A présent, envisageons le cas d'une fonction numérique $f(x)$ de n variables réelles $x = (x_1, \ldots, x_n)$. C'est la deuxième de deux définitions ci-dessus qui s'étend d'une façon naturelle à ce cas.

A savoir, nous disons qu'une fonction $f(x) \equiv f(x_1, \ldots, x_n)$, $f: R_n \to R_1$ est *dérivable à un point* $x = c = (c_1, \ldots, c_n)$, si l'accroissement de $f(x)$ correspondant au passage du point $x = c$ à un point $x = c + h$, $h = (h_1, \ldots, h_n)$ *admet une partie principale linéaire en* h. Le dernier veut dire que, pour certaines constantes $D_1, \ldots, D_n$, on a la relation

$$f(c+h) - f(c) = \sum_{i=1}^{n} D_i h_i + o(h), \quad \lim_{h \to 0} \frac{o(h)}{|h|} = 0. \tag{1}$$

C'est la valeur $\sum\limits_{i=1}^{n} D_i h_i$ qui est la partie linéaire principale de l'accroissement de la fonction $f(x)$ lorsque x varie de c à $c + h$. La quantité $o(h)$ est un infiniment petit d'ordre supérieur par rapport à h. Le sens rigoureux de la notation $\lim\limits_{h \to 0} \dfrac{o(h)}{|h|} = 0$ est : pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que $\dfrac{|o(h)|}{|h|} < \varepsilon$ dès que $|h| < \delta$.

Formellement, la définition (1) dépend du choix de base dans l'espace R_n. Or, si l'on se rappelle (*L5.4*) que la forme linéaire

$\sum_{i=1}^{n} D_i h_i$ reste forme linéaire en passant à une nouvelle base (les coefficients changés), il devient clair que la dérivabilité de la fonction $f(x)$ pour $x = c$ est sa propriété intrinsèque qui ne dépend pas du choix de base dans R_n.

b. Lorsqu'une base, donc aussi un système de coordonnées, sont choisis, on peut dire que, *pour une fonction $f(x)$ dérivable, les coefficients D_i dans la formule* (1) *se déterminent de façon unique.* Afin de le voir, choisissons un nombre entier m entre 1 et n et posons dans (1)

$$h = (0, \ldots, 0, h_m, 0, \ldots, 0).$$

Nous aurons alors $|h| = |h_m|$ et

$$f(c + h) - f(c) = f(c_1, \ldots, c_{m-1}, c_m + h_m, c_{m+1}, \ldots, c_n) - \\ - f(c_1, \ldots, c_{m-1}, c_m, c_{m+1}, \ldots, c_n) = D_m h_m + o(h_m);$$

ceci veut dire que la fonction $f(c_1, \ldots, x_m, \ldots, c_n)$ d'une seule variable x_m est dérivable par rapport à x_m au point $x_m = c_m$ et que le nombre D_m n'est rien d'autre que sa dérivée par rapport à cette variable :

$$D_m = \lim_{h_m \to 0} \frac{f(c_1, \ldots, c_m + h_m, \ldots, c_n) - f(c_1, \ldots, c_m, \ldots, c_n)}{h_m}. \qquad (2)$$

La formule obtenue démontre l'unicité des coefficients D_i dans (1).

Si la limite dans le deuxième membre de (2) existe, elle s'appelle *dérivée partielle de la fonction $f(x)$ par rapport à la variable x_m au point $x = c$.* Ainsi, une fonction $f(x)$ dérivable à un point $x = c$ (au sens de (1)) y possède une dérivée partielle par rapport à chacune des variables $x_1, \ldots, x_n$.

La dérivée partielle d'une fonction $f(x)$ par rapport à une variable x_m à un point $x = c$ est notée $\frac{\partial f(c)}{\partial x_m}$ ou $f'_{x_m}(c)$. Remarquons que les dérivées partielles d'une fonction $f(x)$ à un point c par rapport à toutes les variables peuvent bien exister sans que la fonction $f(x)$ soit dérivable au point c (cf. exercice 3).

c. Les nombres $D_1, \ldots, D_n$ déterminent le vecteur $D = \{D_1, \ldots, D_n\}$ qui s'appelle *gradient* de la fonction $f(x)$ au point $x = c$ et se note $\operatorname{grad} f(c)$. L'expression $\sum_{1}^{n} D_i h_i$ s'appelle *différentielle de la fonction $f(x)$* au point c pour le déplacement $h = \{h_1, \ldots, h_n\}$. Elle est désignée par $df(c)$. La tradition veut que les quantités h et h_i soient désignées par dx et dx_i $(i = 1, \ldots, n)$ respectivement. Si l'on munit l'espace R_n d'un produit scalaire de deux vecteurs $x = \{x_1, \ldots, x_n\}$ et $y = \{y_1, \ldots, y_n\}$ d'après la

formule

$$(x,\ y) = \sum_1^n x_i y_i,$$

alors on peut mettre la différentielle de la fonction $f(x)$ au point c sous n'importe laquelle des formes suivantes:

$$df(c) = \sum_{i=1}^n D_i h_i = (D,\ h) = (\operatorname{grad} f(c),\ h) = \sum_{i=1}^n \frac{\partial f(c)}{\partial x_i}\, dx_i. \qquad (3)$$

Les aspects respectifs de la formule (1) sont:

$$f(c+h) - f(c) = df(c) + o(h) = (\operatorname{grad} f(c),\ h) + o(h) =$$

$$= \sum_{i=1}^n \frac{\partial f(c)}{\partial x_i}\, dx_i + o(dx). \qquad (4)$$

d. En cas de fonction de deux variables $y = f(x_1,\ x_2)$ $(R_2 \to R_1)$, on utilise souvent les désignations plus traditionnelles: x, y pour les variables indépendantes, z pour la fonction, soit

$$z = f(x,\ y).$$

Les formules (3) et (4) ont alors les formes respectives:

$$dz = \frac{\partial z}{\partial x}\, dx + \frac{\partial z}{\partial y}\, dy\,, \qquad (5)$$

$$f(x+dx,\ y+dy) - f(x,\ y) = \frac{\partial z}{\partial x}\, dx + \frac{\partial z}{\partial y}\, dy + o(|dx| + |dy|)\,. \qquad (6)$$

e. Comme la dérivée partielle d'une fonction de plusieurs variables est sa dérivée ordinaire par rapport à l'une des variables indépendantes (les valeurs des autres étant fixées), le calcul de dérivées partielles se ramène à celui de dérivées ordinaires. Ainsi, pour la fonction $z = x^2/y^2$ $(R_2 \to R_1)$, on a:

$$\frac{\partial z}{\partial x} = \frac{2x}{y^2} \quad (y \text{ est fixe}),$$

$$\frac{\partial z}{\partial y} = -\frac{2x^2}{y^3} \quad (x \text{ est fixe}).$$

La différentielle de cette fonction a la forme:

$$dz = \frac{\partial z}{\partial x}\, dx + \frac{\partial z}{\partial y}\, dy = \frac{2x}{y^2}\, dx - \frac{2x^2}{y^3}\, dy.$$

Son gradient au point $(x,\ y)$ est:

$$\operatorname{grad} z\ (x,\ y) = \left\{ \frac{2x}{y^2},\ -\frac{2x^2}{y^3} \right\}.$$

1.23. a. Formulons maintenant la définition *générale* d'une fonction dérivable. Soit $y = f(x)$ une fonction définie dans un domaine $G \subset X$ d'un espace normé X, à valeurs dans un espace normé Y. Cette fonction est dite dérivable à un point $x = c \in G$,

lorsque l'accroissement de $f(x)$ correspondant au déplacement du point c à un point $c + h \in G$ admet une partie principale linéaire en h, c'est-à-dire lorsqu'on a la relation

$$f(c+h) - f(c) = Dh + o(h), \tag{1}$$

D étant un opérateur linéaire continu de l'espace X dans Y, $o(h)$ un vecteur dans l'espace Y qui satisfait à la condition

$$\lim_{h \to 0} \frac{o(h)}{|h|} = 0.$$

L'expression Dh est donc, dans le présent cas, la *partie linéaire principale de l'accroissement de la fonction* $f(x)$ lorsque x varie de c à $c + h$.

L'égalité (1) admet l'interprétation géométrique suivante. La fonction $f(x)$ réalise une application du domaine $G \subset X$ dans l'espace Y; si l'on met l'origine des coordonnées de l'espace X au point c, et l'origine des coordonnées de l'espace Y au point $f(c)$, c'est-à-dire que l'on choisit le vecteur $h = x - c$ pour variable indépendante et le vecteur $k = y - f(c)$ pour sa fonction, alors l'application $h \to k$ que l'on obtient de la sorte admet l'approximation par l'application linéaire $k = Dh$ (à un infiniment petit $o(h)$ d'ordre supérieur par rapport à h près). On peut donc dire que l'application $y - f(c) = f(x - c)$ elle-même admet une *partie linéaire principale* au voisinage du point $x - c = 0$.

b. Montrons que l'opérateur D qui figure dans la formule (1) est défini d'une façon univoque. Supposons que, à côté de (1), il y ait une autre représentation analogue de la différence $f(c+h) - f(c)$, notamment:

$$f(c+h) - f(c) = D_1 h + o_1(h), \quad \lim_{h \to 0} \frac{o_1(h)}{|h|} = 0. \tag{2}$$

En retranchant (2) de (1) et désignant $D_2 = D - D_1$, on trouve

$$D_2 h = o_2(h), \quad \lim_{h \to 0} \frac{o_2(h)}{|h|} = 0.$$

Pour $\varepsilon > 0$ donné, choisissons $\delta > 0$ tel que $|o_2(h)| \leqslant \varepsilon |h|$ du moment que $|h| \leqslant \delta$. Alors, vu *U12.71b*, on a

$$\|D_2\| = \sup_{|h| \leqslant \delta} \frac{|D_2 h|}{|h|} = \sup_{|h| \leqslant \delta} \frac{|o_2(h)|}{|h|} \leqslant \varepsilon.$$

Puisque $\varepsilon > 0$ est arbitraire, $\|D_2\| = 0$ d'où $D_2 = D - D_1 = 0$, donc $D = D_1$, ce qu'il fallait.

c. L'opérateur D dans la formule (1) qui est, on vient de le voir, unique s'appelle *dérivée* de la fonction $f(x)$ pour $x = c$ et se note $f'(c)$. La quantité Dh, i.e. le vecteur dans l'espace Y obtenu

en appliquant l'opérateur dérivée au vecteur déplacement h, s'appelle *différentielle de la fonction* $f(x)$ *au point* c pour le déplacement h. Par analogie avec les fonctions numériques d'une variable réelle, on note

$$df(c) = Dh = f'(c)\,h = f'(c)\,dx,$$

où $dx = h$ désigne n'importe quel vecteur de l'espace X.

La foncion $f(x)$ est dite *dérivable dans le domaine* G si elle l'est en tout point de ce domaine. Sa dérivée $f'(x)$ est un opérateur linéaire de X dans Y, fonction du point $x \in G$. Sa différentielle $df(x) = f'(x)\,h$ est une *fonction de deux variables*: du vecteur $h \in X$ et du point $x \in G$.

Le passage de la fonction $f(x)$ à sa dérivée $f'(x)$ est la *dérivation de la fonction* $f(x)$, le passage de $f(x)$ à sa différentielle $f'(x)\,dx$, la *différentiation* de $f(x)$.

En cas de fonction d'une variable réelle, la définition générale (1) coïncide, évidemment, avec la définition ordinaire de la dérivée et de la différentielle (*U7.11* et *U12.61*). Pour une fonction numérique de n variables réelles, elle coïncide avec la définition donnée plus haut (*1.22*).

1.24. Soit $f(x)$, $f: G \to R_1$ une fonction numérique dérivable définie dans un domaine G d'un espace normé X. Dans ce cas, l'opérateur $D = f'(c)$ de *1.23* (1) est une fonctionnelle linéaire continue $(X \to R_1)$ (qui dépend, en général, du point c); la fonction $o(h)$ prend, elle aussi, des valeurs numériques. Si X est de dimension n, nous revenons à la définition *1.22* car l'expression $\sum_1^n D_i h_i$ représente la forme générale de fonctionnelle dans R_n. La fonctionnelle linéaire $D = f'(c)$ est également appelée *gradient de la fonction* $f(x)$ *au point* $x = c$. Ainsi donc, la définition générale de la dérivée d'une fonction numérique se confond avec la définition de son gradient.

1.25. a. Etudions plus en détail les fonctions vectorielles dérivables dans le cas de dimension finie.

Supposons qu'une fonction $f(x): G \subset R_n \to R_m$ soit définie dans un domaine G d'un espace n-dimensionnel $X = R_n$ et prenne ses valeurs dans un espace m-dimensionnel $Y = R_m$. En choisissant une base dans chacun de deux espaces et notant $x = \{x_1, \ldots, x_n\} \in R_n$, $y = \{y_1, \ldots, y_m\} \in R_m$, on est en mesure d'exprimer la fonction vectorielle $y = f(x)$ à l'aide de m fonctions numériques:

$$\left.\begin{aligned} y_1 = f_1(x) &\equiv f_1(x_1, \ldots, x_n), \\ &\cdots\cdots\cdots\cdots \\ y_m = f_m(x) &\equiv f_m(x_1, \ldots, x_n). \end{aligned}\right\} \qquad (1)$$

Supposons que $f(x)$ soit dérivable pour $x = c$. L'égalité

$$f(c+h) - f(c) = f'(c)\,h + o(h) \tag{2}$$

définit l'opérateur linéaire $f'(c)\colon R_n \to R_m$. Comme à tout opérateur linéaire de R_n dans R_m, on peut lui faire correspondre une $n \times m$-matrice. Pour le faire, on doit exprimer l'égalité (2) en coordonnées par rapport aux bases fixées dans R_n et R_m ; on obtient :

$$f_i(c+h) - f_i(c) = \sum_{j=1}^{n} f'_{ij}(c)\,h_j + o_i(h) \quad (i = 1, \ldots, m). \tag{3}$$

Ce sont les valeurs $f'_{ij}(c)$ $(j = 1, \ldots, n;\ i = 1, \ldots, m)$ qui forment ladite $n \times m$-matrice de l'opérateur $f'(c)$ par rapport aux bases mentionnées. Les formules (3) montrent que *les composantes* $f_i(x)$ *d'une fonction vectorielle* $f(x)$ *dérivable* (pour $x = c$) *sont elles-mêmes des fonctions numériques dérivables* (pour $x = c$). La réciproque est immédiate : *si les fonctions numériques* $f_i(x)$, $i = 1, \ldots, m$ *sont dérivables pour* $x = c$, *il en est de même de la fonction vectorielle* $f(x)$. Nous voyons, de plus, comment sont les éléments de la matrice de l'opérateur linéaire $f'(c)$: les coefficients de la partie linéaire principale de l'accroissement de n'importe quelle fonction numérique étant ses dérivées partielles par rapport aux coordonnées respectives, nous avons :

$$f'_{ij}(c) = \frac{\partial f_i(c)}{\partial x_j} \quad (i = 1, \ldots, m,\ j = 1, \ldots, n). \tag{4}$$

Définitivement, la matrice de l'opérateur linéaire $f'(c)$ $(R_n \to R_m)$ est composée des dérivées partielles et a la forme

$$f'(c) \cong \left\| \begin{array}{cccc} \frac{\partial f_1(c)}{\partial x_1} & \frac{\partial f_1(c)}{\partial x_2} & \cdots & \frac{\partial f_1(c)}{\partial x_n} \\ \frac{\partial f_2(c)}{\partial x_1} & \frac{\partial f_2(c)}{\partial x_2} & \cdots & \frac{\partial f_2(c)}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f_m(c)}{\partial x_1} & \frac{\partial f_m(c)}{\partial x_2} & \cdots & \frac{\partial f_m(c)}{\partial x_n} \end{array} \right\|.$$

Ici le symbole $\cong$ signifie : à l'opérateur $f'(c)$, il correspond la matrice dans le second membre, du moment que les bases dans R_n et R_m sont choisies. Cette matrice s'appelle *matrice de Jacobi*. On la désigne également par $\left\| \frac{\partial f_i(c)}{\partial x_j} \right\|$ ou $\left\| \frac{\partial (f_1, \ldots, f_m)}{\partial (x_1, \ldots, x_n)} \right\|$. Dans le cas où $n = m = 1$ elle se compose d'un seul élément identique à la dérivée ordinaire d'une fonction réelle par rapport à l'unique variable réelle. Pour une fonction numérique de n variables réelles, on a $m = 1$ et la matrice de Jacobi n'a qu'une seule ligne. Pour m fonctions d'une

variable réelle (courbe dans l'espace m-dimensionnel, *U16.12*), on a $n = 1$ et la matrice de Jacobi se réduit à une seule colonne.

Dans le cas $m = n$, la matrice de Jacobi est carrée

$$f'(c) \cong \left\| \begin{matrix} \frac{\partial f_1(c)}{\partial x_1} & \dots & \frac{\partial f_1(c)}{\partial x_n} \\ \dots & \dots & \dots \\ \frac{\partial f_n(c)}{\partial x_1} & \dots & \frac{\partial f_n(c)}{\partial x_n} \end{matrix} \right\|.$$

Son déterminant, on le verra plus bas, représente une importante caractéristique de l'application $y = f(x)$ pour $x = c$; on l'appelle *jacobien de l'application* $f(x)$ pour $x = c$ et on le note

$$\det \left\| \frac{\partial (f_1, \dots, f_n)}{\partial (x_1, \dots, x_n)} \right\|.$$

b. La situation est particulièrement simple dans le cas où $y = f(x)$ est une fonction linéaire, donc a la forme développée suivante:

$$\begin{aligned} y_1 &= D_{11}x_1 + \dots + D_{1n}x_n, \\ &\dots\dots\dots\dots \\ y_m &= D_{m1}x_1 + \dots + D_{mn}x_n, \end{aligned}$$

D_{ij} étant des nombres constants. Ici la matrice de Jacobi est composée des nombres D_{ij}:

$$f'(c) \cong \left\| \begin{matrix} D_{11} & \dots & D_{1n} \\ \dots & \dots & \dots \\ D_{m1} & \dots & D_{mn} \end{matrix} \right\|$$

et ne dépend pas du point c.

1.26. a. Rappelons que (*U7.11*), dans le cas d'une variable réelle, l'existence de la dérivée d'une fonction numérique $y = f(x)$ à un point $x = c$ signifie, du point de vue géométrique, l'existence de la tangente au graphique de $f(x)$ au point $(c, f(c))$. Là, on sous-entend par tangente ou bien la position limite d'une sécante, conformément à la définition analytique *1.21* (1), ou bien la droite dont les points sont éloignés des points correspondants (pour les mêmes x) sur la courbe $y = f(x)$ d'une distance infiniment petite d'ordre supérieur par rapport à $h = x - c$; ceci est conforme à la définition analytique *1.21* (2).

Le coefficient angulaire de la tangente coïncide avec la dérivée $f'(c)$, et l'équation de la tangente est

$$y - p = f'(c)(x - c) \qquad (p = f(c)), \tag{1}$$

ou bien, en désignant $x - c = dx$, $y - p = dy$,

$$dy = f'(c)\, dx. \tag{2}$$

b. En cas de fonction numérique $y = f(x_1, \ldots, x_n)$, l'interprétation géométrique de la dérivabilité est liée à l'existence du *plan tangent à la surface* $y = f(x_1, \ldots, x_n)$. Pour arriver à une définition du plan tangent, nous allons généraliser la deuxième des définitions ci-mentionnées de la droite tangente. A savoir, nous appelons *plan tangent à la surface* $y = f(x)$ *à un point* $x = c$ un plan $y - p = \sum A_i (x_i - c_i)$, $p = f(c)$ dans l'espace $(n+1)$-dimensionnel $x_1, \ldots, x_n, y$, si l'écart entre les points $\{x_1, \ldots, x_n, f(x)\}$ et $\{x_1, \ldots, x_n, p + \sum A_i (x_i - c_i)\}$ (le premier sur la surface, le deuxième sur le plan) a un ordre de petitesse supérieur par rapport à celui de $|h| = \sqrt{\sum (x_i - c_i)^2}$. Il est clair que la condition analytique *1.22* (1) de la dérivabilité de la fonction $f(x)$ pour $x = c$ peut maintenant être interprétée comme condition d'existence du plan tangent à la surface $y = f(x)$ pour $x = c$; dans les désignations de *1.22* (1), ce plan tangent a pour équation

$$y - p = \sum_{i=1}^{n} D_i (x_i - c_i) = \sum_{i=1}^{n} \frac{\partial f(c)}{\partial x_i} (x_i - c_i), \tag{3}$$

soit, en différentielles ($dy = y - p$, $dx_i = x_i - c_i$, $i = 1, \ldots, n$):

$$dy = \sum_{i=1}^{n} \frac{\partial f(c)}{\partial x_i} dx_i. \tag{4}$$

Le vecteur $\{dx_1, \ldots, dx_n, -1\} \in R_{n+1}$ est orthogonal au plan (4) (au sens de produit scalaire ordinaire); nous le dirons normal à la surface $y = f(x)$ au point $\{c, p\}$, la droite qu'il détermine et qui passe par ce point sera appelée normale à la surface $y = f(x)$.

Ainsi, pour la surface correspondant à la fonction $z = x^2/y^2$ $(R_2 \to R_1)$ et considérée dans *1.22d*, l'équation du plan tangent est

$$dz = \frac{2x}{y^2} dx - \frac{2x^2}{y^3} dy \tag{5}$$

(ce qui coïncide formellement avec l'expression de la différentielle). Si l'on désigne les coordonnées courantes du point du plan tangent par X, Y, Z tout en gardant les symboles x, y, z pour les coordonnées du point de tangence, on peut mettre l'équation (5) sous la forme

$$Z - z = \frac{2x}{y^2}(X - x) - \frac{2x^2}{y^3}(Y - y). \tag{6}$$

Le vecteur normal au point $\{x, y, z\}$ a $\frac{2x}{y^2}$, $-\frac{2x^2}{y^3}$, -1 pour composantes; c'est pourquoi l'équation de la normale au point $\{x, y, z\}$ est

$$\frac{X - x}{2x/y^2} = \frac{Y - y}{-2x/y^3} = \frac{Z - z}{-1},$$

où, cette fois, X, Y, Z sont les coordonnées courantes du point de la normale.

c. On peut aussi bien formuler la définition du plan tangent dans le cas général d'une fonction $y = f(x)$: $G \subset X \to Y$ dérivable à un point $x = c$. Soit $f(c) = p$. Alors nous appelons « plan tangent » à la surface $y = f(x)$ au point c la variété linéaire (hyperplan) dans l'espace $X + Y$ (somme droite) définie par l'équation

$$y - p = f'(c) \cdot (x - c) \qquad (x \in X,\ y \in Y). \tag{7}$$

Là aussi, l'écart entre les points $\{x, f(x)\}$ et $\{x, p + f'(c)(x - c)\}$ (le premier sur la surface $y = f(x)$, le deuxième sur la variété (7)) a un ordre de petitesse supérieur à celui de $|x - c|$. En nous servant de différentielles, nous mettons l'équation sous la forme

$$dy = f'(c)\, dx. \tag{8}$$

d. En cas de fonction numérique d'une variable réelle, nous disposons d'une autre définition de la droite tangente. Précisément, nous considérons un petit angle de sommet au point $\{c, f(c)\}$ découpé par les droites de coefficients angulaires $D - \varepsilon$ et $D + \varepsilon$; la droite $y = p + D(x - c)$ est, par définition, la tangente à la courbe $y = f(x)$ au point $\{c, f(c)\}$ si, pour tout $\varepsilon > 0$ et pour $h = x - c$, $|h| \leqslant \delta(\varepsilon)$ suffisamment petits, tous les points de la courbe $y = f(x)$ sont à l'intérieur de l'angle en question (fig. 1.2-1). Une construction analogue est réalisable en cas de plusieurs variables indépendantes. Au lieu d'angle, on considère un domaine Ω_ε dans l'espace $(n + 1)$-dimensionnel $x_1, \ldots, x_n, y$, qui est défini par les inéquations

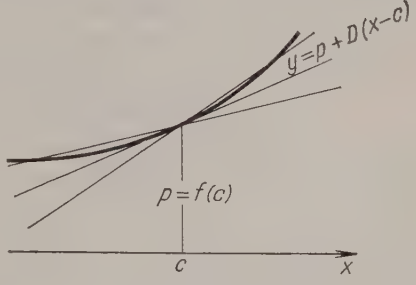


Fig. 1.2-1

$$\sum_1^n D_i (x_i - c_i) - \varepsilon |x - c| \leqslant y - p \leqslant \sum_1^n D_i (x_i - c_i) + \varepsilon |x - c|.$$

Par définition, un plan $y = p + \sum_1^n D_i (x_i - c_i)$ est le plan tangent à la surface $y = f(x)$ pour $x = c$ du moment que, quels que soient $\varepsilon > 0$ et $|h| = |x - c| \leqslant \delta(\varepsilon)$ suffisamment petits, la surface $y = f(x)$ appartient entièrement au domaine Ω_ε (fig. 1.2-2). Il est clair que cette définition du plan tangent est équivalente aux précédentes, la condition de dérivabilité de la fonction $y = f(x)$ pour $x = c$ étant équivalente à la condition d'existence du plan tangent au sens de la dernière définition.

e. Signalons encore le cas de fonction vectorielle $y = f(x)$: $R_1 \to R_n$, soit

$$\left.\begin{aligned} y_1 &= f_1(x), \\ &\dots\dots \\ y_n &= f_n(x), \end{aligned}\right\} \quad a \leqslant x \leqslant b. \tag{9}$$

Cette dernière peut elle aussi être interprétée géométriquement comme courbe dans l'espace $(n+1)$-dimensionnel $(x, y_1, \dots, y_n)$.

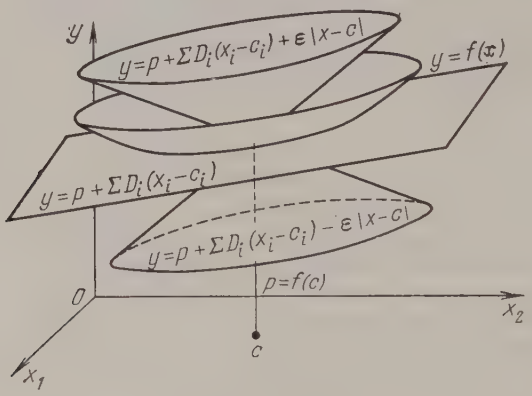


Fig. 1.2-2

La dérivée $f'(c)$ représente le vecteur $f'_1(c), \dots, f'_n(c)$. Il figure dans l'équation de la tangente (7) au point $\{c, f(c)\}$, les y, p et $f'(c)$ étant des vecteurs; dans le présent cas, l'équation (7) prend la forme du système

$$\left.\begin{aligned} y_1 - p_1 &= f'_1(c)(x-c), \\ &\dots\dots\dots\dots \\ y_n - p_n &= f'_n(c)(x-c). \end{aligned}\right\} \tag{10}$$

On peut toujours utiliser les symboles de différentielles en posant $dx = x - c$, $y_i - p_i = dy_i$:

$$dy_i = f'_i(c)\, dx. \tag{11}$$

f. Dans le cas général, une fonction $y = f(x)$: $R_n \to R_m$ peut être conçue géométriquement comme une variété courbée dans l'espace $(n+m)$-dimensionnel des points $\{x_1, \dots, x_n, y_1, \dots, y_m\}$; alors l'équation $y - p = f'(c)(x - c)$ définit une variété linéaire dans R_{n+m} qui contient le point (c, p). Il est naturel de l'appeler variété linéaire « tangente » à la variété $y = f(x)$, en se basant toujours sur le fait que, pour x proches de c, l'écart entre $f(x)$ et

$p + f'(c)(x - c)$ est un infiniment petit d'ordre supérieur par rapport à $|x - c|$.

1.27. Dérivée suivant un vecteur et une direction.

a. Soient $y = f(x)\colon V \subset X \to Y$ une fonction définie dans une boule $V = \{x \in X : |x - c| \leqslant r\}$ d'un espace normé X et $\xi \in X$ un vecteur. Considérons le segment $\Gamma \subset V$ issu du point c dans le sens du vecteur ξ qui a pour équation:

$$x = c + t\xi, \qquad 0 \leqslant t \leqslant t_0.$$

Sur ce segment, la fonction $f(x)$ devient fonction de la variable réelle t. Posons $f[x(t)] = \varphi(t)$, $\varphi\colon [0, r] \to Y$. Pour $t = 0$, on a $f(c) = \varphi(0)$. Supposons que $f(x)$ soit dérivable pour $x = c$; nous avons alors:

$$f(c + t\xi) - f(c) = f'(c)\, t\xi + o(t).$$

Il en résulte la dérivabilité pour $t = 0$ de la fonction $\varphi(t)$ et la relation

$$\varphi'(0) = f'(c)\, \xi. \tag{1}$$

La quantité $\varphi'(0)$ s'appelle *dérivée de la fonction $f(x)$ au point c suivant le vecteur* ξ et se note $f'_\xi(c)$. La même quantité est parfois désignée par $\xi * f(c)$; ici le vecteur ξ représente le symbole de l'opérateur dérivée:

$$\xi * f(c) = f'(c)\, \xi.$$

b. Lorsque ξ est un vecteur unitaire, posons $\xi = e$, $|e| = 1$; $f'_\xi(c)$ s'appelle dérivée suivant une direction (celle du segment Γ) et se désigne $f'_\Gamma(c)$. Si, par exemple, e est le vecteur de base de l'axe des x_1 (dans R_n), alors $f'_\Gamma(c)$ représente, évidemment, la dérivée partielle $\frac{\partial f(c)}{\partial x_1}$.

1.28. Pour une fonction numérique $f(x)\colon R_n \to R_1$, a lieu la relation *1.22* (4):

$$f(c + te) - f(c) = (\operatorname{grad} f(c), te) + o(t), \tag{1}$$

qui implique

$$f'_\Gamma(c) = (\operatorname{grad} f(c), e). \tag{2}$$

Nous savons que le vecteur $\operatorname{grad} f(c) = f'(c)$ est complètement déterminé par les propriétés de la fonction $f(x)$, sans dépendre de la direction suivant laquelle la dérivation est effectuée. Le vecteur e qui détermine ladite direction, à son tour, ne dépend point de la fonction $f(x)$. De la sorte, la formule (2) met en relief les rôles que jouent les deux facteurs: fonction et direction de dérivation.

La formule (2) permet de formuler certaines conclusions sur le comportement de la fonction $f(x)$ au voisinage du point c (à condition que grad $f'(c) \neq 0$). Notamment, si ω est l'angle formé par les vecteurs $G = \text{grad } f(c)$ et e, il résulte de la définition *U4.74* que

$$f'_\Gamma(c) = |G| \cos \omega. \qquad (3)$$

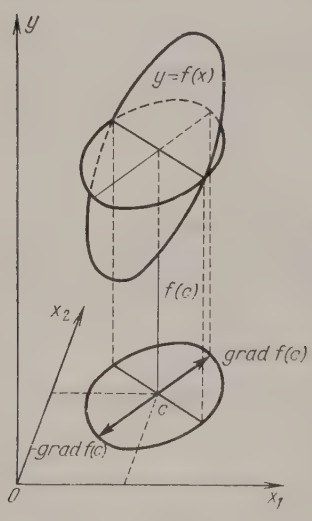


Fig. 1.2-3

On en tire les énoncés *a-d* (fig. 1.2-3, $n = 2$).

a. *Dans la direction du vecteur* $G = \text{grad } f(c)$, *la dérivée* $f'_\Gamma(c)$ *prend sa valeur maximale qui est* $|\text{grad } f(c)|$ (puisque, dans cette direction, $\cos \omega = 1$); *dans la direction opposée,* $f'_\Gamma(c)$ *prend sa valeur minimale qui est* $-|\text{grad } f(c)|$. Ce sont respectivement les *directions de la plus rapide ascendance et de la plus rapide descente de la fonction* $f(x)$ *au point* $x = c$.

b. *La quantité* $f'_\Gamma(c)$ *est nulle suivant toute direction orthogonale au gradient; l'accroissement de la fonction* $f(x)$ *suivant une telle direction est infiniment petit d'ordre supérieur par rapport à* $|h| = |x - c|$.

c. *Les valeurs de* $f'_\Gamma(c)$ *suivant toutes les directions qui restent sont comprises entre* $-|\text{grad } f(c)|$ *et* $+|\text{grad } f(c)|$.

d. Ainsi, *le gradient de la fonction* $f(x)$ *au point* $x = c$ *est le vecteur allant de c dans le sens de la plus rapide ascendance de la fonction* $f(x)$ *et dont la longueur est égale à la dérivée de* $f(x)$ *suivant cette direction.*

e. Exemple. Etudions le comportement de la fonction $y = x_2^2 - x_1^2$ ($R_2 \to R_1$) de deux variables au voisinage du point (1, 1). La valeur de la fonction à ce point même est nulle. L'accroissement de la fonction au passage à un point proche $(1 + h_1, 1 + h_2)$ admet la partie linéaire principale:

$$(1+h_2)^2 - (1+h_1)^2 = -2h_1 + 2h_2 + h_2^2 - h_1^2 = -2h_1 + 2h_2 + o(h),$$

de sorte que la fonction $x_2^2 - x_1^2$ s'avère dérivable au point (1,1) (comme, d'ailleurs, en tout autre point). Ses dérivées partielles $\frac{\partial y}{\partial x_1} = -2x_1$ et $\frac{\partial y}{\partial x_2} = 2x_2$ prennent au point (1,1) les valeurs respectives -2 et 2, de façon que grad $y(1,1) = \{-2, 2\}$; par conséquent, grad $y(1,1)$ fait avec l'axe des x_1 l'angle de 135° (fig. 1.2-4). C'est dans le sens γ_1 de ce vecteur que va la ligne de la plus rapide ascendance de la fonction $y(x_1, x_2)$, la rapidité de l'ascendance valant $|\text{grad } y(1,1)| = 2\sqrt{2}$. Dans la direction γ_2 orthogonale, c'est-à-dire suivant la bissectrice du premier quadrant, nous avons $y'_{\gamma_2}(1,1) = 0$; en effet, la fonction $y(x_1, x_2)$ ne varie pas sur cette bissectrice (elle y reste nulle). Le vecteur $-\text{grad } y(1,1) = \{2, -2\}$ indique la direction γ_3 de la ligne de la plus rapide

descente. Les dérivées partielles $\frac{\partial y\ (1,1)}{\partial x_1} = -2$, $\frac{\partial y\ (1,1)}{\partial x_2} = 2$ fournissent les valeurs de coefficients angulaires de la fonction dans le sens des axes des coordonnées correspondantes. Elles ont une simple interprétation géométrique: si l'on coupe la surface $y = x_2^2 - x_1^2$ par le plan vertical $x_2 = 1$, on obtient en

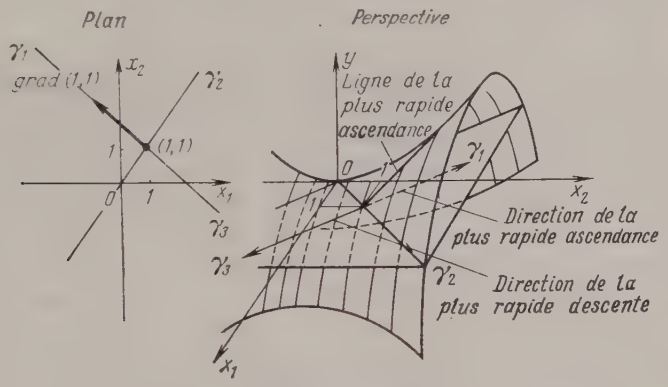


Fig. 1.2-4

section la courbe $y = 1 - x_1^2$ dont le nombre -2 est le coefficient angulaire de la tangente (dans le plan $x_2 = 1$). De même, $\frac{\partial y\ (1,1)}{\partial x_2} = 2$ est le coefficient angulaire de la tangente à la courbe $y = x_2^2 - 1$ que l'on obtient en coupant notre surface par le plan vertical $x_1 = 1$. La section de la surface par le plan vertical $x_1 + x_2 = 2$ qui contient le vecteur gradient représente une courbe de coefficient angulaire $\pm 2\sqrt{2}$.

§ 1. 3. Théorèmes généraux sur les fonctions dérivables

1.31. Dans ce qui suit, les fonctions $f(x)$, $g(x)$, . . . sont supposées agissant d'un voisinage V d'un point c d'un espace vectoriel normé X dans un espace vectoriel normé Y.

a. *Si* $f(x) = \text{const}$ (i.e. les valeurs $f(x)$, pour tous les $x \in V$, représentent un même élément de l'espace Y), *alors* $f'(c) = 0$. Ceci découle de l'égalité $f(c + h) - f(c) \equiv 0$ et de l'unicité de dérivée (*1.23*).

b. *Si* $f(x)$ *est une fonction linéaire* Fx (plus exactement, s'il existe un opérateur linéaire $F: X \to Y$ dont les valeurs sur $V(c)$ coïncident avec les valeurs respectives de $f(x)$), *alors* $f'(c) = F$, $df(c) = F\,dx$.

En effet, on a par hypothèse

$$f(c + h) - f(c) = F(c + h) - Fc = Fh,$$

et le résultat cherché découle toujours de l'unicité de dérivée (1.23).

c. En particulier, la dérivée $f'(x)$ de la fonction $f(x) = x\ (X \to X)$ est l'opérateur identique $f'(x) = E\ (X \to X)$, et $df(x) \equiv dx = dh$.

d. *Toute fonction $f(x)$ dérivable pour $x = c$ est continue pour $x = c$.*

En effet, étant donnée l'égalité

$$f(c+h) - f(c) = f'(c)\,h + o(h)$$

et $\varepsilon > 0$, choisissons $\delta > 0$ suffisamment petit pour avoir $|o(h)| < \varepsilon/2$ et $\|f'(c)\|\,|h| < \varepsilon/2$ dès que $|h| < \delta$; nous verrons que

$$|f(c+h) - f(c)| \leqslant \|f'(c)\|\,|h| + |o(h)| < \varepsilon,$$

d'où la continuité de $f(x)$ pour $x = c$.

1.32. a. *Si les fonctions $f(x): V \to Y$ et $g(x): V \to Y$ sont dérivables pour $x = c \in V$, alors il en est de même de la fonction $s(x) = f(x) + g(x)$ $(V \to Y)$; de plus:*

$$s'(c) = f'(c) + g'(c),$$

ou bien

$$ds(c) = df(c) + dg(c).$$

En effet, il résulte des égalités

$$f(c+h) - f(c) = f'(c)\,h + o(h),$$
$$g(c+h) - g(c) = g'(c)\,h + o(h)$$

que

$$s(c+h) - s(c) = f(c+h) - f(c) + g(c+h) - g(c) = [f'(c) + g'(c)]\,h + o(h),$$

d'où le résultat annoncé.

b. *Soient une fonction $y(x): V \to Y$ dérivable pour $x = a$ et A un opérateur linéaire continu de l'espace Y dans un espace (normé) Z. La fonction $z(x) = Ay(x)$ est alors dérivable pour $x = a$ et*

$$z'(a) = Ay'(a), \qquad dz(a) = A\,dy(a).$$

En effet, dans l'hypothèse formulée on a

$$z(a+h) - z(a) = A[y(a+h) - y(a)] = A[y'(a)\,h + o(h)] = [Ay'(a)]\,h + o(h)$$

ce qu'il fallait prouver.

c. Soit Y la somme droite de sous-espaces $Y_1, \ldots, Y_n$, de sorte que, pour chaque fonction $y(x): X \to Y$, tout comme dans *1.15c*, il y ait les composantes

$$y_1(x) = P_1 y(x)\ (X \to Y_1), \ \ldots, \ y_n(x) = P_n y(x)\ (X \to Y_n).$$

Si Y est un espace complet et que les sous-espaces $Y_1, \ldots, Y_n$ sont fermés, alors la dérivabilité de la fonction $y(x)$ pour $x = a$ implique la dérivabilité pour $x = a$ de chaque fonction $y_k(x)$ $(k = 1, \ldots$

$\ldots, n)$; cela résulte de la continuité de l'opérateur P_k (*1.14d*) et de *b*.

Inversement, si les fonctions $y_1(x)$ $(X \to Y_1)$, $\ldots$, $y_n(x)$ $(X \to Y_n)$ sont dérivables pour $x = a$, il en est de même de la fonction $y(x) = \{y_1(x), \ldots, y_n(x)\}$: $X \to Y$, ce qui résulte de *a* puisque

$$y(x) = y_1(x) + \ldots + y_n(x). \tag{1}$$

La dérivée $y'(a)$ est un opérateur linéaire $X \to Y$, soit $y'(a) \in L(X, Y)$; de même $y_k'(a) \in L(X, Y_k)$ $(k = 1, \ldots, n)$. L'espace $L(X, Y)$ est la somme droite des sous-espaces $L(X, Y_k)$ (*1.14f*), et il découle immédiatement de (1) que les composantes de l'élément $y'(a) \in L(X, Y)$ sont les quantités $y_k'(a)$:

$$y'(a) = \{y_1'(a), \ldots, y_n'(a)\}. \tag{2}$$

1.33. Dérivée et différentielle d'une fonction composée.

a. Théorème. *Soient $y = y(x)$ une fonction qui applique un domaine G d'un espace normé X dans un espace normé Y, $a \in G$, $y(a) = b$, et $z = z(y)$ une fonction définie au voisinage du point b de l'espace Y, à valeurs dans un espace normé Z. Si la fonction $y(x)$ est dérivable pour $x = a$, et la fonction $z(y)$ pour $y = b$, alors leur composée $\zeta(x) \equiv z[y(x)]$, définie dans un voisinage du point $a \in G$ et à valeurs dans Z, est dérivable elle aussi pour $x = a$; on a*

$$\zeta'(a) = z'(b)\, y'(a). \tag{1}$$

Notons que $y'(a)$ et $z'(b)$ sont des opérateurs linéaires de X dans Y et de Y dans Z respectivement, de sorte que le second membre dans (1) soit défini en tant qu'opérateur linéaire de X dans Z.

En abordant la démonstration nous avons:

$$\begin{aligned}
&\zeta(a+h) - \zeta(a) \equiv z[y(a+h)] - z[y(a)] = \\
&= z'[y(a)]\,[y(a+h) - y(a)] + o[y(a+h) - y(a)] = \\
&= z'(b)\,[y'(a)\,h + o(h)] + o[y'(a)\,h + o(h)] = \\
&\qquad = z'(b)\, y'(a)\, h + o(h). \qquad (2)
\end{aligned}$$

Par conséquent, l'expression $z'(b)\, y'(a)\, h$ constitue la partie linéaire principale de l'accroissement de la fonction $\zeta(x)$ en passant de $x = a$ à $x = a + h$, ce qu'il fallait prouver.

En remplaçant a par x, b par $y(x)$ et passant aux opérateurs, on peut mettre la formule obtenue sous la forme:

$$\{z[y(x)]\}' = z'(y)\, y'(x). \tag{3}$$

b. Supposons, par exemple, que les espaces X, Y, Z soient de dimensions finies n, m et p respectivement. Après avoir fixé d'une

quelconque façon leurs bases, les fonctions $y(x)$ et $z(y)$ sont données par les systèmes d'égalités numériques:

$$\left.\begin{array}{l} y_1 = y_1(x_1, \ldots, x_n), \quad z_1 = z_1(y_1, \ldots, y_m), \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_m = y_m(x_1, \ldots, x_n); \quad z_p = z_p(y_1, \ldots, y_m). \end{array}\right\} \tag{4}$$

Aux opérateurs $y'(x)$ et $z'(y)$ il correspond les matrices de Jacobi (*1.25a*):

$$y'(x) \cong \left\| \begin{array}{ccc} \frac{\partial y_1}{\partial x_1} & \ldots & \frac{\partial y_1}{\partial x_n} \\ \ldots & \ldots & \ldots \\ \frac{\partial y_m}{\partial x_1} & \ldots & \frac{\partial y_m}{\partial x_n} \end{array} \right\|, \quad z'(y) \cong \left\| \begin{array}{ccc} \frac{\partial z_1}{\partial y_1} & \ldots & \frac{\partial z_1}{\partial y_m} \\ \ldots & \ldots & \ldots \\ \frac{\partial z_p}{\partial y_1} & \ldots & \frac{\partial z_p}{\partial y_m} \end{array} \right\|. \tag{5}$$

La fonction $\zeta(x) = z[y(x)]$, d'après a, est dérivable; la matrice de Jacobi correspondante, en vertu de (1), est identique au produit des matrices (5):

$$\zeta'(x) \cong \left\| \begin{array}{ccc} \frac{\partial \zeta_1}{\partial x_1} & \ldots & \frac{\partial \zeta_1}{\partial x_n} \\ \ldots & \ldots & \ldots \\ \frac{\partial \zeta_p}{\partial x_1} & \ldots & \frac{\partial \zeta_p}{\partial x_n} \end{array} \right\| = \left\| \begin{array}{ccc} \frac{\partial z_1}{\partial y_1} & \ldots & \frac{\partial z_1}{\partial y_m} \\ \ldots & \ldots & \ldots \\ \frac{\partial z_p}{\partial y_1} & \ldots & \frac{\partial z_p}{\partial y_m} \end{array} \right\| \left\| \begin{array}{ccc} \frac{\partial y_1}{\partial x_1} & \ldots & \frac{\partial y_1}{\partial x_n} \\ \ldots & \ldots & \ldots \\ \frac{\partial y_m}{\partial x_1} & \ldots & \frac{\partial y_m}{\partial x_n} \end{array} \right\|. \tag{6}$$

L'égalité (6) s'appelle *règle de multiplication des matrices de Jacobi*. En abrégé, elle se note comme suit:

$$\left\| \frac{\partial(\zeta_1, \ldots, \zeta_p)}{\partial(x_1, \ldots, x_n)} \right\| = \left\| \frac{\partial(z_1, \ldots, z_p)}{\partial(y_1, \ldots, y_m)} \right\| \left\| \frac{\partial(y_1, \ldots, y_m)}{\partial(x_1, \ldots, x_n)} \right\|. \tag{7}$$

En particulier, quels que soient $j = 1, \ldots, p$; $i = 1, \ldots, n$, on a

$$\frac{\partial \zeta_j}{\partial x_i} = \sum_{k=1}^{m} \frac{\partial z_j}{\partial y_k} \cdot \frac{\partial y_k}{\partial x_i}. \tag{8}$$

Considérons le cas $n = 1$ où l'on peut désigner $x_1 = \ldots = x_n = x$. Ici, les fonctions $y(x)$ et $\zeta(x)$ sont des fonctions d'une seule variable x. La formule (8) prend la forme

$$\frac{d\zeta_j}{dx} = \sum_{k=1}^{m} \frac{\partial z_j}{\partial y_k} \frac{dy_k}{dx}. \tag{9}$$

Nous voyons qu'on peut avoir besoin de dérivées partielles en dérivant des fonctions d'une variable.

c. En remplaçant dans (2) h par dx et compte tenu du fait que la différentielle d'une fonction est la partie linéaire principale de

son accroissement, on trouve

$$d\zeta = z'(b)\, y'(a)\, dx.$$

Or, pour dx donné, on a également

$$dy = y'(a)\, dx,$$

d'où

$$d\zeta = z'(b)\, dy,$$

c'est-à-dire que la différentielle d'une fonction de y a une même forme que y soit variable indépendante ou fonction d'une autre variable x. Cette propriété est appelée *invariance d'une différentielle par rapport au changement de la variable*. Soulignons que, dans le premier cas, on sous-entend par dy l'accroissement arbitraire de la variable y, et dans le deuxième, la valeur de la différentielle de la fonction $y(x)$ pour le vecteur dx.

1.34. Différentielle d'un produit généralisé.

a. Soient deux espaces normés X et Y et un produit généralisé $\langle x, y\rangle$ des éléments $x \in X$, $y \in Y$, i.e. une application bilinéaire continue de l'espace $W = X + Y$ dans un espace Z (*1.15d*).

La forme bilinéaire $\langle x, y\rangle$ étant continue, il existe un $C > 0$ tel que, pour n'importe quels x, y:

$$|\langle x, y\rangle| \leqslant C\,|x|\,|y|.$$

Nous allons voir que *la fonction* $z = \langle x, y\rangle : W \to Z$ *est dérivable en tout point de l'espace* W *et que*

$$dz = \langle dx, y\rangle + \langle x, dy\rangle. \qquad (1)$$

En effet, si l'on donne à la variable $\{x, y\} \in W$ un accroissement $\{dx, dy\}$, et que l'on se sert de la bilinéarité de $\langle x, y\rangle$, on trouve :

$$\langle x + dx,\ y + dy\rangle - \langle x, y\rangle =$$
$$= \langle x, y\rangle + \langle dx, y\rangle + \langle x, dy\rangle + \langle dx, dy\rangle - \langle x, y\rangle =$$
$$= \langle dx, y\rangle + \langle x, dy\rangle + \langle dx, dy\rangle.$$

Les deux premiers termes à droite sont linéaires par rapport au déplacement $\{dx, dy\}$, l'expression $\langle dx, dy\rangle$ admettant la majoration :

$$|\langle dx, dy\rangle| \leqslant C\,|dx|\,|dy| \leqslant \frac{C'}{2}\,(|dx|^2 + |dy|^2) = o(|dx| + |dy|).$$

De la sorte, l'accroissement de la fonction $\langle x, y\rangle$ admet la partie linéaire principale $\langle dx, y\rangle + \langle x, dy\rangle$, ce qui achève la démonstration.

Les formules de dérivation de divers produits que l'on va examiner servent d'exemples d'application de la règle générale que l'on vient d'établir, ainsi que de la règle de dérivation d'une fonction composée.

b. Théorème. *Supposons que, dans un domaine G d'un espace T, soient données deux fonctions dérivables $x = x(t)$, $x: G \to X$ et $y = y(t)$, $y: G \to Y$. Composons, comme dans a, le produit généralisé*

$$\langle x(t),\ y(t)\rangle : G \to Z$$

qui s'appelle produit généralisé de la fonction $x(t)$ par la fonction $y(t)$.

Nous affirmons que la fonction $\zeta(t) = \langle x(t),\ y(t)\rangle$ est dérivable dans G et que

$$d\zeta = \langle x'(t)\,dt,\ y(t)\rangle + \langle x(t),\ y'(t)\,dt\rangle. \tag{2}$$

En effet, la fonction $\zeta(t)$ peut être traitée comme fonction composée:

$$\{x, y\} = \{x(t), y(t)\} : G \to W;$$
$$\zeta(t) = \langle x, y\rangle = z(\{x, y\}) : W \to Z.$$

En appliquant le théorème sur la différentielle d'une fonction composée et le résultat a, on obtient

$$d\zeta = \langle x'(t)\,dt,\ y(t)\rangle + \langle x(t),\ y'(t)\,dt\rangle,$$

ce qu'on cherche.

Il suit de (2) que la dérivée de la fonction $\zeta(t)$ est déterminée par la formule:

$$\zeta'(t) = \langle x'(t),\ y(t)\rangle + \langle x(t),\ y'(t)\rangle, \tag{3}$$

où les termes dans le second membre sont les opérateurs linéaires de T dans Z définis comme suit:

$$\langle x'(t),\ y(t)\rangle\,dt = \langle x'(t)\,dt,\ y(t)\rangle,$$
$$\langle x(t),\ y'(t)\rangle\,dt = \langle x(t),\ y'(t)\,dt\rangle.$$

c. Conséquence. *Si une fonction $x(t): G \subset T \to X$ est dérivable pour $t = c$ et si $\lambda(t): G \subset T \to L(X, Y)$ est une fonction opératorielle dérivable pour $t = c$, le produit $g(t) = \lambda(t)\,x(t): G \subset T \to Y$ est dérivable pour $t = c$ et l'on a*

$$dg(c) \equiv g'(c)\,dt = \lambda'(c)\,dt \cdot x(c) + \lambda(c) \cdot x'(c)\,dt. \tag{4}$$

Il est aisé de voir que les deux termes dans le second membre appartiennent à l'espace Y.

d. Conséquence. *Si $\lambda(t): G \subset T \to R_1$ et $x(t): G \subset T \to R_1$ sont deux fonctions numériques dérivables pour $t = c \in G$, alors il en est de même du produit $g(t) = \lambda(t)\,x(t)$ et l'on a:*

$$g'(c) = \lambda'(c)\,x(c) + \lambda(c)\,x'(c). \tag{5}$$

Ici les deux membres sont des fonctionnelles linéaires sur l'espace T.

En tant que démonstration, remarquons que dans le cas en question il est possible de permuter les facteurs numériques $\lambda'(c)\,dt$ et $x(c)$ dans (4), après quoi on obtient

$$dg(c) \equiv g'(c)\,dt = x(c)\,\lambda'(c)\,dt + \lambda(c)\,x'(c)\,dt = \\ = [x(c)\,\lambda'(c) + \lambda(c)\,x'(c)]\,dt,$$

d'où (5).

1.35. La fonction x^{-1} et sa dérivée.

a. Soient x une application d'un ensemble U dans un ensemble V et y une application de l'ensemble V dans l'ensemble U. Ainsi donc, sont définies deux fonctions composées : $xy\colon V \to V$ et $yx\colon U \to U$. Si xy est l'application identique e_V de l'ensemble V sur lui-même, alors y s'appelle l'*inverse à droite de x et x l'inverse à gauche de y*. D'une façon analogue, si yx est l'application identique e_U de l'ensemble U, alors x s'appelle l'*inverse à droite de y et y l'inverse à gauche de x*. Une application x peut ne posséder aucun inverse à droite ou bien en posséder une infinité (*L4.76*). Mais, lorsque l'application $x\colon U \to V$ possède un inverse à droite $y\colon V \to U$ et un inverse à gauche $z\colon V \to U$, on a $z = ze_V = z(xy) = (zx)\,y = e_U y = y$; par conséquent, x possède un *unique inverse à droite* et un *unique inverse à gauche* qui sont identiques ; tout naturellement, cette unique application s'appelle *application inverse par rapport à x* et se note x^{-1}. Ainsi, $x^{-1}x = e_U$, $xx^{-1} = e_V$.

b. Supposons à présent que U et V soient des espaces normés complets et que toutes les applications considérées soient opérateurs linéaires bornés. Certains des opérateurs linéaires $x\colon U \to V$ sont inversibles ; ils forment un ensemble que nous désignons G. Sur cet ensemble est définie la fonction $x^{-1}\colon G \to L(V, U)$. Montrons que, dans l'espace métrique $L(U, V)$ (doté de la métrique opératorielle ordinaire (*U12.71b*)), l'ensemble G est un domaine (ensemble ouvert) et que la fonction x^{-1} est continue sur G. A cet effet, nous nous servirons du fait que les éléments de l'algèbre $L(V)$ proches de l'unité sont inversibles (*U12.82a*). En cherchant l'inverse d'un élément $x + h$, pour h petit, nous multiplierons d'abord $x + h$ par x^{-1}, verrons que le produit sera proche de l'unité, donc inversible ; puis nous montrerons que l'élément $x + h$ sera inversible lui aussi. Réalisons cette idée en détail : pour tout $h \in L(U, V)$ on a $(x+h)\,x^{-1} = \\ = xx^{-1} + hx^{-1} = e_V + hx^{-1}$; par conséquent, lorsque $|h| < \\ < 1/|x^{-1}|$ on a l'inégalité $|(x + h)\,x^{-1} - e_V| = |hx^{-1}| = |h| \times \\ \times |x^{-1}| < 1$, de sorte que l'élément $(x + h)\,x^{-1}$ est éloigné de e_V moins que d'une unité. Dans ces conditions l'opérateur $(x + h)\,x^{-1}\colon V \to V$ est inversible dans l'espace V (*U12.82a*), donc il existe l'opérateur $z_h\colon V \to V$ pour lequel $(x + h)\,x^{-1}z_h = e_V$. Ceci veut dire que $x^{-1}z_h$ est un inverse à droite de $x + h$. De même, en remplaçant $(x + h)\,x^{-1}$ par $x^{-1}(x + h)$ nous démontrons l'existence, pour $|h| < 1/|x^{-1}|$ d'un inverse à gauche de $x + h$. Il résulte mainte-

nant de a que l'élément $x + h$ est inversible pour $|h| < 1/|x^{-1}|$. Nous voyons que l'ensemble G avec tout son élément x contient tous les autres éléments de $L(U, V)$ qui sont éloignés de x d'une distance inférieure à $1/|x^{-1}|$; l'ensemble G est donc ouvert. Ensuite, on sait que $z_h \to e_V$ pour $h \to 0$ (*U12.82b*), d'où $(x + h)^{-1} = (x + h)^{-1} e_V = (x + h)^{-1} (x + h) x^{-1} z_h = x^{-1} z_h \to x^{-1}$ pour $h \to 0$, et la fonction x^{-1} est continue sur son domaine de définition.

c. Montrons que, dans l'hypothèse de *b*, *la fonction* x^{-1} *est dérivable* et trouvons sa différentielle.

En partant de l'identité

$$(x + h)\,[(x + h)^{-1} - x^{-1}]\,x = x - (x + h) = -h,$$

on a

$$(x + h)^{-1} - x^{-1} = -(x + h)^{-1} hx^{-1} = -x^{-1}hx^{-1} + o(h)$$

en vertu de la continuité de la fonction x^{-1} dans le domaine $G(b)$. Il en résulte la dérivabilité de la fonction x^{-1} sur son domaine de définition et l'égalité

$$d(x^{-1}) = (x^{-1})'\,h = -x^{-1}hx^{-1}.$$

Notons que, en général, on ne peut permuter les facteurs dans le résultat obtenu. Même si $U = V$, donc, si la permutation formelle de facteurs a un sens, on ne peut la réaliser à cause de non-commutativité éventuelle des opérateurs dans $L(U)$. Dans le seul cas $U = V = R_1$, on peut permuter les facteurs sans réserve; on retrouve alors la formule classique

$$d\left(\frac{1}{x}\right) = -\frac{dx}{x^2},$$

ou bien

$$\left(\frac{1}{x}\right)' = -\frac{1}{x^2}.$$

1.36. Dérivée d'un rapport. Soient $f(x)$ et $g(x)$ des fonctions numériques dérivables définies dans une boule $V = \{x \in X : |x - a| < r\}$ d'un espace normé X; soit ensuite $g(x) \neq 0$. Montrons que le rapport $f(x)/g(x)$ est une fonction numérique dérivable dans V et trouvons sa dérivée. La fonction $1/g(x)$ $(V \to R_1 - 0)$ peut être traitée comme composée de la fonction dérivable donnée $u = g(x)$ $(V \to R_1 - 0)$ et de la fonction dérivable $1/u$ $(R_1 - 0 \to R_1 - 0)$; d'après *1.33a* et *1.35c*, la fonction $1/g(x)$ est dérivable et sa dérivée vaut

$$\left(\frac{1}{g(x)}\right)' = -\frac{1}{u^2} g'(x) = -\frac{1}{g^2(x)} g'(x). \qquad (1)$$

En appliquant *1.34d*, on obtient que le rapport $f(x)/g(x)$ est dérivable lui aussi avec pour dérivée

$$\left(\frac{f(x)}{g(x)}\right)' = \frac{f'(x)\,g(x) - f(x)\,g'(x)}{g^2(x)} \tag{2}$$

et pour différentielle

$$d\left(\frac{f(x)}{g(x)}\right) = \frac{df(x)\,g(x) - f(x)\,dg(x)}{g^2(x)}. \tag{3}$$

Dans le cas où $X = R_n$, on se rappelle *1.24* et l'on note

$$\frac{\operatorname{grad} f(x)}{g(x)} = \frac{\operatorname{grad} f(x)\cdot g(x) - f(x)\cdot \operatorname{grad} g(x)}{g^2(x)}. \tag{4}$$

1.37. Exemples.

a. Si l'on veut trouver les dérivées partielles $\frac{\partial u}{\partial x}$ et $\frac{\partial u}{\partial y}$ de la fonction $u = \operatorname{arctg}(y/x)$ $(R_2 \to R_1)$, on peut utiliser la propriété *1.33c* et calculer sa différentielle comme celle de la fonction d'une variable y/x puis appliquer *1.36*(3):

$$d\left(\operatorname{arctg}\frac{y}{x}\right) = \frac{1}{1+(y/x)^2}\,d\left(\frac{y}{x}\right) = \frac{x^2}{x^2+y^2}\,\frac{x\,dy - y\,dx}{x^2} = \frac{x\,dy - y\,dx}{x^2+y^2}.$$

On en tire les dérivées en question comme coefficients devant les différentielles respectives:

$$\frac{\partial u}{\partial x} = -\frac{y}{x^2+y^2}, \quad \frac{\partial u}{\partial y} = \frac{x}{x^2+y^2}.$$

Le résultat obtenu peut être écrit à l'aide du symbole de gradient:

$$\operatorname{grad}\operatorname{arctg}\frac{y}{x} = \left\{-\frac{y}{x^2+y^2}, \quad \frac{x}{x^2+y^2}\right\}.$$

b. Soient y un point fixe et x un point variable de l'espace R_n. Trouvons les dérivées partielles de la fonction $r = |x - y| = \sqrt{\sum (x_i - y_i)^2}$. En dérivant les deux membres de l'égalité

$$r^2 = (x_1 - y_1)^2 + \ldots + (x_n - y_n)^2,$$

on obtient

$$2r\,dr = 2(x_1 - y_1)\,dx_1 + \ldots + 2(x_n - y_n)\,dx_n,$$

d'où

$$dr = \frac{1}{r}[(x_1 - y_1)\,dx_1 + \ldots + (x_n - y_n)\,dx_n],$$

donc

$$\frac{\partial r}{\partial x_j} = \frac{x_j - y_j}{r}.$$

1.38. Différentielles partielles. Plus haut (*1.22*), on désignait la dérivée partielle, par exemple par rapport à la variable x_1, d'une fonction $u(x) \equiv u(x_1, \ldots, x_n)$: $(R_n \to R_1)$ par le symbole $\frac{\partial u}{\partial x_1}$. On peut

attribuer un sens aux quantités ∂u et ∂x_1 : ∂x_1 est l'accroissement de la coordonnée x_1 et ∂u la différentielle partielle correspondante, i.e. la partie linéaire principale de l'accroissement correspondant de la fonction, lorsque $x_2, \ldots, x_n$ ne varient pas. Néanmoins, on ne voit pas dans le symbole ∂u quelles sont les coordonnées qui varient et quelles sont les coordonnées constantes ; pour cette raison, ce symbole peut changer de sens suivant les cas, ce qui provoque des confusions. Signalons quelques exemples de paradoxes qui surgissent lorsqu'on traite formellement des différentielles partielles, sans tenir compte de leur sens exact.

a. En simplifiant par ∂x_i dans la formule de dérivation d'une fonction composée $u = u(x_1, \ldots, x_n)$, $x_i \in x_i(t)$ (*1.33b*)

$$\frac{\partial u}{\partial t} = \sum_{i=1}^{n} \frac{\partial u}{\partial x_i} \cdot \frac{\partial x_i}{\partial t} ;$$

on arrive au résultat dépourvu de sens : $\frac{\partial u}{\partial t} = n \frac{\partial u}{\partial t}$.

b. Si x, y sont les coordonnées cartésiennes, et r, φ, coordonnées polaires d'un point du plan, on a :

$$r = \sqrt{x^2 + y^2}, \quad r = x \sec \varphi.$$

De la première expression on tire $\frac{\partial r}{\partial x} = \frac{x}{\sqrt{x^2+y^2}}$, de la seconde, $\frac{\partial r}{\partial x} = \sec \varphi$.

Or, pour $x > 0$, $y > 0$, ces résultats sont différents car, dans ce cas, $\frac{x}{\sqrt{x^2+y^2}} < 1$, $\sec \varphi > 1$.

c. Soit $z = x + y$; alors $\frac{\partial z}{\partial x} = \frac{\partial z}{\partial y} = 1$. De la même équation on a $x = z - y$, $y = z - x$, donc

$$\frac{\partial x}{\partial y} = -1, \quad \frac{\partial x}{\partial z} = 1, \quad \frac{\partial y}{\partial z} = 1, \quad \frac{\partial y}{\partial x} = -1, \qquad 168$$

d'où

$$\frac{\partial x}{\partial y} \cdot \frac{\partial y}{\partial z} \cdot \frac{\partial z}{\partial x} = (-1) \cdot 1 \cdot 1 = -1.$$

La même expression se ramène à $+1$ en simplifiant par ∂x, ∂y, ∂z.

En réalité, toutes les n quantités ∂u de l'exemple *a* sont en général différentes ; dans l'exemple *b*, le premier ∂r correspond au déplacement avec y constant, le second avec φ constant ; dans l'exemple *c*, toutes les six quantités dans le dernier produit ont les sens différents. Donc, lorsqu'il faut considérer les différentielles partielles, on doit tenir compte de leur sens exact *.

1.39. Dérivée suivant une ligne.

a. Soit $\Gamma = \{x \in X : x = x(t), \ \alpha \leqslant t \leqslant \beta\}$ une courbe lisse

---

* Dans les manuels, on recommande habituellement de traiter le symbole $\frac{\partial u}{\partial x_j}$ comme un tout inséparable, sans affecter un sens à part au numérateur ou au dénominateur.

dans un domaine $G \subset X$. Ceci veut dire que la fonction $x(t)$ $(R_1 \to G)$ admet la dérivée $x'(t)$ qui est continue pour $\alpha \leqslant t \leqslant \beta$ (aux extrémités $t = \alpha$ et $t = \beta$, il s'agit naturellement des dérivées unilatérales, *U7.19*). Le vecteur $x'(t)$ est tangent à la ligne Γ au point $x(t)$ (*U16.13*); nous l'appellerons vecteur directeur de la tangente à la courbe Γ, ou vecteur directeur de la courbe Γ, tout court. Soit ensuite $x(\alpha) = a \in G$, $x(\beta) = b \in G$.

Considérons une fonction $y(x)$, $y: G \to Y$ qui est supposée dérivable au moins en tout point de la courbe Γ et posons

$$\varphi(t) = y[x(t)] \quad (R_1 \to Y).$$

Cette fonction est dérivable en t d'après le théorème *1.33a*. Sa dérivée s'appelle *dérivée de la fonction $y(x)$ suivant la courbe Γ*. Vu la formule *1.33* (3), on a

$$\varphi'(t) = y'(x) \cdot x'(t) \quad (x = x(t)), \tag{1}$$

c'est-à-dire que la dérivée de la fonction $y(x)$ suivant la courbe Γ coïncide avec sa dérivée suivant le vecteur directeur de Γ (*1.27*). Dans le cas où la courbe Γ est le segment $x = c + te$, $0 \leqslant t \leqslant 1$, $|e| = 1$, la dérivée suivant Γ, pour $t = 0$, est la dérivée suivant la direction Γ (*1.27a*).

b. Supposons que, dans *a*, la courbe Γ appartienne à une surface de niveau Q (*1.11d*) d'une fonction dérivable $f(x)$. Soit, par exemple, $x(\gamma) = c$. La dérivée de la fonction $f(x)$ suivant cette courbe au point c vaut $f'(c) \cdot x'(\gamma)$ conformément à la formule (1); or, elle est nulle puisque $f(x)$ est constante sur toute la surface Q, en particulier, sur la courbe Γ. Ainsi

$$f'(c) \cdot x'(\gamma) = 0,$$

donc l'opérateur $f'(c)$ s'annule sur tous les vecteurs tangents aux courbes de la surface Q en point c; autrement dit, *$f'(c)$ s'annule sur tout vecteur du plan tangent à la surface Q en point c*. On exprime ce fait d'une façon suivante: *en tout point d'une surface de niveau de la fonction $f(x)$, le gradient de $f(x)$ est orthogonal à cette surface de niveau*. Une telle terminologie est rendue plausible par le cas de fonction numérique $f(x)$ dans un espace hilbertien, en particulier, dans un espace euclidien de dimension finie où l'opérateur $f'(a)$ agit d'après la formule *1.22* (5) qui contient le produit scalaire:

$$f'(a)\, h = (\operatorname{grad} f(a), h).$$

Le fait signalé permet d'établir facilement le sens du gradient

lorsqu'on connaît les surfaces de niveau de la fonction numérique en question. Ainsi, pour une fonction numérique de la forme $f(|x-a|)$, le gradient à tout point $x \neq a$ est orthogonal à la sphère $|x-a| = r$ qui représente une surface de niveau de ladite fonction, c'est dire que le gradient est dirigé suivant le rayon menant du point a au point x.

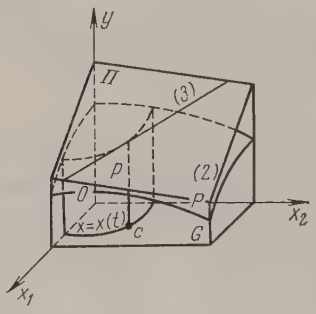


Fig. 1.3-1

c. Appliquons la formule (1) à l'étude de plus détaillée qu'avant (*1.26c*) du plan tangent Π à une surface $P = \{x \in G,\ y \in Y:\ y = y(x)\}$ de l'espace $G \times Y$. Au point $x = c$, le plan tangent a pour équation (*1.26* (7)):

$$y - p = y'(c)(x - c) \quad (p = y(c)). \tag{2}$$

Menons par le point c, dans le domaine G, toutes les courbes dérivables possibles (fig. 1.3-1). Leurs équations sont de la forme $x = x(t)$ où $x(\gamma) = c$. A l'aide de l'équation $y = y(x)$, ces courbes sont « transportées » sur la surface P et y ont les représentations paramétriques $x = x(t)$, $y = y[x(t)] = \varphi(t)$. La tangente à n'importe quelle de ces courbes au point $x = c$, $y = p$ est décrite par les équations

$$x - c = x'(\gamma)(t - \gamma), \quad y - p = \varphi'(\gamma)(t - \gamma).$$

Avec la formule (1), nous mettons ces équations sous la forme

$$x - c = x'(\gamma)(t - \gamma), \quad y - p = y'(c) \cdot x'(\gamma)(t - \gamma). \tag{3}$$

On voit que *toute droite* (3) *appartient au plan* (2).

Inversement, *toute droite passant dans le plan* (2) *par le point* $\{c, p\}$ *est la tangente, en ce point, à une courbe sur la surface* P. En effet, si $x = c_1$, $y = p_1 = p + y'(c)(c_1 - c)$ est n'importe quel autre point du plan Π, alors la droite passant par ces deux points a pour équation

$$x = c + (c_1 - c)\,t, \quad y = p + y'(c)(c_1 - c)\,t,$$

or, cette droite est la tangente à la courbe

$$x = c + (c_1 - c)\,t, \quad y = y[c + (c_1 - c)\,t]$$

se trouvant sur la surface P.

Définitivement, *le plan tangent* Π *est la réunion des droites tangentes à toutes les courbes dérivables qui passent sur la surface* P *par le point* $\{c, p\}$.

d. Dans les conditions de a, la fonction $\varphi(t) = y[x(t)]$ s'interprète géométriquement par une courbe dérivable L dans l'espace Y. L'égalité (1) montre que, pour tout $t \in [\alpha, \beta]$, le vecteur tangent $\varphi'(t)$ (au point correspondant de la courbe L) se déduit en appliquant l'opérateur linéaire $y'(x(t))$ au vecteur $x'(t)$ tangent à la courbe Γ. Si l'on considère dans le domaine G la famille de toutes les courbes dérivables $x = \overline{x}(t)$ ayant les valeurs communes $\overline{x}(t_0) = x_0 \in X$ et $\overline{x}'(t_0) = z_0 \in X$, alors l'application $y(x)$ la transforme en une famille de courbes dérivables $y = \overline{\varphi}(t)$, toujours avec les valeurs communes $\overline{\varphi}(t_0) = y(x_0)$ et $\overline{\varphi}'(t_0) = y'(x_0) = z_0$. On peut dire que $y'(x_0)$ transforme (linéairement) les vecteurs tangents $x'(t_0)$ en les vecteurs tangents $\varphi'(t_0)$; cela caractérise l'opérateur $y'(x_0)$ d'une manière plus précise que le seul fait que $y'(x_0) \in L(X, Y)$.

e. Soit P une surface lisse dans un domaine $G \subset X$; en d'autres termes, on a une fonction $x = x(u)$, $u = (u_1, \ldots, u_n) \in Q \subset R_n$ continûment dérivable dans le domaine Q, à valeurs dans le domaine G. Si les vecteurs $\frac{\partial z}{\partial u_1}, \ldots, \frac{\partial x}{\partial u_n}$ sont linéairement indépendants en un point $u = (u_1, \ldots, u_n) \in Q$, le point $x(u)$ est dit *ordinaire* pour la surface P; une surface P formée par seuls les points ordinaires est appelée surface *n-dimensionnelle*.

Toute courbe lisse L sur la surface P peut être déterminée par des équations de la forme $x = x(u_1, \ldots, u_n)$ où $u_j = u_j(t)$, $j = 1, \ldots, n$, $\alpha \leqslant t \leqslant \beta$. Un vecteur tangent à la courbe L au point $a = x(u_1^0, \ldots, u_n^0) \in L$ qui correspond à une valeur $t = t_0$ est donné par l'égalité

$$\frac{dx(a)}{dt} = \sum_{i=1}^{n} \frac{\partial x(a)}{\partial u_i} u_i'(t_0),$$

i.e. représente une combinaison linéaire des vecteurs $\frac{\partial x(a)}{\partial u_1}, \ldots, \frac{\partial x(a)}{\partial u_n}$. N'importe quelle combinaison linéaire de ces vecteurs, à son tour, peut être obtenue de cette façon; si les coefficients de la combinaison linéaire sont $C_1, \ldots, C_n$, alors le vecteur $\sum_{i=1}^{n} C_i \frac{\partial x(a)}{\partial u_i}$ est tangent à une courbe déterminée, par exemple, par les équations

$$x = x(u), \quad u_i = u_i^0 + C_i t.$$

Nous voyons que les vecteurs tangents à toutes les courbes lisses

sur la surface P qui passent par un point donné a remplissent une variété linéaire n-dimensionnelle $\Pi(a)$ dans X ; elle s'appelle *variété linéaire tangente à la surface P* au point a.

f. Exemple. Dans R_3, les équations

$$x_1 = \sin\theta\cos\varphi, \quad x_2 = \sin\theta\sin\varphi, \quad x_3 = \cos\theta,$$

avec deux paramètres θ et φ, déterminent une surface bidimensionnelle P (la sphère de rayon 1 centrée en l'origine des coordonnées). Construisons le plan tangent à cette surface en un point $a(\theta_0, \varphi_0)$. Les vecteurs de base $\frac{\partial x}{\partial\theta}$, $\frac{\partial x}{\partial\varphi}$ sont :

$$\frac{\partial x}{\partial\theta} = \{\cos\theta\cos\varphi, \cos\theta\sin\varphi, \sin\theta\},$$

$$\frac{\partial x}{\partial\varphi} = \{-\sin\theta\sin\varphi, \sin\theta\cos\varphi, 0\}.$$

Le plan passant par le point $a(\theta_0, \varphi_0)$ et contenant les vecteurs $\frac{\partial x(a)}{\partial\theta}$, $\frac{\partial x(a)}{\partial\varphi}$ a pour équation

$$\begin{vmatrix} x_1 - x_1(\theta_0, \varphi_0) & x_2 - x_2(\theta_0, \varphi_0) & x_3 - x_3(\theta_0, \varphi_0) \\ \cos\theta_0\cos\varphi_0 & \cos\theta_0\sin\varphi_0 & -\sin\theta_0 \\ -\sin\theta_0\sin\varphi_0 & -\sin\theta_0\cos\varphi_0 & 0 \end{vmatrix} = 0.$$

C'est le plan tangent cherché.

§ 1.4. Théorème de la moyenne

1.41. Considérons, dans un domaine $G \subset X$, une courbe lisse

$$L = \{x \in G : x = x(t), \ \alpha \leqslant t \leqslant \beta\}; \ x(\alpha) = a, \quad x(\beta) = b.$$

Soit, en outre, une fonction dérivable $y = y(x) : G \to Y$. Supposons d'abord que $Y = R_1$, c'est-à-dire que $y(x)$ soit une fonction numérique.

a. Théorème de la moyenne. *Si la fonction numérique $y(x)$, $y : G \to R_1$ est dérivable aux points de la courbe L, alors il existe $\theta \in (\alpha, \beta)$ tel que*

$$y(b) - y(a) = y'(c)\, x'(\theta)\,(\beta - \alpha) \quad (c = x(\theta)). \tag{1}$$

Démonstration. Posons, comme dans *1.39a*, $\varphi(t) = y[x(t)]$, $\alpha \leqslant t \leqslant \beta$. La fonction $\varphi(t) : R_1 \to R_1$ étant dérivable, nous avons d'après le théorème de Lagrange (*U7.44*)

$$y(b) - y(a) = \varphi(\beta) - \varphi(\alpha) = \varphi'(\theta)\,(\beta - \alpha)$$

pour un $\theta \in (\alpha, \beta)$. En appliquant la formule *1.39* (1) nous obtenons

$$\varphi'(\theta) = y'(c)\, x'(\theta), \quad c = x(\theta)$$

d'où (1).

b. La relation (1) prend une forme particulièrement simple lorsque la courbe L est le segment joignant les points a et b, de sorte que l'on peut écrire $L = \{x \in X : x(t) = (1-t)\,a + tb,\ 0 \leqslant t \leqslant 1\}$. Dans ce cas $x'(t) = b - a$ et (1) se ramène à

$$y(b) - y(a) = y'(c)\,(b-a), \quad c \in L.$$

1.42. Si la fonction $y(x)$, $y: G \to Y$ n'est plus numérique, alors le théorème de Lagrange n'est plus applicable à la fonction $\varphi(t) = y[x(t)]: R_1 \to Y$. Par conséquent, pour une telle fonction, l'égalité *1.41* (1) n'est en général pas valable (voir l'exercice 7).

a. Tout de même *une estimation a lieu*:

$$|y(b) - y(a)| \leqslant \sup_{x \in L} \|y'(x)\| \cdot s(L), \tag{1}$$

où $s(L)$ est la longueur de la courbe L (U16.19), $|y(b) - y(a)|$ la norme du vecteur $y(b) - y(a) \in Y$, $\|y'(x)\|$ la norme de l'opérateur $y'(x)$ $(X \to Y)$.

Pour le démontrer, désignons $M = \sup\limits_{x \in L} \|y'(x)\|$. D'après la définition de la dérivée, quels que soient $\varepsilon > 0$ et $\tau \in [\alpha, \beta]$, on peut trouver $\delta = \delta(\varepsilon, \tau) > 0$ tel que l'inégalité

$$|\varphi(t) - \varphi(\tau)| \leqslant |\varphi'(\tau)| \cdot |t - \tau| + \varepsilon |t - \tau| \leqslant$$
$$\leqslant M |x'(\tau)|\,|t - \tau| + \varepsilon |t - \tau|$$

est valable du moment que $t \in \Delta(\tau) = \{t : |t - \tau| < \delta\}$.

En vertu du lemme sur le recouvrement fini *U3.97*, il existe un recouvrement de l'intervalle $[\alpha, \beta]$ par un nombre fini d'intervalles de la forme ci-indiquée. Nous les désignons par $\Delta(\tau_1)$, $\Delta(\tau_3)$, ..., $\Delta(\tau_{2n+1})$ en supposant que $\alpha = \tau_1 < \tau_3 < \ldots < \tau_{2n+1} = \beta$. Choisissons ensuite les points $\tau_2 < \tau_n < \ldots < \tau_{2n}$ de façon que chaque point τ_{2k} appartienne à l'intersection des intervalles $\Delta(\tau_{2k-1})$ et $\Delta(\tau_{2k+1})$ $(k = 1, \ldots, n)$. Nous aurons

$$|y(b) - y(a)| = |\varphi(\beta) - \varphi(\alpha)| \leqslant \sum_{i=1}^{2n} |\varphi(\tau_{i+1} - \varphi(\tau_i)| \leqslant$$
$$\leqslant M \sum_{i=1}^{2n} |x'(\tau_j)|\,|\tau_{i+1} - \tau_i| + \varepsilon(\beta - \alpha) \tag{2}$$

où j désigne celui des nombres i et $i + 1$ qui est impair. En passant dans (2) à la limite pour $\varepsilon \to 0$ nous obtenons l'inégalité cherchée (1).

b. Dans le cas où L est le segment joignant les points a et b, l'inégalité (1) a la forme bien simple:

$$|y(b)-y(a)| \leqslant \sup_{x\in L} \|y'(x)\| \, |b-a|. \tag{3}$$

c. Pour le même cas, on peut démontrer une inégalité plus forte:

$$|y(b)-y(a)| \leqslant \sup_{x\in L} |y'(x)(b-a)|, \tag{4}$$

dont l'avantage devant (3) se voit, par exemple, lorsque $X = R_n$, $Y = R_1$ et le vecteur $y'(x) = \operatorname{grad} y(x)$ fait avec le vecteur $b - a$ un angle proche de l'angle droit.

d. Dans les suppositions *b-c*, considérons la fonction

$$g(x) = y(x) - y'(a)(x - a).$$

Elle est dérivable avec $y(x)$, donc, d'après *1.32a-b*, on a

$$g'(x) = y'(x) - y'(a).$$

En appliquant à $g(x)$ l'inégalité (4), on trouve

$$|y(b)-y(a)-y'(a)(b-a)| \equiv |g(b)-g(a)| \leqslant \sup_{x\in L} |g'(x)(b-a)|,$$

ou bien

$$|y(b)-y(a)-y'(a)(b-a)| \leqslant \sup_{x\in L} |[y'(x)-y'(a)]\cdot(b-a)|, \tag{5}$$

ce qui est déjà beaucoup plus fort que (2).

e. L'inégalité (5) peut être utilisée sous une forme moins forte:

$$|y(b)-y(a)-y'(a)(b-a)| \leqslant \sup_{x\in L} \|y'(x)-y'(a)\| \, |b-a|, \tag{6}$$

ce qui reste tout de même beaucoup plus fort que (2).

1.43. a. Théorème. *Si, dans une boule* $V = \{x \in X : |x - a| \leqslant r\}$, *on a* $y'(x) \equiv 0$, *alors* $y(x)$ *est une constante.*

En effet, il suit de *1.42* (1) que, pour tout $b \in V$ et pour le segment L joignant les points a et b, on a l'inégalité *1.42* (3):

$$|y(b)-y(a)| \leqslant \sup_{x\in L} \|y'(x)\| \, |b-a| = 0,$$

d'où $y(b) \equiv y(a)$.

b. Conséquence. *Si, dans une boule* $V = \{x \in X : |x - a| \leqslant r\}$, *deux fonctions* $y_1(x)$ et $y_2(x)$ *ont les dérivées identiques, alors la différence de* $y_1(x)$ *et de* $y_2(x)$ *est constante dans cette boule.*

En effet, $y_1(x) - y_2(x)$ a la dérivée nulle, et on peut appliquer *a*.

c. Conséquence. *Si, dans une boule* $V = \{x \in X : |x - a| \leqslant r\}$, *l'opérateur* $y'(x)$ *est constant* ($y'(x) \equiv y'(a)$),

alors

$$y(x) = y'(a)(x - a) + y(a).$$

La démonstration se déroule de la même façon que dans *a*, sauf que l'on utilise l'inégalité *1.42* (5) au lieu de *1.42* (1). On peut se passer de *1.42* (5): la fonction $g(x) = y'(a)(x - a) + y(a)$ a la dérivée $g'(x) = y'(a)$; c'est pourquoi, conformément à *b*, la différence des fonctions $g(x)$ et $y(x)$ est constante; en posant $x = a$ nous voyons que cette constante est nulle.

1.44. Dérivée et condition de Lipschitz.

a. Par définition, une fonction $y = y(x)$ $(G \subset X \to Y)$ satisfait à la *condition de Lipschitz* dans une boule $V = \{x \in X : |x - a| \leqslant r\} \subset G$ s'il existe une constante $C > 0$ telle que l'inégalité

$$|y(x_1) - y(x_2)| \leqslant C|x_1 - x_2| \qquad (1)$$

a lieu pour tous les $x_1 \in V$, $x_2 \in V$.

Supposons que la fonction $y(x)$ soit dérivable dans la boule V et que

$$\sup_{x \in V} \|y'(x)\| = B.$$

Alors, en vertu de *1.42* (3), nous avons

$$|y(x_1) - y(x_2)| \leqslant B|x_1 - x_2|,$$

c'est-à-dire que la fonction $y(x)$ satisfait dans la boule V à la condition de Lipschitz (1) avec la constante B.

Inversement, si la fonction $y(x)$ a la dérivée continue $y'(x)$ et satisfait dans la boule V à la condition de Lipschitz (1), on peut affirmer que $\|y'(x)\| \leqslant C$. En effet, soit $|x - a| = \rho < r$; pour $\varepsilon > 0$ trouvons $\delta > 0$, $\delta < r - \rho$ de façon à avoir l'inégalité

$$|y(x_1) - y(x) - y'(x)(x_1 - x)| < \varepsilon |x_1 - x|$$

pour tout x_1 tel que $|x_1 - x| \leqslant \delta$. Puisque $|y(x_1) - y(x)| \leqslant C|x_1 - x|$ par hypothèse, nous avons

$$|y'(x)(x_1 - x)| \leqslant (C + \varepsilon)|x_1 - x|$$

pour $|x_1 - x| \leqslant \delta$, d'où l'inégalité

$$\|y'(x)\| \leqslant C + \varepsilon$$

pour la norme de l'opérateur $y'(x)$; $\varepsilon > 0$ étant arbitraire, il en résulte $\|y'(x)\| \leqslant C$, ce qui achève la preuve.

Remarquons que, en général, la condition de Lipschitz (1) n'est pas suffisante pour la dérivabilité de la fonction $y(x)$ (même dans le cas $X = R_1$; exemple: $y = |x|$ $(R_1 \to R_1)$).

b. Un cas spécial très intéressant se présente lorsque $\sup_{x \in V} \|y'(x)\| < 1$. Ceci veut dire, on l'a vu, que la fonction $y(x)$ satisfait à la condition de Lipschitz avec une constante C inférieure

à 1. En d'autres termes, *l'application* $y = y(x)$ *diminue les distances*: la distance entre les points $y(x)$ et $y(a)$ est strictement inférieure à la distance entre les points x et a. Si, de plus, les valeurs de la fonction $y(x)$ appartiennent au domaine V et que $\sup \| y'(x) \| \leqslant \theta < 1$, alors l'application $y(x)$: $V \to V$ se trouve *contractante* (*U13.22*):

$$| y(x') - y(x'') | \leqslant \theta | x' - x'' |.$$

Sur la boule fermée V (sur laquelle cette inégalité se prolonge par continuité), l'application $y(x)$ possède forcément un point fixe en vertu de *U13.22*.

c. On peut formuler une simple condition suffisante pour qu'une fonction $y(x)$ $(V \subset X \to X)$ qui satisfait, dans une boule $V = \{x \in X: | x - a | \leqslant r\}$, à la condition de Lipschitz avec une constante $\theta < 1$ applique cette boule dans elle-même. Notamment, la fonction doit vérifier l'inégalité $| y(a) - a | \leqslant (1 - \theta) r$. Dans cette condition, nous avons pour tout $x \in V$:

$$| y(x) - a | \leqslant | y(x) - y(a) | + | y(a) - a | \leqslant$$
$$\leqslant \theta | x - a | + | 1 - \theta | r \leqslant \theta r + (1 - \theta) r = r,$$

de sorte que toutes les valeurs de la fonction $y(x)$ pour $x \in V$ appartiennent à la boule V.

d. En combinant les résultats a, b et c nous aboutissons au théorème:

Théorème. *Si, dans une boule* $V = \{x \in X: | x - a | \leqslant r\}$, *une fonction* $y(x)$ $(V \to X)$ *est dérivable, et les inégalités*

$$\sup_{x \in V} \| y'(x) \| \leqslant \theta, \quad | y(a) - a | \leqslant (1 - \theta) r$$

sont remplies avec un $\theta < 1$, *alors il existe dans la boule* V *un, et un seul, point* x_0 *tel que* $y(x_0) = x_0$. Nous nous servirons dans la suite de cette méthode de démonstration d'existence des points fixes.

1.45. Théorème. *Supposons que, dans une boule* $V = \{x \in X: | x - a | \leqslant r\}$, *il y ait une suite* $y_1(x)$, $y_2(x)$, ... *de fonctions dérivables, à valeurs dans un espace complet* Y, *dont les dérivées* $y_1'(x)$, $y_2'(x)$, ... $(X \to L(X, Y))$ *soient continues et convergent uniformément dans* V *vers une fonction* $g(x)$ $(X \to L(X, Y))$. *Si les vecteurs* $y_1(a)$, $y_2(a)$, ... *de l'espace* Y *tendent vers une limite, alors la suite* $y_1(x)$, $y_2(x)$, ... *converge uniformément dans* V *vers une fonction* $y(x)$ *(à valeurs dans* Y*) dérivable à l'intérieur de la boule* V, *et l'on a* $y'(x) = g(x)$.

Démonstration. Ecrivons la formule *1.42* (3) en y remplaçant y par $y_n - y_m$ et b par x:

$$| [y_n(x) - y_m(x)] - [y_n(a) - y_m(a)] | \leqslant \sup_{x \in V} \| y_n'(x) - y_m'(x) \| | x - a |.$$

Vue la convergence uniforme des fonctions $y'_n(x)$, le second membre tend vers 0 lorsque $n \to \infty$, $m \to \infty$. Les vecteurs $y_n(a)$ étant convergents, nous voyons que la différence $y_n(x) - y_m(x)$ tend vers 0 pour $n \to \infty$, $m \to \infty$, uniformément par rapport à $x \in V$. Ainsi, la suite $y_n(x)$ est de Cauchy dans Y et, Y étant complet, il existe la fonction limite $y(x) = \lim_{n \to \infty} y_n(x)$. Elle est continue en vertu de la convergence uniforme de la suite $y_n(x)$ (*1.13b*). Ecrivons maintenant la formule *1.42* (6) pour la fonction $y_n(x)$ dans une boule de rayon ρ et de centre à un point b, $|b - a| \leqslant r - \rho$:

$$|y_n(x) - y_n(b) - y'_n(b)(x-b)| \leqslant \sup_{|\xi - b| \leqslant \rho} \|y'_n(\xi) - y'_n(b)\| \, |x - b|. \quad (1)$$

Pour $\varepsilon > 0$, on peut choisir $\delta > 0$, $\delta \leqslant \rho$ de façon à avoir

$$\sup_{|\xi - b| \leqslant \delta} \|y'_n(\xi) - y'_n(b)\| \leqslant \varepsilon$$

pour tous les $n = N, N+1, \ldots$ suffisamment grands. Afin de s'en convaincre il suffit de noter

$$y'_n(\xi) - y'_n(b) = [g(\xi) - g(b)] - [g(\xi) - y'_n(\xi)] + [g(b) - y'_n(b)]$$

et d'utiliser l'uniformité de la convergence de $y'_n(x)$ vers $g(x)$ et la continuité de la fonction $g(x)$.

En remplaçant dans (1) ρ par δ et passant à la limite pour $n \to \infty$, on voit que, pour tous les x, $|x - b| \leqslant \delta$:

$$|y(x) - y(b) - g(b)(x - b)| \leqslant \varepsilon |x - b|;$$

c'est ce qui démontre la dérivabilité de $y(x)$ pour $x = b$ et l'égalité $y'(b) = g(b)$.

Dans tous les résultats *1.44-1.45*, la boule V peut être remplacée par un domaine connexe (i.e. un domaine dont deux points quelconques x_0 et x_1 peuvent être joints par un polygone d'un nombre fini de côtés).

1.46. Dérivées par rapport à un sous-espace.

a. D'après la définition, si une fonction $y = y(x)$ $(G \subset X \to Y)$ est dérivable pour $x = c \in G$, alors l'opérateur linéaire $y'(c)$ est défini sur tout l'espace X, et la partie linéaire principale de l'accroissement de la fonction $y(x)$ correspondant à l'accroissement h de la variable indépendante vaut $y'(c)h$. Or, on peut poser la question sur la dérivabilité de la fonction $y(x)$ en se bornant aux accroissements de la variable indépendante qui appartiennent à un sous-espace vectoriel $X_1 \subset X$.

Nous dirons qu'une fonction $y(x)$ $(G \subset X \to Y)$ est *dérivable par rapport à un sous-espace* $X_1 \subset X$ si l'accroissement de $y(x)$ en passant d'un point $x = c$ à un point $x = c + h$, $h \in X_1$, admet

une partie linéaire principale par rapport à $h \in X_1$:

$$y(c+h) - y(c) = D_1(c)\,h + o(h),$$

où $D_1(c)$ est un opérateur linéaire défini sur le sous-espace X_1. Cet opérateur s'appelle *opérateur dérivée partielle par rapport au sous-espace* X_1. On désigne parfois la variable x appartenant au sous-espace X_1 par un symbole spécial, par exemple x_1 (tout en gardant la désignation x pour les vecteurs de l'espace X), alors on désigne l'opérateur $D_1(c)$ par $\frac{\partial y(c)}{\partial x_1}$.

b. Voyons, par exemple, qu'est-ce que la dérivabilité d'une fonction $y(x)$ $(G \subset R_n \to Y)$ par rapport au sous-espace unidimensionnel X_k défini par l'axe des coordonnées x_k. Dans ce cas, $h = (0, \ldots, h_k, \ldots, 0)$ et nous avons l'égalité

$$y(c+h) - y(c) = y(c_1, \ldots, c_k + h_k, \ldots, c_n) - y(c_1, \ldots, c_k, \ldots, c_n) = D_k(c)\,h_k + o(h_k),$$

$$D_k(c): X_k \to Y.$$

Or, cela est équivalent à l'existence de la dérivée partielle ordinaire $\frac{\partial y(c)}{\partial x_k}$ (*1.22b*) qui coïncide avec la quantité $D_k(c)$. En général, la dérivée par rapport à un sous-espace unidimensionnel n'est rien d'autre que la dérivée suivant la direction correspondante (*1.27a*).

c. La dérivabilité d'une fonction $y(x)$: $R_n \to R_m$ à un point $x = c$ par rapport au sous-espace R_k engendré par k premiers vecteurs d'une base dans R_n implique l'existence de toutes les dérivées partielles figurant dans la matrice:

$$\frac{\partial(y_1, \ldots, y_n)(c)}{\partial(x_1, \ldots, x_k)} \cong \left\| \begin{array}{ccc} \frac{\partial y_1(c)}{\partial x_1} & \cdots & \frac{\partial y_1(c)}{\partial x_k} \\ \cdots & \cdots & \cdots \\ \frac{\partial y_n(c)}{\partial x_1} & \cdots & \frac{\partial y_n(c)}{\partial x_k} \end{array} \right\|.$$

L'opérateur linéaire $(R_k \to R_m)$ correspondant à cette matrice représente la dérivée partielle de la fonction $y(x)$ par rapport au sous-espace R_k.

d. Il est clair que, si la fonction $y(x)$ est dérivable à un point c au sens original *1.23*, elle sera dérivable à ce point par rapport à n'importe quel sous-espace $X_1 \subset X$, et l'opérateur linéaire correspondant dérivée partielle $\frac{\partial y(c)}{\partial x_1}$ sera tout simplement la restriction de l'opérateur $\frac{\partial y(c)}{\partial x}$ au sous-espace X_1.

1.47. La dérivabilité d'une fonction $y(x)$ par rapport à un sous-espace (propre) $X_1 \subset X$ n'implique en général pas sa dérivabilité par rapport à tout le X. Même plus, dans le cas de $X = R_n$, la

dérivabilité d'une fonction $y(x)$ à un point $x = c$ par rapport à *tout* sous-espace de dimension $k < n$ n'implique pas sa dérivabilité par rapport à l'espace R_n (cf. l'exercice 3). Le théorème suivant est valable :

a. Théorème. *Si un espace X est la somme droite de deux sous-espaces fermés X_1 et X_2 et si une fonction $y(x)$: $G \subset X \to Y$ est dérivable dans un voisinage d'un point $c \in G$ par rapport aux sous-espaces X_1 et X_2, les dérivées $\frac{\partial y(x)}{\partial x_1}$ et $\frac{\partial y(x)}{\partial x_2}$ étant continues au point c, alors la fonction $y(x)$ est dérivable au point c par rapport à tout l'espace X.*

Démonstration. Pour tout $h \in X$ on peut écrire $h = h_1 + h_2$, $h_1 \in X_1$, $h_2 \in X_2$, et la relation $h \to 0$ est équivalente à $h_1 \to 0$, $h_2 \to 0$ (*U12.72e*). Ensuite, pour h assez petits, on a

$$y(c+h) - y(c) = [y(c+h_1+h_2) - y(c+h_1)] + \\ + [y(c+h_1) - y(c)],$$

et d'après le théorème de la moyenne *1.42e* on trouve :

$$\left| y(c+h) - y(c) - \frac{\partial y(c+h_1)}{\partial x_2} h_2 - \frac{\partial y(c)}{\partial x_1} h_1 \right| \leqslant \\ \leqslant \sup_{0 \leqslant t \leqslant 1} \left\| \frac{\partial y(c+h_1+th_2)}{\partial x_2} - \frac{\partial y(c+h_1)}{\partial x_2} \right\| |h_2| + \\ + \sup_{0 \leqslant \tau \leqslant 1} \left\| \frac{\partial y(c+\tau h_1)}{\partial x_1} - \frac{\partial y(c)}{\partial x_1} \right\| |h_1|.$$

En vertu de la continuité des dérivées $\frac{\partial y(x)}{\partial x_1}$ et $\frac{\partial y(x)}{\partial x_2}$ au point c, on a :

$$\frac{\partial y(c+h_1)}{\partial x_2} = \frac{\partial y(c)}{\partial x_2} + o(1),$$

$$\frac{\partial y(c+h_1+th_2)}{\partial x_2} - \frac{\partial y(c+h_1)}{\partial x_2} = o(1),$$

$$\frac{\partial y(c+\tau h_1)}{\partial x_1} - \frac{\partial y(c)}{\partial x_1} = o(1),$$

où $o(1)$ tend vers 0 lorsque $h_1 \to 0$, $h_2 \to 0$. Par conséquent,

$$y(c+h) - y(c) = \frac{\partial y(c)}{\partial x_1} h_1 + \frac{\partial y(c)}{\partial x_2} h_2 + o(h).$$

On peut mettre ce résultat sous la forme

$$y(c+h) - y(c) = Dh + o(h),$$

où l'égalité $Dh = \frac{\partial y(c)}{\partial x_1} h_1 + \frac{\partial y(c)}{\partial x_2} h_2$ définit D en tant qu'opérateur continu sur l'espace X (*U12.72f*). Ceci démontre le théorème.

b. Le théorème suivant plus général se déduit facilement de *a* par récurrence.

Théorème. *Si un espace X est la somme droite de ses sous-espaces fermés $X_1, \ldots, X_n$ et si une fonction $y(x)$ $(G \subset X \to Y)$ est dérivable dans un voisinage d'un point $x = c$ par rapport aux sous-espaces $X_1, \ldots, X_n$, les dérivées $\frac{\partial y(x)}{\partial x_1}, \ldots, \frac{\partial y(x)}{\partial x_n}$ étant continues au point c, alors la fonction $y(x)$ est dérivable au point c par rapport à l'espace X.*

c. Appliquons le théorème *b* au cas où $X = R_n$ et $X_1, \ldots, X_n$ sont les sous-espaces unidimensionnels répondant aux axes des coordonnées. La dérivée par rapport à chaque sous-espace unidimensionnel X_k étant la dérivée partielle $\frac{\partial}{\partial x_k}$ (*1.46b*), le théorème *b* nous amène au résultat suivant:

Théorème. *Si une fonction $y(x)$ $(G \subset R_n \to Y)$ possède, dans un voisinage d'un point $x = c$, les dérivées partielles $\frac{\partial y(x)}{\partial x_1}, \ldots, \frac{\partial y(x)}{\partial x_n}$ qui sont continues pour $x = c$, alors la fonction $y(x)$ est dérivable au point c.*

Ce théorème fournit les conditions suffisantes de la dérivabilité d'une fonction $y(x): G \subset R_n \to Y$, puisqu'il ne demande que l'existence des dérivées partielles (par rapport à toutes les variables) continues à un point donné; de telles conditions sont souvent plus faciles à vérifier.

D'autre part, ce théorème peut être mis sous forme d'une condition nécessaire et suffisante: *pour que la dérivée $y'(x)$ d'une fonction $y(x): G \subset R_n \to Y$ existe et soit continue dans le domaine G, il faut et il suffit qu'existent et soient continues dans G ses dérivées partielles $\frac{\partial y(x)}{\partial x_1}, \ldots, \frac{\partial y(x)}{\partial x_n}$.*

d. Signalons une condition qui permet de conclure sur la dérivabilité d'une fonction $y(x)$ à partir de l'existence de ses dérivées suivant les directions:

Théorème. *Supposons que, dans un domaine $G \subset X$, on ait une fonction vectorielle $y(x)$ $(G \to Y)$ et une fonction opératorielle continue $D(x)$ $(G \to L(X, Y))$. Supposons, de plus, que la fonction $y(x)$ ait, en tout point $c \in G$, la dérivée suivant toute direction $\Gamma = \{x = c + th,\ 0 \leqslant t < \infty\}$, cette dérivée agissant sur tout vecteur $h \in X$ comme opérateur $D(c)$:*

$$y'_\Gamma(c)\, h = D(c)\, h.$$

Alors la fonction $y(x)$ est dérivable dans le domaine G et $y'(x) = D(x)$.

Démonstration. Pour $h \in X$ donné, nous avons au point $c \in G$:

$$y(c + h) - y(c) = y'_\Gamma(c)\, h + o(h) = D(c)\, h + o(h),$$

et il reste à démontrer que la quantité $o(h)$ est infiniment petite, uniformément par rapport à tous les h, quelles que soient leurs directions. Sur toute demi-droite Γ la fonction $y(x)$ est dérivable et l'on peut appliquer l'estimation *1.42d*:

$$|y(c+h)-y(c)-D(c)h| \leqslant \sup_{0\leqslant\theta\leqslant 1} |[y'(c+\theta h)-y'(c)]h| =$$
$$= \sup_{0\leqslant\theta\leqslant 1} |[D(c+\theta h)-D(c)]h| \leqslant$$
$$\leqslant \sup_{0\leqslant\theta\leqslant 1} \|D(c+\theta h)-D(c)\| \, |h|. \qquad (1)$$

A présent, $\varepsilon > 0$ étant donné, trouvons $\delta > 0$ de façon à avoir $\|D(c+k)-D(c)\| < \varepsilon$ dès que $|k| < \delta$; c'est possible par hypothèse de continuité de la fonction opératorielle $D(x)$ au point c. Alors, quel que soit h, $|h| < \delta$, nous tirons de (1)

$$|y(c+h)-y(c)-D(c)h| \leqslant \varepsilon |h|,$$

où ε ne dépend plus de la direction de h. Nous avons déjà vu qu'il en résulte la dérivabilité de la fonction $y(x)$ pour $x = c$, de même que la relation $y'(c) = D(c)$, ce qu'il fallait démontrer.

1.48. Le théorème de la moyenne permet parfois d'établir la dérivabilité de fonctions compliquées à partir de la dérivabilité de fonctions plus simples.

a. Soient M un espace métrique, V un domaine dans un espace normé Y et $\Phi(x, y)$ une fonction bornée et uniformément continue dans $M \times V$, à valeurs dans un espace normé Z. Considérons les espaces métriques $Y(M)$ et $Z(M)$ (*1.13c*) composés des fonctions bornées et continues en $x \in M$, à valeurs dans Y et Z respectivement; leurs métriques indiquées dans *1.13c* peuvent être données moyennant les normes respectives $\|y(x)\| = \sup_{x\in M} |y(x)|_Y$, $\|z(x)\| = \sup_{x\in M} |z(x)|_Z$. Naturellement, c'est par $V(M)$ que nous désignons l'ensemble des $y(x) \in Y(M)$ dont les valeurs appartiennent à V; il est clair que $V(M)$ représente une partie de l'espace $Y(M)$. D'après *1.13e*, la fonction $\Phi(x, y)$ détermine une application continue $Fy: V(M) \to Z(M)$ selon la formule $Fy(x) = \Phi(x, y(x))$. Supposons encore que la fonction $\Phi(x, y)$ ait la dérivée $\frac{\partial \Phi(x, y)}{\partial y}$ bornée et uniformément continue dans $M \times V$. Montrons qu'alors la fonction $F(y)$ est elle aussi dérivable dans $V(M)$ et trouvons l'expression de sa dérivée. Pour tout x fixe et pour $y(x) \in V(M)$ et $y(x) + h(x) \in V(M)$ donnés, nous avons

$$\Phi(x, y(x)+h(x))-\Phi(x, y(x)) = \frac{\partial \Phi(x, y(x))}{\partial y} h(x) + R, \qquad (1)$$

où, d'après *1.24e*,

$$|R| \leqslant \sup_{0 \leqslant \theta(x) \leqslant 1} \left| \frac{\partial \Phi(x, y(x) + \theta(x) h(x))}{\partial y} - \frac{\partial \Phi(x, y(x))}{\partial y} \right| \times$$

$$\times |h(x)| \leqslant \sup_{\substack{0 \leqslant \theta(x) \leqslant 1 \\ x \in M}} \left| \frac{\partial \Phi(x, y(x) + \theta(x) h(x))}{\partial y} - \frac{\partial \Phi(x, y(x))}{\partial y} \right| \cdot \sup_{x \in M} |h(x)|. \quad (2)$$

Le premier membre de (1) est une fonction bornée et continue de x. Le premier terme du second membre l'est également (*1.15c* et *U12.61*). Par conséquent, le second terme du second membre est encore une fonction bornée et continue de x. On peut donc considérer l'égalité (1) comme égalité dans l'espace $Z(M)$. Le facteur $\frac{\partial \Phi(x, y(x))}{\partial y}$, pour tout x fixe, est un opérateur linéaire (borné uniformément par rapport à x) de Y dans Z; pour cette raison, la fonction $\frac{\partial \Phi(x, y(x))}{\partial x}$ peut être traitée en tant qu'opérateur linéaire borné $\left(\text{dont la norme ne dépasse pas } \sup_{x \in M} \left\| \frac{\partial \Phi(x, y)}{\partial y} \right\|\right)$ agissant de $Y(M)$ dans $Z(M)$. Le premier terme à droite dans (1) est donc linéaire en $h \in Y(M)$. Le second terme a dans $Z(M)$ la norme d'ordre $o(\|h\|)$, ce qui résulte de l'estimation (2) et de la continuité uniforme de $\frac{\partial \Phi(x, y)}{\partial y}$. Il s'ensuit que *l'application $F(y)$ est dérivable dans le domaine $V(M)$ et sa différentielle coïncide avec le premier terme à droite dans* (1). On peut aussi écrire

$$F'(y) = \frac{\partial \Phi(x, y(x))}{\partial y} \quad (3)$$

en considérant le second membre comme opérateur linéaire cidécrit de $Y(M)$ dans $Z(M)$.

b. Choisissons en tant que M dans a l'intervalle $a \leqslant x \leqslant b$ de la droite réelle. Alors on peut définir sur l'espace $Z(M)$ l'opérateur d'intégration (*U12.62*):

$$(Iz)(x) = \int_a^x z(\xi)\, d\xi$$

qui est linéaire et borné (on le sait de *U12.62c*), avec la norme ne dépassant pas $b - a$. En composant cet opérateur avec l'application $F(y)$ de a nous obtenons une nouvelle application

$$[IF(y)](x) = \int_a^x \Phi(\xi, y(\xi))\, d\xi \quad (V(M) \to Z(M)). \quad (4)$$

De même que F, l'application IF est continue (*1.15b*), dérivable (*1.32b*), et sa dérivée, conformément à *1.32b*, vaut

$$(IF)'\, y = IF'\,(y) = \int_a^x \frac{\partial \Phi\,(\xi,\, y\,(\xi))}{\partial y}\, d\xi. \tag{5}$$

c. Dans les raisonnements ci-dessus, on peut remplacer l'intégrale à la limite supérieure variable par une intégrale aux limites d'intégration constantes, par exemple a et b. L'opérateur I sera alors un opérateur linéaire de $Z\,(M)$ dans Z, et l'opérateur IF agira de $V\,(M)$ toujours dans Z. Il sera continu et dérivable (sur $V\,(M)$) avec pour dérivée

$$(IF)'\, y = \int_a^b \frac{\partial \Phi\,(\xi,\, y\,(\xi))}{\partial y}\, d\xi. \tag{6}$$

d. Il est possible de généraliser les constructions données en permettant la dépendance de la fonction $\Phi\,(x,\, y)$ d'un paramètre λ qui parcourt un espace métrique Δ ; nous noterons $\Phi\,(x,\, y,\, \lambda)$ à la place de $\Phi\,(x,\, y)$. Dans ce cas, la fonction $Ff\,(x) = \Phi\,(x,\, f\,(x),\, \lambda)$ dépend elle aussi de λ. Nous affirmons : si la fonction $\Phi\,(x,\, y,\, \lambda)$ est continue en λ, uniformément par rapport à $x \in M$ et $y \in V$, alors la fonction $\Phi\,(x,\, f\,(x),\, \lambda)$ est continue en λ sur la métrique de l'espace $Z\,(M)$. En effet, pour $\varepsilon > 0$ donné, il existe par hypothèse un $\delta > 0$ tel que l'inégalité $|\,\Phi\,(x,\, y,\, \lambda') - \Phi\,(x,\, y,\, \lambda)\,| \leqslant \varepsilon$ a lieu pour tous les $x \in M$, $y \in V$, $\rho\,(\lambda',\, \lambda) < \delta$. Il en découle pour $\rho\,(\lambda',\, \lambda) < \delta$:

$$\|\,\Phi\,(x,\, f\,(x),\, \lambda') - \Phi\,(x,\, f\,(x),\, \lambda)\,\|_{Z(M)} = \\ = \sup_{x \in M} |\,\Phi\,(x,\, f\,(x),\, \lambda') - \Phi\,(x,\, f\,(x),\, \lambda)\,| \leqslant \varepsilon,$$

ce qui exprime la continuité de la fonction $\Phi\,(x,\, f\,(x),\, \lambda)$ par rapport au paramètre λ dans l'espace $Z\,(M)$.

Ensuite, supposons que l'espace métrique Δ soit un domaine dans un espace normé Λ et que la fonction $\Phi\,(x,\, y,\, \lambda)$ ait la dérivée partielle $\dfrac{\partial \Phi}{\partial \lambda}$ bornée et uniformément continue dans $M \times V \times \Delta$. Nous affirmons : la fonction $\Phi\,(x,\, f\,(x),\, \lambda)$ en tant qu'élément de l'espace $Z\,(M)$ possède alors la dérivée en λ qui a la forme $\dfrac{\partial \Phi\,(x,\, f\,(x),\, \lambda)}{\partial \lambda}$. En effet, conformément à *1.42* (5) et en vertu de notre hypothèse, il existe, pour $\varepsilon > 0$ donné, un $\delta > 0$ tel que, une fois $x \in M$, $y \in V$, $|\,\lambda' - \lambda\,| < \delta$, $\Delta\lambda = \lambda' - \lambda$, on a

$$\left|\Phi\,(x,\, y,\, \lambda') - \Phi\,(x,\, y,\, \lambda) - \frac{\partial \Phi\,(x,\, y,\, \lambda)}{\partial \lambda}\,\Delta\lambda\right| \leqslant \\ \leqslant \sup_{0 \leqslant \theta \leqslant 1} \left|\frac{\partial \Phi\,(x,\, y,\, \lambda + \theta\Delta\lambda)}{\partial \lambda} - \frac{\partial \Phi\,(x,\, y,\, \lambda)}{\partial \lambda}\right| |\,\Delta\lambda\,| \leqslant \varepsilon\,|\,\Delta\lambda\,|.$$

On en tire

$$\left\| \Phi(x, f(x), \lambda') - \Phi(x, f(x), \lambda) - \frac{\partial \Phi(x, f(x), \lambda)}{\partial \lambda} \Delta\lambda \right\|_{Z(M)} =$$

$$= \sup_{x \in M} \left| \Phi(x, f(x), \lambda') - \Phi(x, f(x), \lambda) - \right.$$

$$\left. - \frac{\partial \Phi(x, f(x), \lambda)}{\partial \lambda} \Delta\lambda \right| \leqslant \varepsilon |\Delta\lambda|,$$

donc $\frac{\partial \Phi(x, f(x), \lambda)}{\partial \lambda}$ est la dérivée par rapport à λ de $\Phi(x, f(x), \lambda)$ dans l'espace $Z(M)$, ce qu'il nous fallait.

Enfin, si $M = [a, b]$ alors, d'après b et c, les mêmes propriétés restent valables en appliquant à $\Phi(x, f(x), \lambda)$ l'opération d'intégration par rapport à x.

§ 1. 5. Théorème sur la fonction implicite

1.51. Fonction inverse; position du problème. Soit $x = \varphi(y)$: $F \to E$ une fonction qui applique un ensemble F d'une façon biunivoque sur un ensemble E; alors, sur l'ensemble E est définie d'une façon naturelle sa fonction *inverse* $y = f(x)$: $E \to F$ telle que $f[\varphi(y)] = y$, $\varphi[f(x)] = x$. La question suivante surgit souvent dans l'analyse: quelles sont les conditions à imposer à la fonction $x = \varphi(y)$ pour qu'il existe sa fonction inverse? On peut préciser le problème comme suit: on suppose que E et F soient des domaines dans des espaces normés et que la fonction $x = \varphi(y)$ soit continue et dérivable dans un voisinage d'un point $b \in F$; si $\varphi(b) = a$, il s'agit d'indiquer au moins un voisinage du point a dans lequel la fonction inverse $y = f(x)$ soit définie d'une façon univoque, continue et dérivable. Tout cela exige, naturellement, certaines conditions supplémentaires pour la fonction $\varphi(x)$. La nature de ces conditions sera mise à jour en considérant, en tant qu'exemple, la fonction linéaire $x = \varphi(y)$ d'un espace m-dimensionnel $F = R_m$ dans un espace n-dimensionnel $E = R_n$. Les bases dans ces espaces étant fixées d'une façon quelconque, nous mettons l'égalité $x = \varphi(y)$ sous la forme d'un système d'équations linéaires à coefficients numériques:

$$\left.\begin{aligned} x_1 &= \varphi_{11} y_1 + \ldots + \varphi_{1m} y_m, \\ &\ldots\ldots\ldots\ldots \\ x_n &= \varphi_{n1} y_1 + \ldots + \varphi_{nm} y_m. \end{aligned}\right\} \qquad (1)$$

On se demande, dans quelles conditions ce système a-t-il une unique solution $y_1, \ldots, y_m$ pour n'importe quels $x_1, \ldots, x_n$ (choisis

arbitrairement dans un voisinage au moins du point $(0, \ldots, 0) \in \in R_n$)? L'algèbre nous donne la réponse: *la* $n \times m$*-matrice* $\Phi = \| \varphi_{ij} \|$ $(i = 1, \ldots, n;\ j = 1, \ldots, m)$ *doit être inversible* (en particulier, la condition $m = n$ est nécessaire), alors la solution cherchée s'obtient d'après les formules de Kramer. Notons ensuite que la matrice Φ est celle d'un opérateur, à savoir de la dérivée de la fonction $\varphi(y)$ (*1.25b*). Donc, en cas de fonction linéaire $x = \varphi(y)$ $(R_m \to R_n)$ *c'est l'inversibilité de l'opérateur* $\varphi'(y)$ *qui sert de condition nécessaire et suffisante pour l'existence de la fonction inverse* (l'opérateur $\varphi'(y)$ est dans ce cas constant, et il n'est pas nécessaire d'indiquer le point b). Dans le cas général d'une fonction dérivable $\varphi(y) : F \to E$, il est naturel d'imposer la condition d'inversibilité de $\varphi'(b)$; si la partie linéaire principale d'une application est inversible, l'inversibilité de toute l'application semble bien probable.

1.52. Fonction implicite; position du problème. Nous verrons plus bas que l'hypothèse exprimée dans *1.51* est juste: si une fonction $\varphi(y)$ est dérivable et l'opérateur linéaire $\varphi'(b)$ représentant la dérivée de φ en un point b est inversible, alors, dans un voisinage du point $a = \varphi(b)$, il existe effectivement l'unique fonction continue et dérivable $f(x)$ telle que $f[\varphi(y)] = y$, $\varphi[f(x)] = x$ (*1.55*). Or, il s'avère qu'a lieu un théorème plus général où il s'agit de l'existence de solution non pas de l'équation spéciale $x - \varphi(y) = 0$, mais d'une équation beaucoup plus générale $\Phi(x, y) = 0$.

Dans le cas linéaire, cette équation générale prend la forme du système:

$$\left.\begin{aligned} a_{11}x_1 + \ldots + a_{1n}x_n + b_{11}y_1 + \ldots + b_{1m}y_m = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{p1}x_1 + \ldots + a_{pn}x_n + b_{p1}y_1 + \ldots + b_{pm}y_m = 0. \end{aligned}\right\} \qquad (1)$$

C'est ainsi que se pose la question: dans quelles conditions imposées aux coefficients a_{ij}, b_{ij} le système (1) détermine-t-il de façon unique les nombres $y_1, \ldots, y_m$, quels que soient les nombres donnés $x_1, \ldots, x_n$? La réponse est toujours formulée dans l'algèbre: il faut que la matrice $\| b_{ij} \|$ des coefficients de $y_1, \ldots, y_m$ soit inversible (en particulier, la condition $m = p$ est nécessaire). Cette réponse peut être formulée en termes de la fonction $z = \Phi(x, y)$ appliquant la somme droite des sous-espaces R_n et R_m dans l'espace R_p; la fonction mentionnée s'exprime en coordonnées comme suit:

$$\begin{aligned} z_1 &= a_{11}x_1 + \ldots + a_{1n}x_n + b_{11}y_1 + \ldots + b_{1m}y_m, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ z_p &= a_{p1}x_1 + \ldots + a_{pn}x_n + b_{p1}y_1 + \ldots + b_{pm}y_m. \end{aligned}$$

C'est bien la matrice $\| b_{ij} \|$ qui est la matrice de l'opérateur dérivée partielle $\frac{\partial z}{\partial y}$, et l'inversibilité de cette matrice signifie l'inversibilité de l'opérateur.

1.53. A présent, il est naturel d'énoncer le théorème suivant:

Théorème sur la fonction implicite. *Soient M un espace métrique et Y un espace normé complet; soit $z = \Phi(x, y)$ une fonction définie sur le produit de M par une boule $V_r = \{y \in Y : | y - b | \leqslant r\}$, à valeurs dans un espace normé Z. Supposons que cette fonction soit bornée, uniformément continue et possède la dérivée bornée et uniformément continue $\frac{\partial \Phi(x, y)}{\partial y}$. Supposons encore que, pour un $a \in M$, on ait $\Phi(a, b) = 0$ et que l'opérateur $\frac{\partial \Phi(a, b)}{\partial y}$ $(Y \to Z)$ soit inversible. Alors il existe une boule $U_\delta = \{x \in M : \rho(x, a) \leqslant \delta\}$ et une fonction $y = f(x) : U_\delta \to V$ définie et continue dans la boule U_δ et telle que $f(a) = b$ et $\Phi(x, f(x)) \equiv 0$ pour tous les $x \in U_\delta$. Cette fonction est unique au sens suivant: quelle que soit une autre fonction $f_1(x)$, à valeurs dans Y, définie et continue au voisinage du point $a \in M$ et telle que $f_1(a) = b$ et que $\Phi(x, f_1(x)) \equiv 0$, il existe une boule $\{x \in M : \rho(x, a) \leqslant \delta'\}$ dans laquelle $f_1(x) \equiv f(x)$.*

La fonction $y = f(x)$ s'appelle *fonction implicite définie par l'équation $\Phi(x, y) = 0$ et par la condition $f(a) = b$.*

Démonstration. Si $y = f(x)$ est la fonction cherchée, c'est-à-dire que $\Phi(x, f(x)) \equiv 0$, alors on a

$$\left[\frac{\partial \Phi(a, b)}{\partial y}\right]^{-1}\left[\frac{\partial \Phi(a, b)}{\partial y} f(x) - \Phi(x, f(x))\right] = f(x), \qquad (1)$$

donc on peut chercher la fonction $f(x)$ dans l'espace des fonctions $y(x) : U_\delta \subset M \to Y$ en tant que point fixe de la transformation $F(y)$ définie par la formule

$$Fy(x) = \left[\frac{\partial \Phi(a, b)}{\partial y}\right]^{-1}\left[\frac{\partial \Phi(a, b)}{\partial y} y(x) - \Phi(x, y(x))\right]. \qquad (2)$$

Afin d'utiliser cette idée, choisissons un $\delta > 0$ et considérons l'espace normé $Y(U_\delta)$ de toutes les fonctions bornées et continues $y(x)$, à valeurs dans l'espace Y (cf. *1.48a*), avec la norme $\| y(x) \| = \sup_{x \in U_\delta} | y(x) |_Y$. L'espace $Y(U_\delta)$ est complet (*1.13c*). Désignons par $V(U_\delta)$ l'ensemble des fonctions $y(x) \in Y(U_\delta)$ dont les valeurs, pour $x \in U_\delta$, appartiennent à la boule $V \subset Y$; l'ensemble $V(U_\delta)$ est la boule fermée dans l'espace $Y(U_\delta)$, de rayon r et de centre au point $b(x) \equiv b$, donc représente lui-même un espace métrique complet. L'application (2) est définie pour tous les $y(x) \in V(U_\delta)$ et prend ses valeurs dans l'espace $Y(U_\delta)$. En vertu de *1.13e*, *1.48a* et des propriétés de la fonction $\Phi(x, y)$, l'application F est continue

et dérivable dans la boule $V(U_\delta)$. D'après *1.48a* et *1.31-1.32*, sa dérivée a la forme:

$$F'(y) = E + \left[\frac{\partial \Phi(a, b)}{\partial y}\right]^{-1} \frac{\partial \Phi(x, y(x))}{\partial y} =$$

$$= E + \left[\frac{\partial \Phi(a, b)}{\partial y}\right]^{-1} \left[\frac{\partial \Phi(x, y(x))}{\partial y} - \frac{\partial \Phi(a, b)}{\partial y}\right] +$$

$$+ \left[\frac{\partial \Phi(a, b)}{\partial y}\right]^{-1} \frac{\partial \Phi(a, b)}{\partial y} =$$

$$= \left[\frac{\partial \Phi(a, b)}{\partial y}\right]^{-1} \left[\frac{\partial \Phi(x, y(x))}{\partial y} - \frac{\partial \Phi(a, b)}{\partial y}\right].$$

Pour la norme de $F'(y)$, on a l'estimation suivante:

$$\|F'(y)\| \leqslant \left\|\left[\frac{\partial \Phi(a, b)}{\partial y}\right]^{-1}\right\| \cdot \sup_{x \in U_\delta} \left\|\frac{\partial \Phi(x, y(x))}{\partial y} - \frac{\partial \Phi(a, b)}{\partial y}\right\|. \quad (3)$$

En utilisant la continuité de la fonction $\frac{\partial \Phi(x, y)}{\partial y}$ pour $x = a$, $y = b$ on peut trouver δ_1 et ρ tels que, du moment que $|x - a| \leqslant \delta_1$, $|y - b| \leqslant \rho$, le second membre dans (3) est inférieur à $\frac{1}{2}$. Ainsi nous avons $\|F'(y)\| < \frac{1}{2}$ dans $V_\rho(U_{\delta_1})$.

Remarquons ensuite que, pour $b(x) \equiv b$,

$$F[b(x)] - b(x) = -\left[\frac{\partial \Phi(a, b)}{\partial y}\right]^{-1} \Phi(x, b),$$

donc,

$$\|F[b(x)] - b(x)\| \leqslant \left\|\left[\frac{\partial \Phi(a, b)}{\partial y}\right]^{-1}\right\| \cdot \sup_{x \in U_\delta} |\Phi(x, b)|.$$

Comme $\Phi(a, b) = 0$ et la fonction $\Phi(x, y)$ est continue au point $x = a$, $y = b$, avec ρ déjà choisi, on peut trouver δ_2 tel que

$$\|F[b(x)] - b(x)\|_{Y(U_{\delta_2})} \leqslant \frac{1}{2}\rho.$$

Posons $\delta = \min(\delta_1, \delta_2)$; alors, pour l'application $F(y)$ considérée dans la boule $V_\rho(U_\delta)$, est remplie l'hypothèse du théorème *1.45d* (avec $\theta = \frac{1}{2}$). En appliquant ce théorème nous prouvons l'existence dans la boule $V_\rho(U_\delta)$ d'un point fixe de l'application $F(y)$. Désignons ce point par $f(x)$; il vérifie l'égalité (1), donc aussi l'égalité $\Phi(x, f(x)) \equiv 0$, pour $x \in U_\delta$. Montrons que $f(a) = b$. Remarquons que $y(a) = b$ implique $Fy(a) = b$. Rappelons maintenant que le point fixe d'une application contractante est construit comme limite des itérations y_0, Fy_0, F^2y_0, . . ., (*U13.22*) où y_0 est un point quelconque de l'espace métrique complet dans lequel

agit l'application contractante. Choisissons pour point initial du procédé itératif une quelconque fonction $y(x) \in V_\rho(U_\delta)$ avec $y(a) = b$, par exemple $y(x) \equiv b$; alors toutes les itérations $y(x)$ possèdent la propriété $y(a) = b$; il en est de même de la fonction limite $f(x)$ qui s'avère ainsi un point fixe cherché de l'application $F(y)$, ce qu'il nous fallait.

C'est l'unicité de la solution trouvée qu'il reste à démontrer. Notons que l'identité $F(f(x)) = f(x)$ obtenue pour les points x de la boule U_δ reste valable pour toute boule $U_{\delta'}$ avec $\delta' < \delta$. Par conséquent, la restriction de la fonction $f(x)$ sur cette plus petite boule représente un point fixe de l'application $F(y)$ dans la boule $V_\rho(U_{\delta'})$ aussi. Soit $f_1(x)$ une autre solution du problème sur la fonction implicite; de toute façon, il existe un $\delta' \leqslant \delta$ tel que la fonction $f_1(x)$ soit définie, continue et vérifie la condition $|f_1(x) - b| \leqslant \rho$ dans la boule $U_{\delta'}$, donc appartienne à la boule $V_\rho(U_{\delta'})$. Or, il n'existe qu'un seul point fixe de l'application $F(y)$ dans la boule $V_\rho(U_{\delta'})$. Nous voyons donc que $f_1(x) \equiv f(x)$ pour $x \in U_{\delta'}$. Le théorème est démontré.

1.54.a. Le théorème *1.53* sur la fonction implicite a un caractère local, c'est-à-dire que l'existence de la fonction $y(x)$ représentant une solution de l'équation $\Phi(x, y) = 0$ n'est garantie que dans un voisinage du point a, la valeur $b = y(a)$ étant donnée. Nous voudrions préciser le domaine d'existence de la fonction $y(x)$. Supposons, par exemple, que l'on sache que la fonction $\Phi(x, y)$ soit définie, continue et dérivable en y *pour tous les* $x \in M$, $y \in Y$, et que $\frac{\partial \Phi(x, y)}{\partial y}$ soit elle aussi *partout* définie, continue et inversible; on se demande, est-ce que, du moment que $\Phi(a, b) = 0$, la fonction implicite correspondante $y(x)$ (dont l'existence n'est garantie par le théorème sur la fonction implicite que dans une boule $|x - a| < \delta$) sera définie, sinon pour tous les $x \in M$, au moins pour une boule $|x - a| < r$ où r est un nombre positif *fixe* qui ne dépend pas de la fonction $\Phi(x, y)$?

Il se trouve que, même dans une telle situation qui semble très avantageuse, on est forcé de répondre négativement. A savoir nous allons indiquer, pour tout $\varepsilon > 0$, une telle fonction $\Phi_\varepsilon(x, y)$ $(R_2 \to R_1)$ qui sera définie, continue et dérivable en y pour tous les $\{x, y\} \in R_2$ et dont la dérivée par rapport à y sera partout continue et inversible; de plus, on aura $\Phi_\varepsilon(0, 0) = 0$. Tout de même, l'intervalle $-h < x < h$ ne sera l'intervalle d'existence de la fonction implicite correspondante que lorsque $h < \varepsilon$. Notamment, posons $\Phi_\varepsilon(x, y) = x + \varepsilon - \varepsilon e^y$; toutes les conditions que l'on vient d'imposer à la fonction $\Phi_\varepsilon(x, y)$ sont, évidemment, remplies, la fonction implicite correspondante $y = \ln \frac{x+\varepsilon}{\varepsilon}$ est définie pour seuls $x > -\varepsilon$.

b. En revanche à ce qui était dit dans *a* relativement au domaine d'*existence* de la fonction implicite, signalons la situation beaucoup plus satisfaisante en ce qui concerne l'*unicité* de la fonction implicite.

Supposons que l'espace métrique M soit *connexe*, i.e. qu'il n'admette aucun sous-ensemble (propre) ouvert et fermé à la fois. Soient $y = f(x)$ et $y = f_1(x)$, $f(a) = f_1(a)$, deux fonctions définies et continues sur M dont chacune satisfait à l'équation $\Phi(x, y) \equiv 0$. Alors, *si l'hypothèse du théorème sur la fonction implicite est remplie en tout point* $\{x, f(x)\} \in M \times Y$, *on a* $f(x) = f_1(x)$ *partout sur* M.

En effet, soit $B = \{x \in M : f(x) = f_1(x)\}$. L'ensemble B est fermé en tant qu'ensemble des racines de la fonction continue $f_1(x) - f(x)$ (*U5.14b*) ; or, il est ouvert, puisque, d'après le théorème sur la fonction implicite, contient un voisinage de tout son point. Il contient le point a, donc n'est pas vide ; vu que l'espace M est connexe, nous concluons que $B = M$, ce qu'il fallait.

1.55. Théorème sur la dérivée d'une fonction implicite. Dans ce qui suit, nous supposons que l'espace métrique M dont il s'agissait dans *1.53-1.54* soit un domaine dans un espace normé X.

a. Théorème. *L'hypothèse du théorème 1.53 étant remplie et, de plus, la fonction* $\Phi(x, y)$ *étant dérivable au point* (a, b) *(par rapport à l'espace* $X \times Y$*), la fonction implicite* $y = f(x)$ *que l'on a construite dans 1.53 est dérivable pour* $x = a$, *et*

$$f'(a) = -\left[\frac{\partial \Phi(a, b)}{\partial y}\right]^{-1} \frac{\partial \Phi(a, b)}{\partial x}. \tag{1}$$

Démonstration. Comme la fonction $\Phi(x, y)$ est dérivable pour $x = a$, $y = b$, nous avons pour Δx suffisamment petits :

$$0 = \Phi(a + \Delta x, f(a + \Delta x)) = \frac{\partial \Phi(a, b)}{\partial x} \Delta x + \\ + \frac{\partial \Phi(a, b)}{\partial y} \Delta y + o(|\Delta x| + |\Delta y|),$$

où $\Delta y = f(a + \Delta x) - f(a)$. Donc

$$\left|\left[\frac{\partial \Phi(a, b)}{\partial y}\right]^{-1} \frac{\partial \Phi(a, b)}{\partial x} \Delta x + \Delta y\right| \leqslant \\ \leqslant \left\|\left[\frac{\partial \Phi(a, b)}{\partial y}\right]^{-1}\right\| |o(|\Delta x| + |\Delta y|)|. \tag{2}$$

Supposons que Δx de même que Δy soient si petits que

$$\left\|\left[\frac{\partial \Phi(a, b)}{\partial y}\right]^{-1}\right\| |o(|\Delta x| + |\Delta y|)| \leqslant \frac{1}{2}(|\Delta x| + |\Delta y|).$$

Nous avons alors :

$$|\Delta y| - \left\| \left[\frac{\partial \Phi (a, b)}{\partial y} \right]^{-1} \right\| \left\| \frac{\partial \Phi (a, b)}{\partial x} \right\| |\Delta x| \leqslant$$

$$\leqslant \left| \left[\frac{\partial \Phi (a, b)}{\partial y} \right]^{-1} \frac{\partial \Phi (a, b)}{\partial x} \Delta x + \Delta y \right| \leqslant \frac{1}{2} |\Delta x| + \frac{1}{2} |\Delta y|,$$

d'où

$$\frac{1}{2} |\Delta y| \leqslant \left(\frac{1}{2} + \left\| \left[\frac{\partial \Phi (a, b)}{\partial y} \right]^{-1} \right\| \cdot \left\| \frac{\partial \Phi (a, b)}{\partial x} \right\| \right) |\Delta x|,$$

c'est-à-dire que $|\Delta y| \leqslant C |\Delta x|$ pour un $C > 0$; ensuite, en portant cette estimation dans (2), nous obtenons

$$\left| \Delta y - \left(- \left[\frac{\partial \Phi (a, b)}{\partial y} \right]^{-1} \frac{\partial \Phi (a, b)}{\partial x} \Delta x \right) \right| \leqslant$$

$$\leqslant \left\| \left[\frac{\partial \Phi (a, b)}{\partial y} \right]^{-1} \right\| o ((C + 1) |\Delta x|) = o (|\Delta x|),$$

c'est-à-dire la fonction $f(x)$ est dérivable pour $x = a$ et la formule (1) est valable ; ce qu'il fallait démontrer.

b. En vertu de *1.48*, pour que la fonction $\Phi (x, y)$ soit dérivable au point (a, b), il suffit (les autres conditions *1.53* sont supposées satisfaites) que, au voisinage du point (a, b), il existe la dérivée continue $\frac{\partial \Phi (x, y)}{\partial x}$. Alors la fonction $\Phi (x, y)$ sera dérivable non seulement au point (a, b), mais aussi dans son voisinage ; la fonction implicite $y = f(x)$ que l'on a construite dans *1.53* sera elle aussi dérivable au voisinage du point $x = a$; sa dérivée, que l'on calcule d'après la formule

$$f' (x) = - \left[\frac{\partial \Phi (x, y (x))}{\partial y} \right]^{-1} \frac{\partial \Phi (x, y (x))}{\partial x}, \tag{3}$$

analogue à (1), sera une fonction continue au voisinage de $x = a$.

c. En appliquant aux deux membres de la formule (3) l'opérateur $\frac{\partial \Phi (x, y)}{\partial y}$, nous la mettons sous une forme équivalente :

$$\frac{\partial \Phi (x, y)}{\partial x} y' (x) + \frac{\partial \Phi (x, y)}{\partial x} = 0. \tag{4}$$

La formule (3) montre que, dans l'hypothèse du théorème, *on peut trouver $y'(x)$ en dérivant l'égalité $\Phi (x, y (x)) = 0$ en x comme fonction composée de x.*

1.56.a. Théorème sur la fonction inverse. *Soit $x = \varphi (y)$: $(F \subset Y) \to (E \subset X)$ une fonction dérivable dans un voisinage d'un point $y = b \in Y$ et telle que l'opérateur $\varphi' (y)$ est continu au point $y = b$ et inversible. Soit $\varphi (b) = a \in X$. Alors*

il existe un voisinage $V_\delta = \{x \in X : |x - a| < \delta\}$ *et une fonction dérivable* $y = f(x) : V_\delta \to Y$ *tels que* $f[\varphi(y)] \equiv y$ *pour tout* $y \in V_\delta$, *l'opérateur* $f'(x)$ $(X \to Y)$ *étant l'inverse de l'opérateur* $\varphi'(y)$:

$$f'(x) = [\varphi'(y)]^{-1} \text{ où } y = f(x).$$

La démonstration du théorème découle directement du théorème sur la fonction implicite si l'on y pose $\Phi(x, y) = x - \varphi(y)$ et que l'on remarque que $\frac{\partial \Phi(x, y)}{\partial x}$ est l'opérateur unité et $\frac{\partial \Phi(x, y)}{\partial y} = -\varphi'(y)$.

b. Une application $y = f(x) : G \subset X \to Y$ ayant la dérivée continue $f'(x)$ s'appelle *difféomorphisme dans* G si elle est biunivoque (de G sur $f(G)$) et si l'application inverse $x = \varphi(y) : f(G) \to G$, définie d'une façon naturelle, possède elle aussi la dérivée continue.

En vertu de *a*, pour qu'une application $y = f(x)$ ayant la dérivée continue $f'(x)$ soit un difféomorphisme dans un voisinage $U(a)$ d'un point $a \in G$, *il suffit* que l'opérateur $f'(a)$ soit inversible. Cette condition est en même temps *nécessaire*, puisque, du moment que l'application inverse dérivable $x = \varphi(y)$ existe, les égalités $f'(a)\varphi'(b) = E_Y$, $\varphi'(b) f'(a) = E_X$ ont lieu d'après *1.33a*, donc, l'opérateur $f'(a)$ est inversible.

S'il y a un difféomorphisme $y = f(x) : G \subset X \to Y$, alors toute fonction dérivable $z = \psi(x) : G \to Z$ peut être représentée comme fonction dérivable de $f(x)$; cela découle de la formule

$$z = \psi(x) = \psi(\varphi(f(x))) = g(f(x)) \text{ où } g(y) = \psi(\varphi(y)).$$

c. Supposons que les espaces X et Y dans *b* soient de dimension finie, $X = R_n$, $Y = R_m$. Choisissons dans l'espace X les coordonnées $x_1, \ldots, x_n$, et dans l'espace Y les coordonnées $y_1, \ldots, y_m$. L'application $y = f(x)$ est alors représentée par les formules de la forme

$$y_i = f_i(x_1, \ldots, x_n) \quad (i = 1, \ldots, m).$$

L'existence de la dérivée continue $f'(x)$ dans le domaine G équivaut à l'existence et à la continuité dans ce domaine de toutes les dérivées $\frac{\partial f_i(x)}{\partial x_j}$ (*1.25a*). Supposons que $m = n$ et que la matrice de Jacobi $\left\| \frac{\partial f_i(a)}{\partial x_j} \right\|$ soit inversible; l'opérateur $f'(a)$ est donc lui aussi inversible. D'après *b*, l'application $y = f(x)$ est un difféomorphisme d'un voisinage $U(a)$ sur un voisinage $V(b)$ du point $b = f(a) \in Y$. Pour tout point $y = (y_1, \ldots, y_n) \in V(b)$, il existe un point $x = (x_1, \ldots, x_n) \in U(a)$ tel que $f(x) = y$. Par conséquent, les nombres $y_1, \ldots, y_n$, de même que les nombres $x_1, \ldots, x_n$, déterminent la position du point x dans le domaine G d'une façon univoque. Ainsi, les nombres $y_1, \ldots, y_n$ peuvent servir de *nouvelles coordonnées* du point x.

En particulier, toute fonction $\varphi(x)$: $U(a) \to R_1$ peut être représentée comme fonction $\Phi(y_1(x), \ldots, y_n(x))$ des fonctions $y_1(x), \ldots, y_n(x)$.

Ainsi, les formules

$$x_1 = r\cos\theta, \quad x_2 = r\sin\theta \tag{1}$$

déterminent dans le plan x_1, x_2 les nouvelles coordonnées r, θ qui sont les coordonnées polaires (*U5.81*). Comme

$$\det \left\| \begin{matrix} \frac{\partial x_1}{\partial r} & \frac{\partial x_1}{\partial \theta} \\ \frac{\partial x_2}{\partial r} & \frac{\partial x_2}{\partial \theta} \end{matrix} \right\| = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r,$$

on peut choisir les nombres r, θ pour nouvelles coordonnées dans un voisinage de tout point distinct du point $r = 0$ ($x_1 = x_2 = 0$); en ce point même la biunivocité de l'application est violée d'une façon évidente.

S'il y a un difféomorphisme $y = f(x)$: $G \subset R_n \to R_n$, ou bien

$$y_i = f_i(x_1, \ldots, x_n), \quad i = 1, \ldots, n,$$

alors, d'après b, toute fonction dérivable $z = \psi(x)$: $G \to Z$ peut être représentée comme fonction dérivable de $f(x)$, i.e. de $f_1(x), \ldots, f_n(x)$:

$$z = \psi(x) = g(f_1(x), \ldots, f_n(x)).$$

d. Supposons encore une fois qu'il y ait des fonctions continûment dérivables

$$y_i = f_i(x_1, \ldots, x_n), \quad i = 1, \ldots, n \tag{1}$$

qui déterminent dans un domaine $G \subset R_n$ les nouvelles coordonnées $\{y\} = \{y_1, \ldots, y_n\}$, $\det \left\| \frac{\partial f_i(x)}{\partial x_j} \right\| \neq 0$. Une ligne L_j d'équations

$$y_i = f_i(a_1, \ldots, a_{j-1}, x_j, a_{j+1}, \ldots, a_n), \quad i = 1, \ldots, n$$

s'appelle *j-ème ligne de coordonnées du système* $\{y\}$ *passant par le point* a. Les vecteurs directeurs correspondants

$$g_j = \left\{ \frac{\partial f_1(a)}{\partial x_j}, \ldots, \frac{\partial f_n(a)}{\partial x_j} \right\}, \quad j = 1, \ldots, n,$$

sont linéairement indépendants; par définition, ils forment *la base locale du système de coordonnées* $\{y\}$ au point a. N'importe quel vecteur $\xi = \sum_{i=1}^{n} \xi_i e_i$ peut être développé suivant les vecteurs de la base locale:

$$\xi = \sum_{j=1}^{n} \eta_j g_j. \tag{2}$$

Les expressions des composantes η_j par ξ_i s'obtiennent comme suit. Désignons $p_{ij} = \frac{\partial f_i(a)}{\partial x_j}$ et supposons que q_{ij} soient les éléments de la matrice inverse; il découle alors des égalités $g_j = \sum_{i=1}^{n} p_{ij} e_i$ que $e_i = \sum_{j=1}^{n} q_{ij} g_j$ et que $\sum_{j=1}^{n} \eta_j g_j = \xi = \sum_{i=1}^{n} \xi_i e_i = \sum_{j=1}^{n} \sum_{i=1}^{n} \xi_i q_{ij} g_j$, d'où, en raison de l'indépendance linéaire des vecteurs g_j, on a :

$$\eta_j = \sum_{i=1}^{n} q_{ij} \xi_i. \tag{3}$$

Les lignes de coordonnées analogues passent par chaque point x du domaine G, et en tout point $x \in G$ on peut construire une base locale; soulignons que, contrairement à la base fixe $e_1, \ldots, e_n$ de R_n, *la base locale* $\{g_i(x)\}$ *varie, en général, avec* x.

1.57. Cas de fonction numérique. Lorsque $z = \Phi(x, y)$ dans l'énoncé du théorème *1.53* est une fonction numérique d'une variable $x \in E \subset X$ et de la variable numérique $y \in F \subset R$, l'opérateur $\frac{\partial \Phi(a, b)}{\partial y}$ représente la multiplication par un nombre, son inversibilité équivalant à l'inégalité $\frac{\partial \Phi(a, b)}{\partial y} \neq 0$. Dans ce cas, on peut donner une autre démonstration du théorème qui n'utilise pas le principe des applications contractantes mais part du fait que les nombres réels sont ordonnés.

Voici cette démonstration. Supposons pour fixer les idées que $\frac{\partial \Phi(a, b)}{\partial y} > 0$. Comme $\frac{\partial \Phi(x, y)}{\partial y}$ est continu, on peut dire que $\frac{\partial \Phi(x, y)}{\partial y} > 0$ partout dans un voisinage suffisamment petit W du point (a, b), ce voisinage pouvant être choisi comme produit d'une boule $v = \{x \in X : |x - a| \leqslant r\}$ par un intervalle fermé $b_2 \leqslant y \leqslant b_1$ (fig. 1.5-1). Puisque $\Phi(a, b) = 0$, alors $\Phi(a, y) > 0$ pour $y > b$ et $\Phi(a, y) < 0$ pour $y < b$. En particulier, $\Phi(a, b_1) > 0$, $\Phi(a, b_2) < 0$. En vertu de la continuité de la fonction $\Phi(x, y)$ dans la boule V, on peut y trouver deux voisinages du point a

$$V_+ = \{x : |x - a| < \delta_1\}$$

et

$$V_- = \{x : |x - a| < \delta_2\}$$

tels que

$$\Phi(x, b_1) > 0 \quad \text{pour} \quad x \in V_+,$$
$$\Phi(x, b_2) < 0 \quad \text{pour} \quad x \in V_-.$$

Soit $\delta = \min\{\delta_1, \delta_2\}$; posons

$$U_\delta = \{x : |x - a| < \delta\}.$$

Montrons que l'on peut considérer U_δ comme voisinage cherché. Soit $x \in U_\delta$. Etant donné que $\Phi(x, b_2) < 0$, $\Phi(x, b_1) > 0$ et que la fonction $\Phi(x, y)$ d'une seule variable y croît sur l'intervalle $b_2 \leqslant y \leqslant b_1$ (d'après la condition $\frac{\partial \Phi(x, y)}{\partial y} > 0$ dans W), il existe l'unique valeur $y = y(x)$ pour laquelle $\Phi(x, y(x)) = 0$ (*U5.22* et *U5.32*). L'existence et l'unicité de solution de l'équation $\Phi(x, y) = 0$ pour tout $x \in U_\delta$ sont ainsi démontrées. Il nous reste d'établir la continuité de la fonction obtenue $y = y(x)$. Notons que, par construction, la fonction $y(x)$ vérifie les inégalités $b_2 \leqslant y(x) \leqslant b_1$.

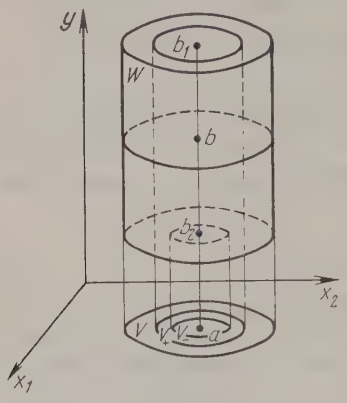


Fig. 1.5-1

Si la fonction $y(x)$ n'est pas continue en un point $c \in U_\delta$, alors il existe une suite $x_1, \ldots, x_m, \ldots$ de points de U_δ qui converge vers c et telle que la suite $y(x_m)$ ne tend pas vers $y(c)$. Il est possible d'extraire de la première suite une sous-suite $x'_1, \ldots, x'_m, \ldots$ admettant une limite $\lim y(x'_m) = A$ autre que $B = y(c)$; il résulte de $b_2 \leqslant y(x'_m) \leqslant b_1$ que $b_2 \leqslant A \leqslant b_1$. Comme $x_m \to c$ et $\Phi(x_m, y(x_m)) = 0$, on a aussi $\Phi(c, A) = 0$ en vertu de la continuité de la fonction $\Phi(x, y)$. Or, $\Phi(c, B) = \Phi(c, y(c)) = 0$.

De la sorte, sur la verticale $x = c$, $b_2 \leqslant y \leqslant b_1$, la fonction $\Phi(c, y)$ s'annule en deux points A et B, ce qui est impossible par construction. Ainsi, $y(x)$ est continue et le théorème est démontré.

En cas de fonction numérique, la valeur de la dérivée $y'(a)$ $(X \to R_1)$ peut être mise sous la forme (*1.55*):

$$y'(a) = -\frac{\partial \Phi(a, b)/\partial x}{\partial \Phi(a, b)/\partial y}.$$

1.58. Exemples élémentaires.

a. La courbe

$$x^3 + y^3 - 2xy = 0 \qquad (1)$$

dans le plan $\{x, y\}$ passe par le point (1, 1) (fig. 1.5-2). Trouver la tangente à cette courbe au point (1, 1).

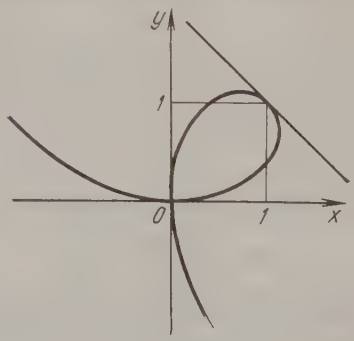


Fig. 1.5-2

Posons $\Phi(x, y) = x^3 + y^3 - 2xy$ $(R_2 \to R_1)$; évidemment, $\Phi(x, y)$ est une fonction dérivable avec ses dérivées $\frac{\partial \Phi}{\partial x}$ et $\frac{\partial \Phi}{\partial y}$ et l'on a $\Phi(1, 1) = 0$, $\frac{\partial \Phi(1, 1)}{\partial y} = 3y^2 - 2x\,|_{\{1,1\}} = 1 \neq 0$. On peut donc appliquer le théorème sur la fonction implicite: il existe une fonction $y = y(x)$, solution de l'équation $\Phi(x, y) = 0$ dans un voisinage du point (1, 1), et telle que $y(1) = 1$. Cette fonction est dérivable et, comme $\frac{\partial \Phi(1, 1)}{\partial x} = 3x^2 - 2y|_{\{1,1\}} = 1$, on a

$$y'(1) = -\frac{\partial \Phi(1,1)/\partial x}{\partial \Phi(1,1)/\partial y} = -1.$$

Par conséquent, l'équation de la tangente a la forme

$$y - 1 = -(x - 1), \tag{2}$$

soit, en différentielles ($dx = x - 1$, $dy = y - 1$) :

$$dy = -dx. \tag{3}$$

Formellement, on pourrait arriver au même résultat en dérivant directement l'équation (1), ce qui amène à l'égalité

$$3x^2\,dx + 3y^2\,dy - 2x\,dy - 2y\,dx = 0,$$

d'où, pour $x = y = 1$, on retrouve (3). Mais cette méthode ne devient légitime qu'en vertu du théorème sur la fonction implicite.

Signalons que le théorème sur la fonction implicite ne répond pas à la même question pour le point (0, 0), car $\dfrac{\partial \Phi\,(0,\ 0)}{\partial y} = 0$. En effet, c'est un point commun aux deux branches de la courbe (1) (fig. 1.5-2).

b. La surface

$$x^3 + y^3 + z^3 - 6xyz = 0 \tag{4}$$

dans l'espace $\{x,\ y,\ z\}$ passe par le point $\{1,\ 2,\ 3\}$; trouver le plan tangent en ce point.

De même que plus haut, considérons la fonction

$$\Phi\,(x,\ y,\ z) = x^3 + y^3 + z^3 - 6xyz;$$

il est évident que

$$\Phi\,(1,\ 2,\ 3) = 0, \quad \frac{\partial \Phi\,(1,\ 2,\ 3)}{\partial z} = 3z^2 - 6xy\Big|_{\{1,\ 2,\ 3\}} = 15 \neq 0,$$

$$\frac{\partial \Phi\,(1,\ 2,\ 3)}{\partial x} = 3x^2 - 6yz\Big|_{\{1,\ 2,\ 3\}} = -33,$$

$$\frac{\partial \Phi\,(1,\ 2,\ 3)}{\partial y} = 3y^2 - 6xz\Big|_{\{1,\ 2,\ 3\}} = -6.$$

En vertu du théorème sur la fonction implicite, il existe une fonction $z = z\,(x,\ y)$, solution de l'équation $\Phi\,(x,\ y,\ z) = 0$ dans un voisinage du point $x = 1$, $y = 2$, et telle que $z\,(1,\ 2) = 3$.

Cette fonction est dérivable et l'on a

$$\operatorname{grad} z\,(1,\ 2) = \left\{ \frac{\partial z\,(1,\ 2)}{\partial x},\ \frac{\partial z\,(1,\ 2)}{\partial y} \right\} =$$

$$= \left\{ -\frac{\partial \Phi\,(1,\ 2,\ 3)/\partial x}{\partial \Phi\,(1,\ 2,\ 3)/\partial z},\ -\frac{\partial \Phi\,(1,\ 2,\ 3)/\partial y}{\partial \Phi\,(1,\ 2,\ 3)/\partial z} \right\} = \left\{ \frac{33}{15},\ \frac{6}{15} \right\}.$$

Par conséquent, le plan tangent au point (1, 2, 3) a pour équation

$$z - 3 = \frac{1}{5}\,[11\,(x-1) + 2\,(y-2)], \tag{5}$$

soit, en différentielles ($dx = x - 1$, $dy = y - 2$, $dz = z - 3$),

$$dz = \frac{1}{5}\,(11dx + 2dy). \tag{6}$$

Ce résultat peut également être obtenu en dérivant d'une façon formelle l'équation (4):

$$3x^2\,dx + 3y^2\,dy + 3z^2\,dz - 6dx\cdot yz - 6x\,dy\cdot z - 6xy\cdot dz = 0. \qquad (7)$$

Pour $x = 1$, $y = 2$, $z = 3$, on retrouve (6) de (7).

1.59. Cas de fonction $\Phi(x, y)$: $R_{n+m} \to R_m$. Ici le théorème sur la fonction implicite *1.53* admet une interprétation en coordonnées qu'il est bon de comparer avec le théorème pour le système linéaire (*1.52*).

a. Théorème. *Soit*

$$\left.\begin{aligned} z_1 &= f_1(x_1, \ldots, x_n, y_1, \ldots, y_m), \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ z_m &= f_m(x_1, \ldots, x_n, y_1, \ldots, y_m) \end{aligned}\right\} \qquad (1)$$

un système de fonctions définies dans un domaine de l'espace R_{n+m}. *Supposons que les conditions suivantes soient remplies:*

(1) *Il existe un point* $\{a, b\} = \{a_1, \ldots, a_n, b_1, \ldots, b_m\}$ *tel que*

$$\left.\begin{aligned} f_1(a_1, \ldots, a_n, b_1, \ldots, b_m) &= 0, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ f_m(a_1, \ldots, a_n, b_1, \ldots, b_m) &= 0. \end{aligned}\right\} \qquad (2)$$

(2) *Il existe un voisinage* W *du point* $\{a, b\}$ *où sont définies et continues les dérivées partielles*

$$\frac{\partial f_i(x_1, \ldots, x_n, y_1, \ldots, y_m)}{\partial y_j}, \quad i = 1, \ldots, m; \quad j = 1, \ldots, m.$$

(3) *Le jacobien de la matrice*

$$\frac{\partial f(x, y)}{\partial y} = \left\| \begin{matrix} \frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_m} \\ \cdots & \cdots & \cdots \\ \frac{\partial f_m}{\partial y_1} & \cdots & \frac{\partial f_m}{\partial y_m} \end{matrix} \right\|$$

ne s'annule en aucun point du voisinage mentionné du point $\{a, b\}$.

Alors il existe un voisinage U_δ *du point* $a = \{a_1, \ldots, a_n\} \in R_n$ *où est défini un système*

$$\left.\begin{aligned} y_1 &= y_1(x_1, \ldots, x_n), \\ &\cdots\cdots\cdots\cdots \\ y_m &= y_m(x_1, \ldots, x_n) \end{aligned}\right\} \qquad (3)$$

de fonctions continues tel que

$$y_1(a_1, \ldots, a_n) = b_1, \ \ldots, \ y_m(a_1, \ldots, a_n) = b_m, \tag{4}$$

$$\left.\begin{aligned} &f_1(x_1, \ldots, x_n, y_1(x_1, \ldots, x_n), \ldots, y_m(x_1, \ldots, x_n)) \equiv 0, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &f_m(x_1, \ldots, x_n, y_1(x_1, \ldots, x_n), \ldots, y_m(x_1, \ldots, x_n)) \equiv 0. \end{aligned}\right\} \tag{5}$$

Le système possédant les propriétés ci-indiquées est unique.

Dans l'hypothèse supplémentaire d'existence de la dérivée

$$\frac{\partial f(x, y)}{\partial x} = \left\| \begin{matrix} \frac{\partial f_1}{\partial x_1} & \ldots & \frac{\partial f_1}{\partial x_n} \\ \ldots & \ldots & \ldots \\ \frac{\partial f_m}{\partial x_1} & \ldots & \frac{\partial f_m}{\partial x_n} \end{matrix} \right\|,$$

dans le voisinage W, les fonctions $y_1, \ldots, y_m$ sont dérivables pour $x \in U_\delta$ et l'on a:

$$y'(x) = \left\| \begin{matrix} \frac{\partial y_1}{\partial x_1} & \ldots & \frac{\partial y_m}{\partial x_1} \\ \ldots & \ldots & \ldots \\ \frac{\partial y_1}{\partial x_n} & \ldots & \frac{\partial y_m}{\partial x_n} \end{matrix} \right\| = -\left[\frac{\partial f(x, y)}{\partial y}\right]^{-1} \cdot \left[\frac{\partial f'(x, y)}{\partial x}\right]. \tag{6}$$

b. Exemple. La courbe qui est définie dans l'espace $\{x, y, z\}$ par le système d'équations

$$x^2 - yz = 0, \quad 3x^3 - y - 2z = 0 \tag{7}$$

passe par le point $\{1, 1, 1\}$; trouver la tangente à cette courbe au point mentionné.

Solution. Considérons la fonction $\Phi(x, y, z)$: $R_3 \to R_2$ définie par les formules $\Phi(x, y, z) = \{u, v\}$, où

$$u = x^2 - yz, \quad v = 3x^3 - y - 2z.$$

Il est évident que $\Phi(1, 1, 1) = \{0, 0\}$. On a

$$\frac{\partial \Phi(1, 1, 1)}{\partial (y, z)} = \left\| \begin{matrix} \frac{\partial u(1, 1, 1)}{\partial y} & \frac{\partial u(1, 1, 1)}{\partial z} \\ \frac{\partial v(1, 1, 1)}{\partial y} & \frac{\partial v(1, 1, 1)}{\partial z} \end{matrix} \right\| = \left\| \begin{matrix} -1 & -1 \\ -1 & -2 \end{matrix} \right\|,$$

$$\left[\frac{\partial \Phi(1, 1, 1)}{\partial (y, z)}\right]^{-1} = \left\| \begin{matrix} -2 & 1 \\ 1 & -1 \end{matrix} \right\|.$$

La matrice $\dfrac{\partial \Phi(1, 1, 1)}{\partial (y, z)}$ est inversible et l'on peut appliquer le théorème sur la fonction implicite: il existe une solution $y = y(x)$, $z = z(x)$ du système $u = 0$, $v = 0$ pour laquelle $y(1) = 1$, $z(1) = 1$. Ensuite,

$$\frac{\partial \Phi(1, 1, 1)}{\partial x} = \{2x, 9x^2\}\big|_{\{1, 1, 1\}} = \{2, 9\},$$

de sorte que

$$\{y'(1), z'(1)\} = -\left[\frac{\partial \Phi(1, 1, 1)}{\partial (y, z)}\right]^{-1} \cdot \frac{\partial \Phi(1, 1, 1)}{\partial x} =$$

$$= -\begin{Vmatrix} -2 & 1 \\ 1 & -1 \end{Vmatrix} \{2, 9\} = \{-5, 7\}.$$

Par conséquent, les équations de la tangente au point $\{1, 1, 1\}$ sont

$$\left.\begin{aligned} y - 1 &= -5(x - 1), \\ z - 1 &= 7(x - 1). \end{aligned}\right\} \tag{8}$$

La solution formelle s'obtient en différentiant les équations données (7):

$$2x\,dx - dy \cdot z - y \cdot dz = 0,$$

$$9x^2\,dx - dy - 2dz = 0,$$

en portant les valeurs $x = y = z = 1$ dans les égalités obtenues et en résolvant le système qui en résulte par rapport à dy et dz. On trouve ainsi

$$dy = -5dx, \quad dz = 7dx,$$

ce qui coïncide avec (8) en substituant $x - 1 = dx$, $y - 1 = dy$, $z - 1 = dz$.

c. Racines d'une équation algébrique comme fonctions continues et dérivables de ses coefficients. Considérons un polynôme algébrique de n-ème degré

$$z^n - a_1 z^{n-1} + a_2 z^{n-2} - \ldots + (-1)^n a_n \tag{9}$$

avec les coefficients complexes $a_1, a_2, \ldots, a_n$ et les racines $\lambda_1, \ldots, \lambda_n$. Puisqu'à tout ensemble de coefficients $a = (a_1, \ldots, a_n) \in C_n$ il correspond un unique ensemble de racines $\lambda = (\lambda_1, \ldots, \lambda_n) \in C_n$, nous avons affaire à la fonction $\lambda = \lambda(a)$: $C_n \to C_n$. Montrons que *cette fonction est continue (et dérivable) en tout point $a \in C_n$ pour lequel les racines $\lambda_1, \ldots, \lambda_n$ sont deux à deux distinctes.*

Les coefficients du polynôme (9) sont liés avec ses racines par les formules de Viète

$$\left.\begin{aligned} a_1 &= \lambda_1 + \lambda_2 + \ldots + \lambda_n, \\ a_2 &= \lambda_1\lambda_2 + \lambda_1\lambda_3 + \ldots + \lambda_{n-1}\lambda_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ a_n &= \lambda_1\lambda_2 \ldots \lambda_n; \end{aligned}\right\} \tag{10}$$

la fonction obtenue $a = a(\lambda)$ est, évidemment, partout continue et dérivable. Si l'on désigne l'accroissement de la variable λ par $d\lambda = (d\lambda_1, \ldots, d\lambda_n)$, alors la partie linéaire principale de la fonction $a(\lambda)$ est déterminée tout naturellement:

$$da(\lambda) = \frac{\partial (a_1, \ldots, a_n)}{\partial (\lambda_1, \ldots, \lambda_n)}\, d\lambda, \tag{11}$$

où $\frac{\partial (a_1, \ldots, a_n)}{\partial (\lambda_1, \ldots, \lambda_n)}$ est la matrice jacobienne complexe du système (10). L'opérateur (11) est inversible si $\det \frac{\partial (a_1, \ldots, a_n)}{\partial (\lambda_1, \ldots, \lambda_n)} \neq 0$. On peut alors appliquer le théorème sur la fonction inverse *1.56* qui garantit l'existence, la continuité et la dérivabilité de la fonction inverse $\lambda (a)$ *.

Ainsi, nous n'avons qu'à montrer que la relation $\det \frac{\partial (a_1, \ldots, a_n)}{\partial (\lambda_1, \ldots, \lambda_n)} \neq 0$ a lieu, à condition que les nombres $\lambda_1, \ldots, \lambda_n$ soient deux à deux distincts.

La démonstration se fait par récurrence sur $m = 1, \ldots, n, \ldots$ Notre affirmation est, évidemment, juste pour $m = 1$, car $\frac{\partial a_1}{\partial \lambda_1} = 1$. Supposons qu'elle le soit pour $m = n - 1$. Désignons

$$\left.\begin{aligned} b_1 &= \lambda_1 + \ldots + \lambda_{n-1}, \\ b_2 &= \lambda_1\lambda_2 + \lambda_1\lambda_3 + \ldots + \lambda_{n-2}\lambda_{n-1}, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ b_{n-1} &= \lambda_1\lambda_2 \ldots \lambda_{n-1}. \end{aligned}\right\} \qquad (12)$$

Si l'on forme le polynôme

$$z^{n-1} - b_1 z^{n-2} + b_2 z^{n-3} - \ldots + (-1)^{n-1} b_{n-1}, \qquad (13)$$

les nombres $\lambda_1, \ldots, \lambda_{n-1}$ lui serviront de racines. S'ils sont deux à deux distincts, on a $\det \frac{\partial (b_1, \ldots, b_{n-1})}{\partial (\lambda_1, \ldots, \lambda_{n-1})} \neq 0$ d'après l'hypothèse de récurrence. Donc, en complétant les formules (12) par la formule $\lambda_n = \lambda_n$, nous aurons

$$\det \frac{\partial (b_1, \ldots, b_{n-1}, \lambda_n)}{\partial (\lambda_1, \ldots, \lambda_{n-1}, \lambda_n)} \neq 0. \qquad (14)$$

Compte tenu de (12), les formules (10) prennent la forme

$$\begin{aligned} a_1 &= b_1 + \lambda_n, \\ a_2 &= b_2 + b_1\lambda_n, \\ a_3 &= b_3 + b_2\lambda_n, \\ &\ldots\ldots\ldots \\ a_n &= b_{n-1}\lambda_n. \end{aligned}$$

* Dérivabilité au sens général *1.23a*; en fait, la fonction $\lambda (a)$ est même analytique, i.e. dérivable par rapport aux variables complexes $a_1, \ldots, a_n$, ce qui ne découle pas encore des considérations de *1.23a*.

On voit que

$$\frac{\partial(a_1, \ldots, a_n)}{\partial(b_1, \ldots, b_{n-1}, \lambda_n)} = \begin{Vmatrix} 1 & & & & & 1 \\ \lambda_n & 1 & & & & b_1 \\ & \lambda_n & 1 & & & b_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & & \lambda_n & 1 & b_{n-2} \\ & & & & \lambda_n & b_{n-1} \end{Vmatrix}.$$

Calculons le déterminant de cette matrice. Pour le faire, retranchons de chaque ligne, à partir de la deuxième, la précédente multipliée par λ_n; nous aurons

$$\det \frac{\partial(a_1, \ldots, a_n)}{\partial(b_1, \ldots, b_{n-1}, \lambda_n)} =$$

$$= \begin{vmatrix} 1 & & & & 1 \\ 0 & 1 & & & b_1 - \lambda_n \\ & 0 & 1 & & b_2 - b_1\lambda_n + \lambda_n^2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ & & 0 & b_{n-1} - b_{n-2}\lambda_n + \ldots + (-1)^{n-1}\lambda_n^{n-1} \end{vmatrix} =$$

$$= (-1)^{n-1}(\lambda_n^{n-1} - b_1\lambda_n^{n-2} + \ldots + (-1)^{n-1} b_{n-1}) \neq 0,$$

puisque λ_n, par hypothèse, est distinct de chacun des nombres $\lambda_1, \ldots, \lambda_{n-1}$ et, par conséquent, n'est pas une racine du polynôme (13). Ensuite, vu (14), nous aboutissons à

$$\det \frac{\partial(a_1, \ldots, a_n)}{\partial(\lambda_1, \ldots, \lambda_n)} = \det \frac{\partial(a_1, \ldots, a_n)}{\partial(b_1, \ldots, b_{n-1}, \lambda_n)} \det \frac{\partial(b_1, \ldots, b_{n-1}, \lambda_n)}{\partial(\lambda_1, \ldots, \lambda_{n-1}, \lambda_n)} \neq 0,$$

ce qu'il fallait.

§ 1. 6. Structure locale d'une fonction dérivable

Le problème de structure locale d'une fonction dérivable $y = f(x)$: $G \subset X \to Y$ est étroitement lié au théorème sur la fonction implicite. Dans ce qui suit, les espaces X et Y sont supposés complets.

Si l'opérateur $f'(a)$ est inversible à un point donné $a \in G$, la fonction $f(x)$ applique un voisinage du point a, d'une façon biunivoque et bidérivable, sur un voisinage du point $b = f(a) \in Y$, ce qui découle du théorème *1.56* sur la fonction implicite. Comment est la situation, si l'opérateur $f'(a)$: $X \to Y$ n'est pas inversible?

1.61. Nous allons considérer quelques cas. Etablissons d'abord un lemme de la théorie générale des opérateurs linéaires continus dans les espaces normés.

Lemme. *Supposons que la restriction d'un opérateur linéaire continu* $A: X \to Y$ *sur un sous-espace fermé* $X_1 \subset X$ *représente une application inversible de* X_1 *sur tout l'espace* Y. *Il existe alors un sous-espace fermé* $X_2 \subset X$ *qui en somme droite avec* X_1 *donne tout le* X *et tel que sur* X_2 *l'opérateur* A *représente l'opérateur nul.*

Démonstration. Posons

$$X_2 = \{x \in X: Ax = 0\}.$$

Evidemment, X_2 est un sous-espace fermé dans X. Il n'a pas d'éléments communs avec le sous-espace X_1 autres que 0; en effet, si $x_0 \in X_1 \cap X_2$, alors $0 = Ax_0 = z_0 \in AX_1$, d'où $x_0 = A^{-1}z_0 = 0$. Enfin, pour tout $x \in X$, on a $Ax = Ax_1$ où $x_1 \in X_1$. Soit $x_2 = x - x_1$; nous avons $Ax_2 = Ax - Ax_1 = 0$, de sorte que $x_2 \in X_2$; ainsi la décomposition $x = x_1 + x_2$, $x_1 \in X_1$, $x_2 \in X_2$, existe pour tout $x \in X$. Elle est unique, puisque $X_1 \cap X_2 = \{0\}$. Le lemme est démontré.

1.62. Ici et dans les deux points suivants nous étudions quelques cas où l'opérateur $f'(a)$ n'est pas inversible.

a. Dans le premier cas, la non-inversibilité de l'opérateur $f'(a)$ est causée par le fait que l'espace Y est « beaucoup plus petit » que l'espace X et le domaine des valeurs de $f'(a)$ « recouvre Y plusieurs fois ».

Plus exactement, supposons qu'il existe un sous-espace fermé $X_1 \subset X$ tel que l'opérateur dérivée partielle $\frac{\partial f(a)}{\partial x_1}$ (*1.47a*) applique X_1 sur tout le Y et soit inversible. Un tel point a sera appelé *point ordinaire* de la surface de niveau correspondante $f(x) = C\ (= f(a))$.

D'après le lemme *1.61*, il existe un sous-espace fermé $X_2 \subset X$ dont la somme droite avec X_1 représente tout l'espace X. Dans ce qui suit, il n'importe pas que X_2 a la forme concrète indiquée dans le lemme *1.61*; soit X_2 n'importe quel sous-espace fermé qui en somme droite avec X_1 donne tout le X. Notons la variable $x = x_1 + x_2$ ($x_1 \in X_1$, $x_2 \in X_2$) sous forme de couple $\{x_1, x_2\}$, de sorte que $a = \{a_1, a_2\}$. Posons, comme plus haut, $b = f(a)$. Il est affirmé que la « surface de niveau » $P = \{x \in X: f(x) = b\}$ représente, au voisinage du point ordinaire $x = a$, le lieu géométrique décrit par une équation de la forme $x_1 = \varphi(x_2)$, la fonction $\varphi(x_2)$ $(X_2 \to X_1)$ étant dérivable dans un voisinage du point $a_2 \in X_2$, $\{a_1, a_2\} = a$.

Voici la démonstration.

Ayant donnée la décomposition $X = X_1 + X_2$, nous mettons la fonction « d'une variable » $f(x)$ sous forme de fonction « de deux variables » $y = f(x) = F(x_1, x_2)$, l'opérateur $\frac{\partial F(a_1, a_2)}{\partial x_1}$ étant inversible par hypothèse. D'après le théorème sur la fonction implicite, on peut résoudre l'équation $y = F(x_1, x_2)$ par rapport à x_1 au voisi-

nage du point a_1, ce qui conduit à une équation de la forme

$$x_1 = g\,(x_2, y),$$

la fonction $g\,(x_2, y)$ ayant les dérivées continues, $a_1 = g\,(a_2, b)$. En posant $y = b$, on a l'équation de la « surface de niveau » P:

$$x_1 = g\,(x_2, b),$$

résolue par rapport à la coordonnée x_1, ce qu'il fallait.

b. Cas de fonction numérique. Soit $y = f\,(x)$: $G \subset X \to R_1$ une fonction dérivable et telle que $f'\,(a) \neq 0$ pour un point $a \in G$ avec $f\,(a) = b$. Cela veut dire qu'il existe un vecteur $h \in X$ tel que $f'\,(a)\,h \neq 0$. Dans le présent cas, l'opérateur $f'\,(a)$ est une fonctionnelle linéaire (*1.24*) avec R_1 pour domaine des valeurs; il résulte de l'inégalité $f'\,(a)\,h \neq 0$ que même le sous-espace unidimensionnel engendré par le vecteur h s'applique par l'opérateur $f'\,(a)$ sur tout le R_1. On peut donc choisir ce sous-espace unidimensionnel pour le sous-espace X_1 figurant dans les constructions générales *a*. En tant que X_2 on peut prendre, d'après le lemme *1.61*, le sous-espace annulé par la fonctionnelle $f'\,(a)$. Dans les coordonnées correspondantes x_1 et x_2, on obtient l'équation suivante pour la surface de niveau $f\,(x) = b$:

$$x_1 = g\,(x_2)$$

avec la fonction $g\,(x_2)$ dérivable. En particulier, pour $a = (a_1, a_2)$, on a $a_1 = g\,(a_2)$. On peut également calculer la valeur $g'\,(a_2)$ en dérivant la relation $f\,(g\,(x_2), x_2) \equiv b$:

$$\frac{\partial f}{\partial x_1}\,g'\,(x_2) + \frac{\partial f}{\partial x_2} = 0\,;$$

vu que $\dfrac{\partial f\,(a)}{\partial x_2} = 0$, $\dfrac{\partial f\,(a)}{\partial x_1} \neq 0$, on en tire $g'\,(a_2) = 0$ en posant $(x_1, x_2) = (a_1, a_2) = a$.

La variété X_2 est dans ce cas un plan tangent à la surface $f\,(x) = b$ (*1.39c*).

c. Exemple. Soit $y = f\,(x)$: $R_3 \to R_2$ une fonction vectorielle définie par deux équations numériques

$$\left.\begin{aligned} y_1\,(x) &= f_1\,(x_1, x_2, x_3), \\ y_2\,(x) &= f_2\,(x_1, x_2, x_3). \end{aligned}\right\} \qquad (1)$$

La dérivée

$$f'\,(a) \cong \left\| \begin{matrix} \dfrac{\partial f_1\,(a)}{\partial x_1} & \dfrac{\partial f_1\,(a)}{\partial x_2} & \dfrac{\partial f_1\,(a)}{\partial x_3} \\ \dfrac{\partial f_2\,(a)}{\partial x_1} & \dfrac{\partial f_2\,(a)}{\partial x_2} & \dfrac{\partial f_2\,(a)}{\partial x_3} \end{matrix} \right\|$$

n'est sûrement pas un opérateur inversible. Supposons que le mineur

$$\begin{vmatrix} \frac{\partial f_1(a)}{\partial x_1} & \frac{\partial f_1(a)}{\partial x_2} \\ \frac{\partial f_2(a)}{\partial x_1} & \frac{\partial f_2(a)}{\partial x_2} \end{vmatrix}$$

soit non nul. Alors l'opérateur dérivée partielle par rapport au plan $X_1 = \{x_1, x_2\}$ est inversible. Si l'on choisit pour X_2 l'espace unidimensionnel $\{x_3\}$, les équations (1) prennent la forme comme on l'a vu:

$$x_1 = \varphi_1(x_3;\ y_1,\ y_2),$$
$$x_2 = \varphi_2(x_3;\ y_1,\ y_2).$$

En y fixant $y_1 = b_1$, $y_2 = b_2$, nous aurons la forme de la « surface de niveau » de la fonction $y = f(x)$:

$$x_1 = \varphi_1(x_3;\ b_1,\ b_2),$$
$$x_2 = \varphi_2(x_3;\ b_1,\ b_2);$$

c'est une courbe dans l'espace R_3. Désignons-la par P.

Le « plan tangent » à la courbe P est sa droite tangente. L'orthogonalité du gradient au plan tangent (*1.39c*) exprime le fait que l'opérateur $f'(x)$ annule le vecteur tangent à la courbe P, ce qu'on peut voir directement si l'on utilise les formules de dérivation d'une fonction implicite.

1.63. a. Dans le second cas, la non-inversibilité de l'opérateur $f'(a)$ est due à ce que l'espace Y est « beaucoup plus grand » que l'espace X et que le domaine des valeurs de l'opérateur $f'(a)$ ne recouvre pas tout le Y.

Soient $Y_1 \subset Y$ et $Y_2 \subset Y$ deux sous-espaces fermés dont la somme droite donne tout le Y. Alors, en utilisant la décomposition $y = y_1 + y_2$, $y_1 \in Y_1$, $y_2 \in Y_2$, on peut exprimer la fonction $y = f(x)$ $(X \to Y)$ par deux fonctions

$$y_1 = f_1(x) \quad (X \to Y_1), \tag{1}$$
$$y_2 = f_2(x) \quad (X \to Y_2). \tag{2}$$

Ces fonctions sont dérivables avec $f(x)$ (*1.32c*).

Supposons que la fonction $f_1(x)$ garantie l'inversibilité de l'opérateur $f_1'(a)$ $(X \to Y_1)$. Alors, d'après le théorème sur la fonction inverse, on peut résoudre l'équation (1) par rapport à x; en portant l'expression correspondante $x = \varphi(y_1)$ dans l'équation (2) nous trouvons la relation entre y_1 et y_2:

$$y_2 = f_2[\varphi(y_1)]. \tag{3}$$

C'est une représentation de l'image de la fonction $f(x)$ qui, dans le cas en question, est non pas tout l'espace Y, mais seulement une variété dans Y, le graphique de la fonction dérivable (3).

b. Exemple. Supposons qu'une fonction $y = f(x)$: $R_2 \to R_3$ soit donnée par le système d'équations:

$$\left.\begin{aligned} y_1 &= f_1(x_1, x_2), \\ y_2 &= f_2(x_1, x_2), \\ y_3 &= f_3(x_1, x_2). \end{aligned}\right\} \tag{4}$$

Sa dérivée

$$f'(a) \cong \begin{Vmatrix} \dfrac{\partial f_1(a)}{\partial x_1} & \dfrac{\partial f_1(a)}{\partial x_2} \\ \dfrac{\partial f_2(a)}{\partial x_1} & \dfrac{\partial f_2(a)}{\partial x_2} \\ \dfrac{\partial f_3(a)}{\partial x_1} & \dfrac{\partial f_3(a)}{\partial x_2} \end{Vmatrix}$$

n'est sûrement pas un opérateur inversible. Supposons tout de même que le mineur

$$\begin{vmatrix} \dfrac{\partial f_1(a)}{\partial x_1} & \dfrac{\partial f_1(a)}{\partial x_2} \\ \dfrac{\partial f_2(a)}{\partial x_1} & \dfrac{\partial f_2(a)}{\partial x_2} \end{vmatrix}$$

soit non nul. Alors, pour la fonction $y = F(x)$ $(R_2 \to R_2)$

$$\begin{aligned} y_1 &= f_1(x_1, x_2), \\ y_2 &= f_2(x_1, x_2), \end{aligned}$$

l'opérateur $F'(a)$ sera inversible. Selon ce qu'on a démontré plus haut, le premier couple d'équations du système (4) peut être résolu par rapport à x_1 et x_2 dans un voisinage du point a:

$$\left.\begin{aligned} x_1 &= \varphi_1(y_1, y_2), \\ x_2 &= \varphi_2(y_1, y_2), \end{aligned}\right\} \tag{5}$$

les fonctions φ_1 et φ_2 possédant leurs dérivées continues. Puis, en portant les expressions (5) à la dernière équation du système (4), on a

$$y_3 = f_3[\varphi_1(y_1, y_2), \varphi_2(y_1, y_2)].$$

C'est l'équation d'une surface dans l'espace R_3 qui représente le domaine des valeurs de la fonction $y = f(x)$.

1.64. Il existe d'autres raisons de la non-inversibilité de l'opérateur $f'(a)$.

a. Pour les mettre en évidence et poser correctement le problème, considérons d'abord une transformation linéaire $y = Ax$: $R_n \to R_m$ définie (dans certaines bases fixes de ces espaces) par les formules

$$y_i = \sum_{j=1}^{n} a_{ij} x_j \quad (i = 1, \ldots, m). \tag{1}$$

Lorsque le point x parcourt R_n, le vecteur $y = Ax$ ne parcourt en général pas *tout* l'espace R_m; l'image exacte de l'espace R_n par

l'application (1) est un sous-espace $\operatorname{Im} A = R \subset R_m$. Naturellement, ce sont les questions suivantes qui se posent tout de suite: quelle est la dimension du sous-espace R? Comment le décrire en termes des coordonnées y_i? Une autre question y est étroitement liée. Ce n'est en général pas un seul point qui se transforme en un point donné $y = (y_1, \ldots, y_m) \in R_m$, mais tout un ensemble de points qui s'appelle *image réciproque complète du point y* et se désigne par $A^{-1}y$ *. Pour le point $y = 0$, l'ensemble $A^{-1}y$ est un sous-espace $R_0 \subset X$ qui est appelé *noyau* de l'application A et noté $\ker A$ **. Pour n'importe quel autre point $y \in \operatorname{Im} A$, l'image réciproque complète est la translation du sous-espace R_0 d'un certain vecteur (en vertu du théorème connu: la solution générale d'un système linéaire non homogène est la somme d'une solution quelconque de ce système et de la solution générale du système homogène correspondant). Ainsi, les images réciproques complètes de points différents sont des variétés linéaires d'une même dimension. Quelle est cette dimension? Comment décrire ces variétés linéaires en termes des coordonnées x_j?

Dans l'algèbre linéaire on répond aux deux questions ci-posées. A savoir: le sous-espace $R = \operatorname{Im} A$ est engendré par les colonnes de la matrice $A \cong \| a_{ij} \|$; la dimension de R est égale au nombre maximal de colonnes linéairement indépendantes de la matrice A, c'est-à-dire à son *rang*. Le sous-espace $R_0 = \ker A$ est l'espace des solutions du système linéaire homogène que voici:

$$\sum_{j=1}^{n} a_{ij}x_j = 0, \quad i = 1, \ldots, n. \tag{2}$$

Sa dimension vaut $n - r$, où r est le rang de la matrice A (cf., par exemple, *L3.51*).

Supposons, pour fixer les idées, qu'un mineur de base de la matrice A soit disposé dans r premières lignes et r premières colonnes. Alors chaque ligne de la matrice A, à partir de la $(r + 1)$-ème, s'exprime linéairement par r premières lignes, ce qu'on peut noter comme suit:

$$a_{sj} = C_{s1}a_{1j} + \ldots + C_{sr}a_{rj}, \quad (s = r + 1, \ldots, m), \tag{3}$$

les constantes $C_{s1}, \ldots, C_{sr}$ étant définies d'une façon univoque. Il en résulte que les quantités $y_1, \ldots, y_m$ vérifient les relations

$$y_s = C_{s1}y_1 + \ldots + C_{sr}y_r \quad (s = r + 1, \ldots, m). \tag{4}$$

D'autre part, s'il y a des quantités $y_1, \ldots, y_m$ liées par les formules (4), alors *il existe les valeurs $x_1, \ldots, x_n$ vérifiant, avec les*

* Soulignons que, dans cette situation, l'opérateur A^{-1} n'existe en général pas.

** Du Kernel — noyau (angl.).

y_i *donnés, les relations* (1) ; pour le démontrer, on peut poser $x_{r+1} = \ldots = x_n = 0$ et déterminer $x_1, \ldots, x_r$ de r premières équations du système (1) ; les y_j donnés et x_i ainsi trouvés satisfont aux r premières équations (1), après quoi les $m - r$ équations qui restent sont satisfaites en vertu des formules (3). Par conséquent, les relations (2) fournissent une description *complète* du sous-espace $R \subset Y$. Les équations (2), à leur tour, décrivent complètement le sous-espace $R_0 \subset X$.

b. Discutons à présent les questions analogues pour une fonction dérivable arbitraire $y = f(x)$ qui agit d'un domaine $G \subset X = R_n$ dans l'espace $Y = R_m$. En coordonnées, la fonction $y = f(x)$ est représentée par le système d'équations de la forme

$$y_i = f_i(x_1, \ldots, x_n) \quad (i = 1, \ldots, m),$$

où les fonctions $f_i(x)$ sont définies, continues et possèdent les dérivées partielles continues dans le domaine G. Posons les questions suivantes. *Quelle est la dimension de l'image d'un voisinage d'un point* $a \in G$? *Par quelles équations en coordonnées* y_i *peut-on décrire cette image*? *Quelle est la dimension de l'image réciproque complète du point* $b = f(a)$? *Par quelles équations en coordonnées* x_j *peut-on décrire cette image réciproque*? Nous considérons la dimension dont il s'agit au sens suivant : il sera montré que les ensembles cherchés sont décrits chacun à l'aide d'un système de fonctions d'un certain nombre des variables réelles indépendantes ; c'est ce dernier nombre que nous appelons *dimension* de l'ensemble cherché.

Nous donnerons les réponses à toutes ces questions en supposant que *le rang de la matrice de Jacobi*

$$f'(x) \cong \left\| \begin{array}{ccc} \frac{\partial f_1(x)}{\partial x_1} & \cdots & \frac{\partial f_1(x)}{\partial x_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial f_m(x)}{\partial x_1} & \cdots & \frac{\partial f_m(x)}{\partial x_n} \end{array} \right\|$$

soit une valeur constante r *dans un voisinage* U *du point* a.

En général, le rang de la matrice de Jacobi varie selon les points ; si l'on considère un point a_0, où ce rang prend sa valeur maximale r_0, alors, en raison de continuité, un mineur d'ordre r_0 qui n'est pas nul en a_0 sera non nul dans un voisinage de ce point. Ainsi, la supposition que le rang soit constant au voisinage de a est bien remplie pour certains points $a \in G$.

Sans restreindre la généralité, on peut supposer qu'un mineur de base de la matrice $f'(x)$, pour tous les $x \in U$, se trouve dans les premières lignes et colonnes. En effet, on peut l'y mettre en point a même, en permutant, s'il le faut, les coordonnées dans R_n et R_m ; vu la continuité de la matrice de Jacobi, ce mineur restera non nul dans un voisinage de a.

Théorème du rang. *Dans les suppositions formulées:*

(1) *Pour un voisinage* $U \ni a$, *l'ensemble de toutes les valeurs de la fonction* $f(x)$ *dans un voisinage du point* $b = f(a)$ *est décrit par le système d'équations de la forme*

$$y_s = \varphi_s(y_1, \ldots, y_r); \quad s = r+1, \ldots, m,$$

avec les fonctions $\varphi_{r+1}, \ldots, \varphi_m$ *dérivables, et se détermine donc par* r *paramètres libres* $y_1, \ldots, y_r$.

(2) *Il existe un voisinage* V *du point* b *tel que tout point de l'ensemble* $\operatorname{Im} f \cap V$ *est l'image d'un ensemble de points* x *décrit, dans les limites du voisinage* U, *par un système de la forme*

$$x_j = \psi_j(x_{r+1}, \ldots, x_n) \quad (j = 1, \ldots, r),$$

où les fonctions ψ_j *sont dérivables; il se détermine donc par* $n - r$ *paramètres libres* $x_{r+1}, \ldots, x_n$.

La démonstration du théorème du rang sera donnée dans *1.66*.

1.65. Théorème abstrait du rang (Lang). Soient $X = X_1 + X_2$, $Y = Y_1 + Y_2$ deux espaces normés complets, sommes droites de sous-espaces fermés. Pour tous $x \in X$ et $y \in Y$, il y a les décompositions $x = x_1 + x_2$, $x_1 \in X_1$, $x_2 \in X_2$, $y = y_1 + y_2$, $y_1 \in Y_1$, $y_2 \in Y_2$ définies d'une façon univoque. Toute fonction $y = f(x): X \to Y$ peut être mise sous forme d'un couple d'équations:

$$y_1 = F_1(x_1, x_2): X \to Y_1,$$
$$y_2 = F_2(x_1, x_2): X \to Y_2.$$

A la dérivée $f'(x)$ il correspond la matrice opératorielle

$$f'(x) \cong \left\| \begin{matrix} \frac{\partial F_1(x)}{\partial x_1} & \frac{\partial F_1(x)}{\partial x_2} \\ \frac{\partial F_2(x)}{\partial x_1} & \frac{\partial F_2(x)}{\partial x_2} \end{matrix} \right\|.$$

Fixons les points $a = a_1 + a_2$ et $b = b_1 + b_2 = f(a) \in Y$.

Théorème. *Supposons que la fonction* $f(x)$ *soit dérivable dans un domaine* $G \subset X$ *contenant le point* a *et qu'elle satisfasse aux conditions suivantes:*

(1) $\frac{\partial F_1(x)}{\partial x} h = 0$ *implique* $\frac{\partial F_2(x)}{\partial x} h = 0$.

(2) $\frac{\partial F_1(a)}{\partial x_1}$ *est une application inversible de* X_1 *sur* Y_1.

Il existe alors trois voisinages $U(a) \in X$, $V(b_1) \in Y_1$, $W(a_2) \subset X_2$ *tels que*

(a) *pour* $x \in U(a)$ *et* $y_1 \in V(b_1)$, *le domaine des valeurs de la fonction* $y = f(x)$ *peut être donné par une équation de la forme* $y_2 = \varphi(y_1)$;

(b) *pour tout* $y \in f(U)$, *l'image réciproque complète* $f^{-1}(y)$, *pour* $x_2 \in W(a_2)$ *peut être donnée par une équation de la forme* $x_1 = \psi(x_2)$.

Ici φ *et* ψ *sont des fonctions possédant dans les voisinages correspondants les dérivées continues.*

Démonstration. Considérons la fonction $\Phi(y_1, x_1, x_2) = F_1(x_1, x_2) - y_1$ $(G \times Y_1 \to Y_1)$. Il est évident que $\Phi(b_1, a_1, a_2) = 0$ et que l'opérateur $\frac{\partial \Phi(b_1, a_1, a_2)}{\partial x_1} = \frac{\partial F_1(a)}{\partial x_1}$ est inversible. D'après le théorème *1.53* sur la fonction implicite, il existe trois voisinages $V(b_1) \subset Y_1$, $W(a_2) \subset X_2$, $W(a_1) \subset X_1$ et une fonction $g(x_2, y_1)$: $W(a_2) \times V(b_1) \to W(a_1)$ continûment dérivable et telle que l'équation $F_1(x_1, x_2) - y_1 = 0$ soit équivalente à l'équation $x_1 = g(x_2, y_1)$. Autrement dit, on a

$$F_1(g(x_2, y_1), x_2) \equiv y_1. \tag{1}$$

Posons $U(a) = W(a_1) \times W(a_2) \cap \{x: F_1(x) \in V(b_1)\}$. Maintenant on peut mettre l'équation

$$y_2 = F_2(x_1, x_2),$$

pour $x \in U(a)$, sous forme de

$$y_2 = F_2(g(x_2, y_1), x_2). \tag{2}$$

Montrons que le second membre ne dépend pas de x_2. En dérivant (1) et (2) par rapport à x_2 nous obtenons:

$$\frac{\partial F_1}{\partial x_1}\frac{\partial g}{\partial x_2} + \frac{\partial F_1}{\partial x_2} = 0, \tag{3}$$

$$\frac{\partial F_2}{\partial x_1}\frac{\partial g}{\partial x_2} + \frac{\partial F_2}{\partial x_2} = \frac{\partial y_2}{\partial x_2}. \tag{4}$$

Il suit de (3) que le résultat de l'application de l'opérateur $\frac{\partial F_1(x)}{\partial x}$ au vecteur $\left\{\frac{\partial g}{\partial x_2} h_2, h_2\right\}$ est nul, quel que soit $h_2 \in X_2$. Mais en vertu de la supposition (1), il en est de même pour le résultat de l'application de l'opérateur $\frac{\partial F_2(x)}{\partial x}$ au même vecteur; on a donc $\frac{\partial y_2}{\partial x_2} \equiv 0$, et le second membre de (2) ne dépend pas de x_2. Par conséquent, l'équation (2), pour $y_1 \in V(b_1)$, prend la forme

$$y_2 = \varphi(y_1),$$

la fonction $\varphi(y_1)$ ayant la dérivée continue de même que F_2 et g. La propriété (a) est démontrée.

Envisageons à présent l'image réciproque complète d'un point $y = \{y_1, y_2\} = f(x) \in f(U)$. On a vu qu'un tel point $\{y_1, y_2\}$ est complètement déterminé par la première composante y_1. Or, étant donnés x_2 et y_1, cela détermine d'une façon univoque, dans le voisinage mentionné $W(a_2) \times V(b_1)$, la composante $x_1 = g(x_2, y_1)$ qui, pour y_1 fixe, représente une fonction de $x_2 \in W(a_2)$ possédant la dérivée continue.

Le théorème est démontré.

1.66.a. Démonstration du théorème du rang *1.64b*.

Dans l'énoncé du théorème en question, on s'est donné une fonction dérivable $y = f(x)$ $(G \subset R_n \to R_m)$, soit, en coordonnées:

$$y_i = f_i(x_1, \ldots, x_n) \quad (i = 1, \ldots, m).$$

On a supposé que le rang de la matrice de Jacobi

$$\left\| \begin{matrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{matrix} \right\|$$

est un nombre constant r dans un voisinage U d'un point $a \in R_n$ et qu'un mineur de base de cette matrice est disposé dans ses r premières lignes et r premières colonnes.

Nous établirons le théorème du rang en tant que conséquence de *1.65*. Posons dans *1.65* $X = R_n$, $Y = R_m$. Puis définissons le sous-espace $Y_1 \subset Y = R_m$ par r premiers vecteurs de base de l'espace R_m, et le sous-espace Y_2, par les $m-r$ derniers vecteurs de base. L'égalité $\frac{\partial F_1(x)}{\partial x} h = 0$ est alors équivalente au système d'équations

$$\sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j} h_j = 0 \quad (i = 1, \ldots, r), \tag{1}$$

et $\frac{\partial F_2(x)}{\partial x} h = 0$, au système

$$\sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j} h_j = 0 \quad (i = r+1, \ldots, m). \tag{2}$$

Les lignes de la matrice $f'(x)$, à partir de la $(r+1)$-ème, étant des combinaisons linéaires des lignes précédentes, les équations (2) s'avèrent conséquences de (1). L'hypothèse (1) du théorème *1.65* est donc remplie. Définissons à présent le sous-espace $X_1 \subset X = R_n$ par les r premiers vecteurs de base de l'espace R_n et le sous-espace X_2 par les $n-r$ derniers vecteurs de base. Alors l'opérateur $\frac{\partial F_1(a)}{\partial x_1}$ donné par la matrice

$$\left\| \begin{matrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_r} \\ \cdots & \cdots & \cdots \\ \frac{\partial f_r}{\partial x_1} & \cdots & \frac{\partial f_r}{\partial x_r} \end{matrix} \right\|$$

est inversible car son déterminant est non nul; l'hypothèse (2) du théorème *1.65* est donc remplie elle aussi. Il reste à formuler la conclusion du théorème pour le cas en question: il existe trois voisinages $U(a) \subset R_n$, $V(b_1) \subset Y_1 = R_r$, $W(a_2) \subset X_2 = R_{n-r}$ tels que, pour $x \in U(a)$ et $(y_1, \ldots, y_r) \in V(b_1)$, l'ensemble des valeurs de la fonction $y = f(x)$ peut être représenté par les équations de la forme

$$y_{r+1} = \varphi_{r+1}(y_1, \ldots, y_r), \ldots, y_m = \varphi_m(y_1, \ldots, y_r);$$

pour tout $y \in f(U)$, l'image réciproque complète $f^{-1}(y)$ peut être donnée par les équations de la forme

$$x = \psi_1(x_{r+1}, \ldots, x_n), \ldots, x_r = \psi_r(x_{r+1}, \ldots, x_n)$$

avec $\{x_{r+1}, \ldots, x_n\} \in W(a_2)$, les fonctions φ_i et ψ_j étant continûment dérivables dans les voisinages respectifs.

Or, ce sont les conclusions cherchées du théorème *1.64* qui se trouve donc démontré.

Le théorème du rang peut être interprété comme suit en langage de dimensions (si l'on sous-entend par dimension le nombre minimal de paramètres qu'il faut pour représenter un ensemble donné avec les fonctions dérivables): la dimension de l'image d'un voisinage d'un point par l'application $y = f(x)$ vaut r; celle de l'image réciproque complète $f^{-1}(y)$ de tout point y dans $U(a)$ vaut $n - r$.

b. La notion de *dépendance linéaire* de vecteurs (lignes de nombres, formes linéaires, etc.) formulée en algèbre linéaire peut être étendue sur les fonctions de la manière suivante.

Soit $y = f(x)$ $(G \to R_m)$ une fonction définie et continûment dérivable dans un domaine $G \subset R_n$, de sorte que toutes les composantes

$$y_i(x) = f_i(x_1, \ldots, x_n), \quad i = 1, \ldots, m, \tag{1}$$

possèdent leurs dérivées partielles continues par rapport à $x_1, \ldots, x_n$.

Supposons ensuite que le rang r de la matrice de Jacobi

$$f'(x) \cong \left\| \begin{matrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{matrix} \right\|$$

soit constant dans le domaine G. Si, de plus, $r = m$, alors les fonctions $f_1(x), \ldots, f_m(x)$ sont dites *indépendantes dans* G; si $r < m$ elles sont dites *dépendantes dans* G. Le théorème du rang *1.64b* permet de décrire les propriétés géométriques des fonctions indépendantes et dépendantes. Ainsi, les fonctions $f_1(x), \ldots, f_n(x)$ qui réalisent un difféomorphisme du domaine G sur $f(G)$ sont indépendantes, car

leur matrice $f'(x)$ est non dégénérée, $r = m = n$. La réciproque a lieu sous une forme un peu plus faible : si $r = m = n$, les fonctions $f_1(x), \ldots, f_n(x)$ réalisent un difféomorphisme d'un voisinage $U(a)$ de chaque point $a \in G$ sur le voisinage correspondant du point $f(a) \in f(G)$. Si, de plus, l'application $y = f(x)$ est biunivoque (de G sur $f(G)$), alors les fonctions $f_1(x), \ldots, f_n(x)$ réalisent un difféomorphisme de G sur $f(G)$. Lorsque $r = m < n$, les fonctions (indépendantes) $f_1(x), \ldots, f_m(x)$ appliquent un voisinage de chaque point $a \in G$ sur le voisinage correspondant du point $f(a)$, mais d'une façon non biunivoque, les points $f(x)$ ont leurs images réciproques $(n - m)$-dimensionnelles dans G. Enfin, si $r < m$, les fonctions (dépendantes) $f_1(x), \ldots, f_m(x)$ appliquent un voisinage $U(a)$ de chaque point $a \in G$ sur une surface r-dimensionnelle dans R_m avec pour équation $y_{r+1} = \varphi(y_1, \ldots, y_r)$, de sorte que, pour une numération convenable des coordonnées, les fonctions $f_1(x), \ldots, f_m(x)$ sont liées dans $U(a)$ par les relations de la forme $f_{r+1}(x) = \varphi(f_1(x), \ldots, f_r(x))$.

c. *Pour tout difféomorphisme* $\pi : G \subset R_n \to V \subset Z (= R_n)$, *les fonctions indépendantes* $f_1(x), \ldots, f_m(x)$ *deviennent fonctions indépendantes* $\varphi_1(z) = f_1(\pi^{-1}z), \ldots, \varphi_m(z) = f_m(\pi^{-1}z)$. En effet, la matrice de Jacobi de l'application $\varphi(z) = \{\varphi_1(z), \ldots, \varphi_m(z)\}$ est le produit de la $(n \times m)$-matrice non dégénérée de Jacobi du difféomorphisme π^{-1} et de la $(n \times m)$-matrice de l'application $f(x) = \{f_1(x), \ldots, f_m(x)\}$; or, le rang d'une matrice ne change pas en la multipliant par n'importe quelle matrice non dégénérée (*L4.67*).

1.67. Problème d'équivalence.

a. La question de la structure locale d'une fonction dérivable a encore un aspect important ; de même que dans *1.64a*, nous étudions d'abord le cas de fonction linéaire $y = f(x) : X = R_n \to Y = R_m$ donnée par les relations

$$y_i = \sum_{j=1}^{n} a_{ij} x_j, \quad i = 1, \ldots, m. \tag{1}$$

Supposons que dans les espaces X et Y on soit en droit de passer aux nouvelles coordonnées par n'importe quelle transformation linéaire non dégénérée ; on se demande, sous quelle forme, la plus simple possible, peut-on mettre les équations (1) ?

Afin d'y répondre, supposons encore une fois que le rang du système (1) soit r et qu'un mineur de base de la matrice $A = \| a_{ij} \|$ occupe ses r premières lignes et r premières colonnes. Alors, comme nous l'avons vu, les quantités $y_1, \ldots, y_m$ sont liées par les équations *1.64* (4) :

$$y_s = C_{s1} y_1 + \ldots + C_{sr} y_r \quad (s = r + 1, \ldots, m). \tag{2}$$

Introduisons dans l'espace X les nouvelles coordonnées $\xi_1, \ldots, \xi_n$ selon les formules :

$$\left.\begin{aligned}
\xi_1 &= a_{11}x_1 + \ldots + a_{1r}x_r + \ldots + a_{1n}x_n,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
\xi_r &= a_{r1}x_1 + \ldots + a_{rr}x_r + \ldots + a_{rn}x_n,\\
\xi_{r+1} &= x_{r+1},\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
\xi_n &= x_n.
\end{aligned}\right\} \qquad (3)$$

Les quantités $\xi_1, \ldots, \xi_n$ peuvent effectivement servir de nouvelles coordonnées puisque le déterminant du système (3) est, évidemment, égal au mineur de base de la matrice A, donc n'est pas nul. Introduisons dans l'espace Y les nouvelles coordonnées $\eta_1, \ldots, \eta_m$ selon les formules :

$$\left.\begin{aligned}
\eta_1 &= y_1,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
\eta_r &= y_r,\\
\eta_{r+1} &= -C_{r+1,1}y_1 - \ldots - C_{r+1,r}y_r + y_{r+1},\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
\eta_m &= -C_{m1}y_1 - \ldots - C_{mr}y_r + y_m.
\end{aligned}\right\} \qquad (4)$$

Le déterminant de ce système vaut 1, et les quantités $\eta_1, \ldots, \eta_m$ peuvent effectivement servir de nouvelles coordonnées dans l'espace Y. Vu (2), les égalités (1) prennent la forme

$$\left.\begin{aligned}
\eta_1 &= \xi_1,\\
&\ldots\ldots\ldots\\
\eta_r &= \xi_r,\\
\eta_{r+1} &= 0,\\
&\ldots\ldots\ldots\\
\eta_m &= 0.
\end{aligned}\right\} \qquad (5)$$

Il est évident que c'est la plus simple des formes sous lesquelles on peut mettre la transformation linéaire (1) en passant aux nouvelles coordonnées dans les espaces X et Y.

b. Etudions le problème analogue pour une fonction dérivable $y = f(x)$ qui agit de $G \subset X = R_n$ dans $Y = R_m$. En coordonnées, cette fonction s'écrit comme suit :

$$y_i = f_i(x_1, \ldots, x_n) \quad (i = 1, \ldots, m). \qquad (6)$$

On se demande, sous quelle forme, la plus simple possible, peut-on mettre la fonction $y = f(x)$ du moment qu'on soit en droit de passer, au voisinage d'un point donné $a \in G$ et au voisinage du point

$b = f(a)$, aux nouvelles coordonnées par n'importe quelle transformation dérivable et inversible.

Pour y répondre, supposons remplie l'hypothèse du théorème du rang *1.64b*. Introduisons dans l'espace X les nouvelles coordonnées selon les formules

$$\left.\begin{aligned} \xi_1 &= f_1(x_1, \ldots, x_n), \\ &\ldots\ldots\ldots\ldots \\ \xi_r &= f_r(x_1, \ldots, x_n), \\ \xi_{r+1} &= x_{r+1}, \\ \xi_n &= x_n. \end{aligned}\right\} \tag{7}$$

Les quantités $\xi_1, \ldots, \xi_n$ peuvent effectivement servir de nouvelles coordonnées dans un voisinage du point a, puisque le jacobien $\frac{\partial(\xi_1, \ldots, \xi_n)}{\partial(x_1, \ldots x_n)}$ est égal au mineur de base de la matrice A au point a, donc non nul par hypothèse. En vertu du théorème du rang *1.64b*, les quantités $y_1, \ldots, y_m$ sont liées, dans un voisinage $V_0(b)$ du point $b = f(a)$, par les équations

$$y_s = \varphi_s(y_1, \ldots, y_r), \qquad s = r+1, \ldots, m. \tag{8}$$

Introduisons les quantités

$$\left.\begin{aligned} \eta_1 &= y_1, \\ &\ldots\ldots\ldots\ldots \\ \eta_r &= y_r, \\ \eta_{r+1} &= -\varphi_{r+1}(y_1, \ldots, y_r) + y_{r+1}, \\ &\ldots\ldots\ldots\ldots \\ \eta_m &= -\varphi_m(y_1, \ldots, y_r) + y_m. \end{aligned}\right\} \tag{9}$$

Ici $\frac{\partial(\eta_1, \ldots, \eta_m)}{\partial(y_1, \ldots, y_m)} = 1$, donc on peut choisir les quantités $\eta_1, \ldots, \eta_m$ pour nouvelles coordonnées dans un voisinage $V_1(b)$. A présent, vu (7), les relations (6) se mettent sous la forme

$$\left.\begin{aligned} \eta_1 &= \xi_1, \\ &\ldots\ldots\ldots \\ \eta_r &= \xi_r, \\ \eta_{r+1} &= 0, \\ &\ldots\ldots\ldots \\ \eta_m &= 0. \end{aligned}\right\} \tag{10}$$

C'est, évidemment, la plus simple des formes que la fonction $y = f(x)$ peut prendre en effectuant les transformations dérivables et inversibles aux voisinages des points a et b.

c. Nous allons formuler un théorème qui généralise les constructions a et b au cas de domaines dans un espace de Banach. On ne pourra sûrement plus utiliser les coordonnées ; la généralisation sera basée sur la notion d'*équivalence* de deux applications.

Soient X, Y, Ξ, H des espaces de Banach et $y = \varphi(x) : U \subset X \to V \subset Y$ et $\xi = \psi(\eta) : M \subset \Xi \to N \subset H$ deux fonctions dérivables. Soient ensuite $\omega : U_0 \subset U \to M_0 \subset M$, $\omega(a) = \alpha$ et $\pi : V_0 \subset V \to N_0 \subset N$, $\pi(b) = \beta$ deux difféomorphismes. Les applications φ et ψ sont dites *équivalentes* si, pour tout $x \in U_0$, on a

$$\pi\varphi(x) = \psi(\omega x). \tag{11}$$

On trace parfois le « diagramme des applications » :

$$\begin{array}{ccc} U_0 & \xrightarrow{\varphi} & V_0 \\ {\scriptstyle\omega}\downarrow & & \downarrow{\scriptstyle\pi} \\ M_0 & \xrightarrow{\psi} & N_0 \end{array}$$

On peut traiter la relation (11) comme une sorte de « commutativité » de ce diagramme : en partant de n'importe quel point $x \in U_0$ dans le sens des flèches φ et π on arrive au même point du domaine N_0 qu'en partant dans le sens des flèches ω et ψ. Dans le cas de dimension finie, l'équivalence des applications φ et ψ signifie la possibilité de passer de φ à ψ avec une transformation dérivable et inversible des coordonnées, dans un voisinage $U_0 \subset U$ et dans un voisinage $V_0 \subset V$.

d. Etablissons le théorème d'équivalence suivant :

Théorème. *Dans l'hypothèse du théorème 1.65, il existe deux voisinages $U_0 \subset U$ (a) et $V_0 \subset V$ (b) tels que l'application $y = f(x)$ est équivalente (dans ces voisinages) à l'application $\psi : M_0 \subset \Xi = Y_1 + X_2 \to N_0 \subset H = Y_1 + Y_2$ définie par la formule* $\psi(y_1, x_2) = (y_1, 0)$.

Démonstration. Considérons les applications

$$\omega(x_1, x_2) = (F_1(x_1, x_2), x_2) : U \subset X \to \Xi,$$
$$\pi(y_1, y_2) = (y_1, -\varphi(y_1) + y_2) : V \subset Y \to H.$$

La matrice opératorielle de l'application $\frac{d\omega}{dx}$ a la forme

$$\begin{Vmatrix} \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} \\ 0 & E_2 \end{Vmatrix}$$

et, comme $\frac{\partial F_1(a)}{\partial x_1}$ est inversible par hypothèse, $\frac{d\omega(a)}{dx}$ est inversible (*1.14i*). Par conséquent, d'après le théorème sur la fonction inverse

1.56, l'application ω est un difféomorphisme d'un voisinage $U_0(a) \subset U$ sur un voisinage $M_0 \subset \Xi$.

D'une façon analogue, la matrice opératorielle de l'application $\frac{d\pi}{dy}$ a la forme

$$\left\| \begin{matrix} E_1 & 0 \\ -\varphi'(y_1) & E_2 \end{matrix} \right\|$$

d'où, par *1.14i*, l'opérateur $\frac{d\pi}{dy}$ s'avère lui aussi inversible. Donc, l'application π est, elle aussi, un difféomorphisme d'un voisinage $V_0(b) \subset V$ sur un voisinage $N_0 \subset H$.

Posons $\psi(y_1, x_2) = (y_1, 0)$. Montrons que, pour tout $x \in U_0(a)$, on a $\pi f(x) = \psi(\omega x)$.

En effet, il résulte de $\omega(x) = (F_1(x_1, x_2), x_2)$ que

$$\psi(\omega(x)) = (F_1(x_1, x_2), 0);$$

or, $f(x) = (F_1(x_1, x_2), F_2(x_1, x_2))$ implique $y_2 = \varphi(y_1)$; par conséquent,

$$\pi f(x) = (F_1(x_1, x_2), 0) = \psi(\omega(x)).$$

Les applications f et ψ se trouvent donc équivalentes, et le théorème est démontré.

1.68. P l i s. On a démontré le théorème du rang en supposant que le rang de l'opérateur $f'(x)$ $(R_n \to R_m)$ reste constant dans un voisinage du point considéré. En général, en raison de continuité, si le rang de $f'(a)$ vaut r au point a, il existe un voisinage où il n'est pas inférieur à r; tout de même, si $r < \min(n, m)$, alors il existe des points x autant proches que l'on veut du point a et tels que le rang de $f'(x)$ y dépasse r. Disons encore quelques mots.

Envisageons d'abord l'application $y = f(x)$: $R_2 \to R_2$ donnée par les formules

$$y_1 = x_1, \qquad y_2 = x_2^2 \tag{1}$$

et ayant pour dérivée

$$f'(x) \cong \left\| \begin{matrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} \end{matrix} \right\| = \left\| \begin{matrix} 1 & 0 \\ 0 & 2x_2 \end{matrix} \right\|.$$

Le rang de cette matrice vaut 2 pour $x_2 \neq 0$ et 1 pour $x_2 = 0$ (sur l'axe des x_1). L'application (1) transforme tout le plan $x = \{x_1, x_2\}$ en le demi-plan supérieur $y_2 \geqslant 0$ de façon que chaque point $y = \{y_1, y_2\}$ du demi-plan supérieur possède avec $y_2 > 0$ exactement deux images réciproques $x_1 = y_1$, $x_2 = \pm \sqrt{y_2}$, l'un dans le demi-plan supérieur du plan x, l'autre dans l'inférieur. Considérons plus en détail la droite verticale $x_1 = \text{const}$, $-\infty < x_2 < +\infty$; lorsque

le point $\{x_1, x_2\}$ descend suivant cette droite de $+\infty$ vers $-\infty$, son image descend d'abord suivant la verticale $y_1 = x_1$ de $y_2 = +\infty$ jusqu'au point $y_2 = 0$, puis remonte suivant la même droite vers $y_2 = +\infty$. Il est naturel d'appeler une telle application *pli*.

Passons au cas général de l'application $y = f(x)$, $G \subset R_n \to R_n$. Supposons que le jacobien $J(x)$ de l'application $f(x)$ s'annule en un point $a \in G$. En général, le comportement de l'application $f(x)$ au voisinage d'un tel point peut être très compliqué ; nous allons montrer que, dans certaines suppositions supplémentaires, l'application $f(x)$ est du type pli. Plus exactement, nous verrons que si grad $J(a) \neq 0$, ou même si grad $J(a)$ n'appartient pas au « sous-espace gradient » de l'espace $X = R_n$ (la définition est donnée ci-dessous), alors, lorsque le point $x \in X$ se déplace le long d'une courbe perçant la surface $S = \{x \in X,\ J(x) = 0\}$ dans un voisinage du point a, le point correspondant $y = f(x)$ arrive, suivant une courbe, à la surface $f(S) \subset Y = R_n$, puis revient suivant la même courbe.

Soit $f = f(x) : X = R_n \to Y = R_n$ une application qui, dans une base $e_1, \ldots, e_n$ (de sorte que $x = \sum_1^n x_k e_k$), admette la représentation analytique suivante :

$$\begin{aligned} y_1 &= f_1(x_1, \ldots, x_n), \\ &\cdots\cdots\cdots\cdots \\ y_n &= f_n(x_1, \ldots, x_n). \end{aligned}$$

Supposons que $J(a) = 0$. Il en résulte que les vecteurs

$$\begin{aligned} \operatorname{grad} f_1(a) &= \left\{ \frac{\partial f_1(a)}{\partial x_1}, \ \ldots, \ \frac{\partial f_1(a)}{\partial x_n} \right\}, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \operatorname{grad} f_n(a) &= \left\{ \frac{\partial f_n(a)}{\partial x_1}, \ \ldots, \ \frac{\partial f_n(a)}{\partial x_n} \right\} \end{aligned}$$

sont linéairement dépendants, donc engendrent un sous-espace propre Z de l'espace X. On l'appelle *sous-espace gradient*. Soit, ensuite, grad $J(a) \neq 0$. On sait (*U7.14a*) que la dérivée d'un déterminant d'ordre n est égale à la somme de n déterminants le i-ème desquels diffère du déterminant initial par la i-ème colonne formée des dérivées de la i-ème colonne du déterminant donné ; donc, si tous les mineurs d'ordre $n - 1$ de la matrice $\| f'(a) \|$ sont nuls, alors, pour toute direction l, on a $\frac{\partial J(a)}{\partial l} = 0$, puisque chaque déterminant de la somme mentionnée peut être développé suivant la colonne qui contient les dérivées, les coefficients étant égaux à certains mineurs du déterminant initial. Or, dans notre cas, grad $J(a) \neq 0$ par hypothèse, donc il existe une direction l suivant laquelle $\frac{\partial J(a)}{\partial l} \neq 0$; par conséquent,

il existe un mineur $B(x)$ *d'ordre* $n-1$ *de la matrice* $\| f'(x) \|$ *qui vaut, pour* $x = a$, *un nombre* $B \neq 0$. Supposons, pour fixer les idées, que ce mineur soit disposé dans $n-1$ premières lignes et $n-1$ premières colonnes de la matrice $f'(a)$.

Considérons, au voisinage du point $a = (a_1, \ldots, a_n)$, les fonctions

$$\begin{aligned} \xi_1 &= f_1(x_1, \ldots, x_n), \\ &\cdots\cdots\cdots\cdots \\ \xi_{n-1} &= f_{n-1}(x_1, \ldots, x_n), \\ \xi_n &= x_n. \end{aligned}$$

L'égalité $\frac{\partial(\xi_1, \ldots, \xi_n)}{\partial(x_1, \ldots, x_n)} = B(x) \neq 0$ est, évidemment, valable. Pour cette raison, les quantités $\xi_1, \ldots, \xi_n$ peuvent servir de nouvelles coordonnées dans un voisinage du point a (*1.56c*). Dans ces coordonnées, l'application f s'écrit:

$$\begin{aligned} y_1 &= \xi_1, \\ &\cdots\cdots\cdots \\ y_{n-1} &= \xi_{n-1}, \\ y_n &= \varphi(\xi_1, \ldots, \xi_n), \end{aligned}$$

où $\varphi(\xi_1, \ldots, \xi_n) = f_n(x_1(\xi), \ldots, x_{n-1}(\xi), \xi_n)$. En outre, conformément à *1.33*, nous avons $\frac{dy}{dx} = \frac{dy}{d\xi} \cdot \frac{d\xi}{dx}$ et

$$J(x) = \det \left\| \frac{dy}{dx} \right\| = \det \left\| \frac{dy}{d\xi} \right\| \cdot \det \left\| \frac{d\xi}{dx} \right\| = B(x) \frac{\partial \varphi(x)}{\partial \xi_n}; \quad \frac{\partial \varphi(a)}{\partial \xi_n} = 0.$$

Ensuite, en désignant $\psi = \frac{\partial \varphi}{\partial \xi_n}$ et utilisant *1.34d*, nous trouvons

$$\operatorname{grad} J(a) = \operatorname{grad} B(a) \cdot \psi(a) + B(a) \cdot \operatorname{grad} \psi(a) = B(a) \cdot \operatorname{grad} \psi(a) =$$

$$= B(a) \left(\frac{\partial \psi}{\partial \xi_1} \operatorname{grad} \xi_1(a) + \ldots + \frac{\partial \psi}{\partial \xi_n} \operatorname{grad} \xi_n(a) \right) =$$

$$= B(a) \left(\frac{\partial \psi}{\partial \xi_1} \operatorname{grad} y_1(a) + \ldots + \frac{\partial \psi}{\partial \xi_{n-1}} \operatorname{grad} y_{n-1}(a) + \frac{\partial \psi}{\partial \xi_n} e_n \right).$$

Si l'on avait $\frac{\partial \psi(a)}{\partial \xi_n} = 0$, alors grad $J(a)$ s'exprimerait linéairement par grad $y_1, \ldots,$ grad y_{n-1}, ce qui n'est pas par hypothèse. Donc, $\frac{\partial \psi(a)}{\partial \xi_n} \neq 0$. Il en résulte d'après le théorème sur la fonction implicite que l'équation $J(x) = 0$ ou, ce qui est le même, $\psi(\xi_1, \ldots, \xi_n) = 0$ est résoluble par rapport à ξ_n, de sorte que, en coordonnées $\xi_1, \ldots, \xi_n$, la surface $S = \{x: J(x) = 0\}$ peut être mise sous forme de $\xi_n = \omega(\xi_1, \ldots, \xi_{n-1})$, où ω est une fonction ayant les dérivées continues en toutes les variables indépendantes au voisi-

nage du point $\bar{a} = (a_1, \ldots, a_{n-1})$. En dehors de la surface S nous avons $J(x) \neq 0$; supposons, par exemple, que $J(x) > 0$ « au-dessus » de la surface S, i.e. pour $\xi_n > \omega(\xi_1, \ldots, \xi_{n-1})$; alors, puisque grad $J(a) \neq 0$, $J(x) < 0$ « au-dessous » de S, i.e. pour $\xi_n < \omega(\xi_1, \ldots, \xi_{n-1})$.

Soit $c = (c_1, \ldots, c_n)$ un point de la surface S, dans les limites des voisinages indiqués plus haut du point a. Le segment $l = \{\xi_1 = c_1, \ldots, \xi_{n-1} = c_{n-1}, \xi_n = c_n + t\}$ où $|t| \leqslant \varepsilon$ perce la surface S pour $t = 0$. L'image du segment l dans l'espace Y se trouve sur la droite $L = \{y_1 = c_1, \ldots, y_{n-1} = c_{n-1}, y_n = \varphi(c_1, \ldots, c_{n-1}, c_n + t)\}$. De plus, $\frac{\partial y_n}{\partial t} = \frac{\partial \varphi}{\partial \xi_n} = \frac{1}{B} J(x)$, de sorte que $\frac{\partial y_n}{\partial t} > 0$ pour $t > 0$ et $\frac{\partial y_n}{\partial t} < 0$ pour $t < 0$. Nous voyons que, lorsque t varie de ε à $-\varepsilon$, c'est-à-dire que le point $\xi(t) = \{c_1, \ldots, c_{n-1}, c_n + t\}$ parcourt le segment l « du haut vers le bas », le point correspondant $f(\xi(t))$ descend suivant la droite L jusqu'à la surface $f(S)$, pour $t \searrow 0$, et remonte suivant la même droite lorsque t continue de décroître. Ceci veut dire que *l'application $y = f(x)$, dans un voisinage du point a, est du type pli.* A l'heure actuelle, les singularités des applications dérivables (i.e. les points de jacobien nul) attirent une grande attention. Voir, par exemple, В. И. Арнольд, Особенности гладких отображений, УМН, т. XXIII, вып. 1(139) et «Особенности дифференцируемых отображений», «Мир», Москва, 1968.

§ 1. 7. Valeurs stationnaires de fonctions numériques

1.71. Extrémum.

a. Soit $y = f(x)$ une fonction numérique définie dans un domaine G d'un espace normé X. Un point intérieur $a \in G$ s'appelle *point de minimum local de la fonction $f(x)$* si l'inégalité $f(x) \geqslant f(a)$ a lieu partout dans un voisinage du point a. D'une manière analogue, un point intérieur $b \in G$ s'appelle *point de maximum local de la fonction $f(x)$* si l'inégalité $f(x) \leqslant f(b)$ a lieu partout dans un voisinage du point b. Les points de minimum local et de maximum local sont appelés points d'*extrémum local.*

Etant donné un point d'extrémum local, il est en même temps un point d'extrémum local le long de toute droite passant par ce point. Par conséquent, si la fonction $f(x)$ est dérivable, alors sa dérivée suivant n'importe quelle direction s'annule au point d'extrémum local (*1.27a*). En se rappelant l'expression *1.27* (1) de la dérivée suivant une direction on conclut que, quels que soient un point a d'extrémum local de la fonction $f(x)$ et un vecteur $h \in X$, l'égalité $f'(a)\,h = 0$ a lieu; en d'autres termes, l'opérateur $f'(a)$ devient opérateur nul en tout point d'extrémum local:

$$f'(a) = 0. \tag{1}$$

Les points a où la relation (1) est valable sont appelés *points stationnaires de la fonction* $f(x)$. En chacun d'eux la partie linéaire principale de l'accroissement de la fonction s'annule. (Par conséquent, l'accroissement de la fonction est un infiniment petit d'ordre supérieur par rapport à h.)

Ceci ne signifie pas encore, en général, qu'un extrémum se réalise en le point a, mais, en tout cas, les points extrémaux cherchés se trouvent parmi les points stationnaires. Une fois les points stationnaires trouvés, il faut examiner chacun d'eux pour voir s'il est extrémal.

b. Envisageons le cas $X = R_n$, $f(x) = f(x_1, \ldots, x_n)$. L'équation (1) est alors équivalente au système de n équations à n inconnues $a_1, \ldots, a_n$:

$$\frac{\partial f(a_1, \ldots, a_n)}{\partial x_1} = 0,$$

$$\ldots\ldots\ldots\ldots$$

$$\frac{\partial f(a_1, \ldots, a_n)}{\partial x_n} = 0,$$

et la recherche des points stationnaires se réduit à la résolution de ce système.

c. Exemple. Les points stationnaires de la fonction $f(x_1, x_2) = 3x_1x_2 - x_1^3 - x_2^3$ ($R_2 \to R_1$) vérifient le système d'équations:

$$\frac{\partial f}{\partial x_1} \equiv 3x_2 - 3x_1^2 = 0, \quad \frac{\partial f}{\partial x_2} \equiv 3x_1 - 3x_2^2 = 0.$$

Ce système possède deux solutions:

$$a_1^{(1)} = a_2^{(1)} = 0,$$

$$a_1^{(2)} = a_2^{(2)} = 1.$$

Le premier point $a_1^{(1)} = a_2^{(1)} = 0$ n'est pas un point d'extrémum; qui plus est, il n'y a pas d'extrémum sur la droite $x_2 = 0$, car la fonction $f(x_1, x_2)$ a sur cette droite la forme $f(x_1, 0) = -x_1^3$.

Pour le deuxième point $a_1^{(2)} = a_2^{(2)} = 1$, on a $f(x_1, x_2) = 1$. En remplaçant x_1 par $1 + t_1$, x_2 par $1 + t_2$, on obtient

$$f(x_1, x_2) - 1 = 3t_1t_2 - 3t_1^2 - 3t_2^2 - t_1^3 - t_2^3.$$

Or, l'expression $t_1t_2 - t_1^2 - t_2^2 = -\left(t_1 - \frac{1}{2}\, t_2\right)^2 - \frac{3}{4}\, t_2^2 < 0$ pour tous les $t = \{t_1, t_2\} \neq 0$, et les termes du troisième ordre n'y peuvent rien changer pour $|t|$ suffisamment petit. Par conséquent, le point (1, 1) est un point d'extrémum, à savoir de maximum local.

Il est à signaler que la méthode « élémentaire » de l'analyse des points stationnaires que nous venons d'appliquer est loin d'être toujours possible.

Dans le chapitre suivant (*2.15*) on verra des règles plus systématisées de l'analyse des points stationnaires d'une fonction d'un nombre fini de variables.

1.72. Extrémum lié.

a. Définitions. Pour des fonctions numériques d'une variable multidimensionnelle $x \in G \subset X$, un nouveau type de problèmes extrémaux s'impose, à savoir les problèmes d'extrémum lié. En voici la position. On se donne, comme plus haut, une fonction numérique dérivable $y = f(x)$ $(G \subset X \to R_1)$. En outre, on considère un nouvel espace normé Z et une fonction vectorielle dérivable $\varphi(x): G \to Z$; fixons l'une des valeurs $C \in Z$ que cette fonction prend dans le domaine G. La condition

$$\varphi(x) = C \qquad (1)$$

détermine dans G une « surface » P. Un point $a \in G$ s'appelle *point de minimum local lié de la fonction* $f(x)$ *pour la condition* (1), si $\varphi(a) = C$ et si l'inégalité $f(x) \geqslant f(a)$ est valable pour tout point x satisfaisant à la condition (1) et appartenant à un voisinage de a. Autrement dit, un point a de la « surface » (1) est point de minimum local lié de la fonction $f(x)$ si l'inégalité $f(x) \geqslant f(a)$ a lieu pour tous les points de cette « surface » suffisamment proches du point a. Soulignons qu'il n'est nullement exigé que l'inégalité $f(x) \geqslant f(a)$ soit satisfaite pour les points x n'appartenant pas à la « surface » (1) qu'ils soient proches ou non de a.

Un point de *maximum lié* est défini de façon analogue en remplaçant $\geqslant$ par $\leqslant$.

Les points de minimum ou de maximum lié sont appelés *points d'extrémum lié.*

Dans ce qui suit on trouvera une condition nécessaire d'extrémum lié. Supposons que le point considéré a soit un point ordinaire de la surface $\varphi(x) = C$ (*1.62*), c'est-à-dire qu'il existe un sous-espace $X_1 \subset X$ tel que l'opérateur $\frac{\partial \varphi(a)}{\partial x_1}$ $(X_1 \to Z)$ soit inversible. Alors, comme on l'a vu plus haut (*1.61*), l'opérateur $\frac{\partial \varphi(a)}{\partial x_2}$ est nul sur un sous-espace $X_2 \subset X$ qui donne tout le X en somme droite avec X_1.

b. Lemme. *Etant donné un point ordinaire* a *d'extrémum lié de la fonction* $f(x)$, *on a* $f'(a)\, h_2 = 0$ *quel que soit* $h_2 \in X_2$.

Démonstration. Supposons que $f'(a)\, h_2 \neq 0$ pour un $h_2 \in X_2$. Montrons alors que, pour tout $\alpha \in R_1$ suffisamment petit, on peut trouver un $h_1 \in X_1$ tel que le point $a + \alpha h_2 + h_1$ appartienne toujours à la « surface » (1), c'est-à-dire que

$$\varphi(a + \alpha h_2 + h_1) = C, \qquad (2)$$

h_1 étant infiniment petit par rapport à α. Considérons la fonction suivante de α et h_1:

$$\Phi(\alpha, h_1) \equiv \varphi(a + \alpha h_2 + h_1) - C \quad (R_1 \times X_1 \to Z).$$

Elle s'annule pour $\alpha = 0$, $h_1 = 0$. L'opérateur

$$\frac{\partial \Phi(0, 0)}{\partial h_1} = \varphi'(a) \frac{\partial x}{\partial h_1}\Big|_{x=a} = \frac{\partial \varphi(a)}{\partial x_1}$$

est inversible par hypothèse. En appliquant le théorème sur la fonction implicite *1.53* nous avons la possibilité de résoudre l'équation (2) par rapport à h_1, ce qui fournit l'expression

$$h_1 = \psi(\alpha) \quad (R_1 \to X_1)$$

avec la fonction $\psi(\alpha)$ dérivable (*1.55*). De plus, on a

$$\psi'(0) = -\left[\frac{\partial \varphi(a)}{\partial x_1}\right]^{-1} \frac{\partial \Phi(0, 0)}{\partial \alpha} = -\left[\frac{\partial \varphi(a)}{\partial x_1}\right]^{-1} \varphi'(a) h_2 = 0,$$

puisque $h_2 \in X_2$ et $\varphi'(a) h_2 = 0$. Par conséquent, $h_1 = \psi(\alpha)$ est un infiniment petit par rapport à α.

Considérons ensuite $f(a + \alpha h_2 + h_1)$ où h_1 représente la quantité déjà trouvée $\psi(\alpha)$. Nous avons

$$f(a + \alpha h_2 + h_1) - f(a) = f'(a)(\alpha h_2 + h_1) + o(\alpha h_2 + h_1) =$$
$$= \alpha f'(a) h_2 + f'(a) h_1 + o(\alpha h_2 + h_1).$$

Le premier terme à droite est une fonction numérique linéaire de α, avec $f'(a) h_2 \neq 0$ pour coefficient angulaire. Le deuxième et le troisième termes sont infiniment petits par rapport à α avec h_1. Mais alors la fonction $f(x)$ n'a pas d'extrémum lié pour $\alpha = 0$, $h_1 = h_1(\alpha) = 0$: le point $x = a + \alpha h_2 + h_1$ appartient à la « surface » (1), est arbitrairement proche du point a, et la différence $f(x) - f(a)$ change de signe avec α.

Définitivement, $f'(a) h_2 = 0$ dès que $h_2 \in X_2$. Le lemme est démontré.

D'une façon générale, nous appellerons un point ordinaire a de la surface (1) *point stationnaire lié de la fonction* $f(x)$ *pour la liaison* $\varphi(x) = C$ si $f'(a) h_2 = 0$, quel que soit $h_2 \in X_2$ où X_2 est le sous-espace nul de l'opérateur $\varphi'(a)$. Tout point d'extrémum lié de la fonction $f(x)$ en est un point stationnaire lié ; mais un point stationnaire lié n'est pas forcément un point d'extrémum lié (de même que pour l'extrémum non lié).

c. Il est maintenant possible de formuler une condition nécessaire d'un point stationnaire lié (donc aussi d'un point d'extrémum local lié) :

T h é o r è m e. *Si* a *est un point stationnaire lié d'une fonction* $f(x) : G \subset X \to R_1$ *pour la liaison* (1), *alors il existe une fonctionnelle linéaire continue* $\lambda(z)$ *définie sur l'espace* Z *et telle que, pour tout* $h \in X$, on a

$$f'(a) h = \lambda[\varphi'(a) h]. \tag{3}$$

D é m o n s t r a t i o n. Définissons la fonctionnelle $\lambda(z)$ en utilisant la formule (3) et l'inversibilité de l'opérateur $\frac{\partial \varphi(a)}{\partial x_1}$:

$$\lambda(z) = f'(a)\left[\frac{\partial \varphi(a)}{\partial x_1}\right]^{-1} z. \tag{4}$$

La continuité de la fonctionnelle $\lambda(z)$ résulte de celle des opérateurs $\left[\frac{\partial\varphi(a)}{\partial x_1}\right]^{-1}$ et $f'(a)$. D'après le lemme et vu que $h = h_1 + h_2$ où $h_1 \in X_1$, $h_2 \in X_2$, nous avons $\varphi'(a)\,h_2 = 0$, $f'(a)\,h_2 = 0$; il suffit d'établir la relation (3) pour les vecteurs $h_1 \in X_1$; or, pour $h = h_1$ elle découle immédiatement de (4).

d. Le théorème *c* fournit en même temps un moyen de rechercher les points stationnaires liés. En effet, considérons une fonctionnelle linéaire $\lambda(z)$ $(Z \to R_1)$, indéterminée pour le moment, et composons la fonction

$$F(x) = f(x) - \lambda[\varphi(x)].$$

Au point stationnaire lié a que l'on cherche, nous avons l'équation (3):

$$f'(a) - \lambda[\varphi'(a)] = 0,$$

ce qui exprime le fait que a est un point stationnaire (dans tout le domaine G) de la fonctionnelle $F(x)$. Par cela, le problème d'extrémum lié se réduit à la recherche de points stationnaires ordinaires d'une autre fonction avec la fonctionnelle inconnue $\lambda(z)$.

Tous les points d'extrémum lié se trouvent parmi les points stationnaires liés; leur séparation exige, de même qu'en cas d'extrémum non lié, une étude spéciale de chaque point stationnaire lié.

1.73. Cas d'une liaison numérique dans un espace n-dimensionnel.

a. Soit $X = R_n$, $y = f(x)\colon R_n \to R_1$, $z = \varphi(x)\colon R_n \to R_1$. La fonctionnelle $\lambda(z)\colon R_1 \to R_1$ est tout simplement la multiplication par un nombre (inconnu au début) λ. La résolution du problème d'extrémum lié de la fonction $f(x)$ avec la condition

$$\varphi(x) = C \in R_1 \tag{1}$$

se ramène à la recherche des points stationnaires pour la fonction numérique

$$F(x) = f(x) - \lambda\varphi(x).$$

Afin de résoudre ce dernier problème, nous écrivons l'équation

$$F'(x) = f'(x) - \lambda\varphi'(x) = 0,$$

soit, en coordonnées,

$$\left.\begin{aligned} &\frac{\partial f(x_1, \ldots, x_n)}{\partial x_1} - \lambda\frac{\partial\varphi(x_1, \ldots, x_n)}{\partial x_1} = 0,\\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ &\frac{\partial f(x_1, \ldots, x_n)}{\partial x_n} - \lambda\frac{\partial\varphi(x_1, \ldots, x_n)}{\partial x_n} = 0. \end{aligned}\right\} \tag{2}$$

Nous avons à déterminer de ce système de $n+1$ équations (1) et (2) les $n+1$ inconnues $x_1, \ldots, x_n, \lambda$.

b. Exemple. Trouvons les points stationnaires liés pour la fonction

$$f(x)=\sum_{i=1}^{n} b_i x_i^2 \quad (R_n \to R_1) \quad (0<b_1<\ldots<b_n) \tag{3}$$

sur la sphère

$$\varphi(x)=\sum_{i=1}^{n} x_i^2-1=0. \tag{4}$$

Construisons la fonction

$$F(x)=f(x)-\lambda\varphi(x)=\sum_{i=1}^{n} b_i x_i^2-\lambda\sum_{i=1}^{n} x_i^2+\lambda \tag{5}$$

dont nous avons à trouver les points stationnaires ordinaires; pour le faire, annulons ses dérivées premières:

$$\frac{\partial F}{\partial x_j}=2b_j x_j-2\lambda x_j=0 \quad (j=1, \ldots, n). \tag{6}$$

Si le nombre λ est distinct de tous les nombres $b_1, \ldots, b_n$, il résulte des équations (6) que $x_1=\ldots=x_n=0$, ce qui contredit la condition (4). Supposons donc que $\lambda=b_m$, pour un $m=1, 2, \ldots, n$. Nous trouvons alors $x_j=0$ $(j\neq m)$; comme x appartient à la sphère (4), on a $x_m=\pm 1$. Ainsi, les points stationnaires liés de la fonction $f(x)$ sur la sphère (4) sont $\pm e_m=\pm(0, \ldots, 1, \ldots, 0)$. Le minimum lié se réalise aux points $\pm e_1$, le maximum lié aux points $\pm e_n$. Les autres points stationnaires liés ne sont pas extrémaux; en effet, lorsqu'on se déplace en partant d'un point e_m $(1<m<n)$ le long de la circonférence dans le plan des variables x_{m-1}, x_m, la fonction $f(x)$ diminue; si, par contre, on se déplace le long de la circonférence dans le plan x_m, x_{m+1}, alors $f(x)$ croît. Le cas $n=2$ est représenté fig. 1.7-1.

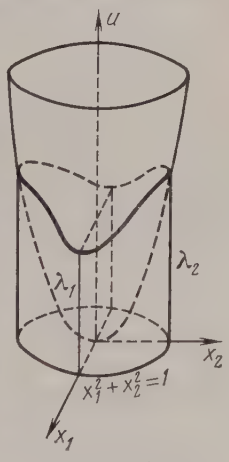


Fig. 1.7-1

1.74. Cas de k liaisons numériques dans un espace n-dimensionnel.

a. Soit

$$X=R_n,\ y=f(x): G\subset R_n \to R_1,\ z=\varphi(x): G\to R_k.$$

La liaison $\varphi(x)=C$ s'écrit donc sous forme de k relations numériques:

$$\left.\begin{aligned} &\varphi_1(x_1, \ldots, x_n)=C_1,\\ &\ldots\ldots\ldots\ldots\\ &\varphi_k(x_1, \ldots, x_n)=C_k. \end{aligned}\right\} \tag{1}$$

La fonctionnelle linéaire $\lambda(z): R_k \to R_1$ se définit en donnant k nombres $\lambda_1, \ldots, \lambda_k$ et agit sur un vecteur $z=\{z_1, \ldots, z_k\}\in R_k$ d'après la formule

$$\lambda(z)=\lambda_1 z_1+\ldots+\lambda_k z_k.$$

La résolution du problème d'extrémum lié se réduit à la recherche des points stationnaires de la fonction

$$F(x) \equiv f(x) - \lambda[\varphi(x)] = f(x_1, \ldots, x_n) - \sum_{i=1}^{k} \lambda_i \varphi_i(x_1, \ldots, x_n).$$

Les nombres λ_i sont appelés *multiplicateurs de Lagrange*. Nous avons à résoudre l'équation

$$F'(x) \equiv f'(x) - \lambda\varphi'(x) = 0,$$

soit, en coordonnées, le système d'équations

$$\left.\begin{aligned} &\frac{\partial f(x_1, \ldots, x_n)}{\partial x_1} - \sum_{i=1}^{k} \lambda_i \frac{\partial \varphi_i(x_1, \ldots, x_n)}{\partial x_1} = 0, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &\frac{\partial f(x_1, \ldots, x_n)}{\partial x_n} - \sum_{i=1}^{k} \lambda_i \frac{\partial \varphi_i(x_1, \ldots, x_n)}{\partial x_n} = 0. \end{aligned}\right\} \quad (2)$$

Le problème est donc réduit à la résolution du système de $k + n$ équations (1) et (2) aux inconnues $x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_k$.

b. Exemple. Sur la courbe $y = x^2$, $z = x^2$ dans l'espace $R_3(x, y, z)$ trouver le point $a = (x, y, z)$ le plus proche du point $c = (0, 0, 1)$.

Nous avons à trouver le minimum de la fonction

$$|c - a|^2 = x^2 + y^2 + (z - 1)^2$$

sous les conditions

$$\varphi_1(x, y, z) \equiv y - x^2 = 0,$$
$$\varphi_2(x, y, z) \equiv z - x^2 = 0.$$

Il faut donc construire la fonction

$$F(x, y, z) \equiv |c - a|^2 - \lambda_1\varphi_1 - \lambda_2\varphi_2 =$$
$$= x^2 + y^2 + (z - 1)^2 - \lambda_1(y - x^2) - \lambda_2(z - x^2)$$

et annuler ses dérivées partielles en x, y, z:

$$\frac{\partial F}{\partial x} \equiv 2x + 2\lambda_1 x + 2\lambda_2 x = 0,$$

$$\frac{\partial F}{\partial y} \equiv 2y - \lambda_1 = 0,$$

$$\frac{\partial F}{\partial z} \equiv 2(z - 1) - \lambda_2 = 0.$$

On en tire $\lambda_1 = 2y$, $\lambda_2 = 2(z - 1)$; il reste à résoudre le système suivant :

$$y - x^2 = 0,$$
$$z - x^2 = 0,$$
$$2x + 4xy + 4(z - 1)x = 0.$$

L'une des solutions est immédiate: $x = y = z = 0$. Et si $x \neq 0$, alors on peut remplacer la troisième équation par

$$1 + 2y + 2(z - 1) = 0.$$

Les deux premières fournissent $y = z$, de sorte que

$$4y = 1,\ y = \frac{1}{4},\ z = \frac{1}{4},\ x = \pm\frac{1}{2}.$$

Le carré de la distance du point $c = (0, 0, 1)$ à chacun des points trouvés a et b vaut $\frac{1}{4} + \frac{1}{16} + \frac{9}{16} < 1$, c'est donc en ces points que le minimum cherché se réalise; le point (0, 0, 0) correspond au maximum local de la distance (fig. 1.7-2).

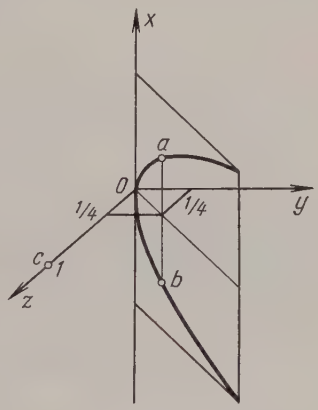


Fig. 1.7-2

§ 1.8. Equations différentielles (théorèmes locaux)

Le théorème fondamental de l'existence et de l'unicité de solution d'une équation différentielle ordinaire $y'(t) = \Phi(t, y)$, avec la condition initiale $y(t_0) = y_0$ et dans certaines suppositions sur la fonction $\Phi(t, y)$, a déjà été démontré dans *U13.35*. Nous revenons ici à ce théorème afin de préciser la dépendance de la solution des paramètres éventuels. Le théorème *1.53* sur la fonction implicite et le théorème *1.55* sur la dérivée d'une fonction implicite nous serviront de moyen principal de l'étude.

1.81. Pour les besoins de la démonstration du théorème sur la continuité de la solution par rapport au paramètre que l'on va considérer, il suffit de disposer le principe du point fixe *U13.22*.

a. Principe du point fixe. Soient M un espace métrique complet et $A : M \to M$ une application qui vérifie la condition

$$\rho(A(u), A(v)) \leqslant \theta\rho(u, v) \text{ quels que soient } u, v \in M$$

avec une « constante de contraction » θ, $0 \leqslant \theta < 1$ fixe.

Alors il existe (et est unique) un point fixe de l'application A, i.e. un point $u_A \in M$ tel que

$$A(u_A) = u_A.$$

Pour deux applications $A(u)$ et $B(u)$ avec une même constante de contraction θ, la distance entre leurs points fixes est estimée par l'inégalité

$$\rho(u_A, u_B) \leqslant \frac{1}{1-\theta} \sup_{u \in M} \rho[A(u), B(u)].$$

En conséquence nous avons: si une application $A = A_\lambda\,(M \to M)$ dépend d'un paramètre λ parcourant un espace métrique Λ et si cette dépendance est continue pour $\lambda = \lambda_0$, à ce sens que

$$\lim_{\lambda \to \lambda_0} \sup_{u \in M} \rho\,[A_\lambda(u),\ A_{\lambda_0}(u)] = 0, \tag{1}$$

la constante de contraction θ étant la même pour tous les opérateurs A_λ, alors le point fixe u_λ de l'opérateur A_λ est continu en fonction de λ pour $\lambda = \lambda_0$, c'est-à-dire

$$\lim_{\lambda \to \lambda_0} u_\lambda = u_{\lambda_0}. \tag{2}$$

b. Soient Λ un espace métrique complet, Y un espace vectoriel normé complet; désignons par Λ_ρ (Y_r) la boule de rayon ρ (r) centrée à un point fixe $\lambda_0 \in \Lambda$ $(y_0 \in Y)$, par T_h l'intervalle $|\,t - t_0\,| \leqslant h$ de la droite réelle.

Théorème. *Soit* $\Phi\,(t,\ y,\ \lambda)\colon T_h \times Y_r \times \Lambda_\rho \to Y$ *une fonction continue majorée par une constante* C *et satisfaisant à la condition de Lipschitz* $|\,\Phi\,(t,\ y_1,\ \lambda) - \Phi\,(t,\ y_2,\ \lambda)\,| \leqslant L\,|\,y_1 - y_2\,|$; *soit ensuite* $\varphi\,(\lambda)\colon \Lambda_\rho \to Y$ *une fonction continue avec la valeur fixe* $\varphi\,(\lambda_0) = y_0$. *Alors il existe deux nombres* $\delta \leqslant h$, $\sigma \leqslant \rho$ *et une fonction* $y\,(t,\ \lambda)\colon T_\delta \times \Lambda_\sigma \to Y_r$ *continue en* λ *et dérivable en* t *qui satisfait à l'équation différentielle*

$$\frac{dy\,(t,\ \lambda)}{dt} = \Phi\,(t,\ y\,(t,\ \lambda),\ \lambda),\quad \{t,\ \lambda\} \in T_\delta \times \Lambda_\sigma \tag{3}$$

et à la condition initiale

$$y\,(t_0,\ \lambda) = \varphi\,(\lambda), \qquad \lambda \in \Lambda_\sigma. \tag{4}$$

Démonstration. Désignons par M_δ, pour $\delta \leqslant h$, l'espace métrique des fonctions continues $u\,(t)\colon T_\delta \to Y_r$, avec la métrique usuelle

$$\rho\,(u_1(t),\ u_2(t)) = \sup_{T_\delta} |\,u_1(t) - u_2(t)\,|.$$

L'espace M_δ est complet (*U12.23f*). Considérons sur M_δ l'opérateur

$$A_\lambda\,(u\,(t)) \equiv \varphi\,(\lambda) + \int_{t_0}^{t} \Phi\,(\tau,\ u\,(\tau),\ \lambda)\,d\tau. \tag{5}$$

Il transforme les fonctions $u\,(t) \in M_\delta$ en des fonctions $T_\delta \to Y$ dont les valeurs n'appartiennent en général pas à la boule Y_r. Néanmoins, pour toutes les valeurs $\lambda \in \Lambda_\sigma$, on a

$$|\,A_\lambda\,(u\,(t)) - y_0\,| \leqslant |\,A_\lambda\,(u\,(t)) - \varphi\,(\lambda)\,| + |\,\varphi\,(\lambda) - y_0\,| =$$

$$= \left|\int_{t_0}^{t} \Phi\,(\tau,\ u\,(\tau),\ \lambda)\,d\tau\right| + |\,\varphi\,(\lambda) - y_0\,| \leqslant C\delta + |\,\varphi\,(\lambda) - y_0\,|,$$

$$|A_\lambda(u_1(t)) - A_\lambda(u_2(t))| \leqslant \int_{t_0}^{t} |\Phi(\tau, u_1(\tau), \lambda) - \Phi(\tau, u_2(\tau), \lambda)|\, d\tau \leqslant$$

$$\leqslant L \sup_{T_\delta} |u_1(\tau) - u_2(\tau)|\, \delta,$$

si l'on choisit maintenant les nombres $\delta \leqslant h$ et $\sigma \leqslant \rho$ suffisamment petits pour avoir

$$C\delta \leqslant r/2,$$

$$|\varphi(\lambda) - y_0| \leqslant r/2 \quad \text{pour} \quad \lambda \in \Lambda_\sigma,$$

$$L\delta \leqslant \theta < 1,$$

alors on trouve

$$|A_\lambda(u(t)) - y_0| \leqslant r,$$

$$\rho[A_\lambda(u_1(t)),\ A_\lambda(u_2(t))] \leqslant \theta\rho[u_1(t),\ u_2(t)].$$

Ainsi, δ et σ étant choisis, l'opérateur $A_\lambda(u(t))$, pour tout $\lambda \in \Lambda_\sigma$ fixe, applique l'espace M_δ dans lui-même et est contractant. D'après a, l'opérateur A_λ possède dans l'espace M_δ un point fixe ; dans notre cas, cela est équivalent à l'existence d'une fonction $y(t, \lambda)$ telle que

$$y(t, \lambda) = \varphi(\lambda) + \int_{t_0}^{t} \Phi(\tau, y(\tau, \lambda), \lambda)\, d\tau. \tag{6}$$

Il suit de cette relation que la fonction $y(t, \lambda)$ est dérivable en t pour λ fixe (*U12.63a*) ; en dérivant nous voyons que la fonction $y(t, \lambda)$ satisfait à la relation (3). La substitution $t = t_0$ dans (6) conduit à la condition initiale (4).

Afin de prouver la continuité de la solution obtenue en fonction de λ, c'est-à-dire la validité de la relation

$$\lim_{\lambda \to \lambda_0} \sup_{t \in T_\delta} |y(t, \lambda) - y(t, \lambda_0)| = 0$$

(pour λ_0 fixe), il suffit de vérifier (1) pour l'opérateur A_λ. Soit $u(t) \in M_\delta$. Alors

$$A_\lambda(u(t)) - A_{\lambda_0}(u(t)) =$$

$$= \varphi(\lambda) - \varphi(\lambda_0) + \int_{t}^{t_0} [\Phi(\tau, u(\tau), \lambda) - \Phi(\tau, u(\tau), \lambda_0)]\, d\tau. \tag{7}$$

Etant donné $\varepsilon > 0$, trouvons $\eta > 0$ ($\eta \leqslant \sigma$) tel que

$$|\Phi(t, u, \lambda) - \Phi(t, u, \lambda_0)| < \varepsilon,$$

$$|\varphi(\lambda) - \varphi(\lambda_0)| < \varepsilon$$

pour $\rho(\lambda, \lambda_0) < \eta$ et pour tous les $t \in T_\delta$, $u \in Y_r$. Ceci fait, il

résulte de (7):

$$| A_\lambda (u(t)) - A_{\lambda_0} (u(t)) | \leqslant \varepsilon + \delta\varepsilon = \varepsilon (\delta + 1).$$

Comme cela est valable pour n'importe quelle fonction $u(t) \in M_\delta$, on a

$$\sup_{u(t) \in M_\delta} | A_\lambda (u(t)) - A_{\lambda_0} (u(t)) | \leqslant \varepsilon (\delta + 1),$$

d'où

$$\lim_{\lambda \to \lambda_0} \sup_{u(t) \in M_\delta} | A_\lambda (u(t)) - A_{\lambda_0} (u(t)) | = 0,$$

ce qu'il nous fallait; le théorème est complètement démontré.

1.82. La solution de l'équation *1.81* (1) avec la condition *1.81* (2) qu'on a trouvée dans *1.81* est *unique* au sens suivant. Soient $y_1 (t, \lambda)$: $: T_{\delta_1} \times \Lambda_{\sigma_1} \to Y_r$ et $y_2 (t, \lambda)$: $T_{\delta 2} \times \Lambda_{\sigma_2} \to Y_r$ deux solutions vérifiant la condition $y_1 (t_0, \lambda) = y_2 (t_0, \lambda) = \varphi (\lambda)$; alors, *pour* $\{t, \lambda\} \in \in T_\delta \times \Lambda_\sigma$ où $\delta = \min (\delta_1, \delta_2)$, $\sigma = \min (\sigma_1, \sigma_2)$ *on a* $y_1 (t, \lambda) \equiv \equiv y_2 (t, \lambda)$.

Pour le prouver, notons d'abord que l'identité *1.81* (3) valable pour tous les points $\{t, \lambda\} \in T_\delta \times \Lambda_\sigma$ est bien valable dans n'importe quel domaine plus petit $T_{\delta'} \times \Lambda_{\sigma'}$, $\delta' \leqslant \delta$, $\sigma' \leqslant \sigma$; c'est pourquoi la restriction de la fonction $y(t, \lambda)$ à ce dernier domaine est le point fixe de l'opérateur correspondant $A_\lambda \equiv A_\lambda^{(\delta')}$ dans l'espace $M_{\delta'}$. Lorsque $\delta = \min (\delta_1, \delta_2)$, l'opérateur $A_\lambda^{(\delta)}$ applique l'espace M_δ dans lui-même, donc, d'après ce qu'on a vu, y possède un point fixe $y^{(\delta)} (t, \lambda)$. Les restrictions à M_δ des fonctions $y_1 (t, \lambda)$ et $y_2 (t, \lambda)$ sont elles aussi des points fixes de l'opérateur $A_\lambda^{(\delta)}$; vu l'unicité du point fixe d'un opérateur contractant, on en tire l'égalité $y(t, \lambda) = = y_1 (t, \lambda) = y_2 (t, \lambda)$ ($\{t, \lambda\} \in T_\delta \times \Lambda_\sigma$), ce qu'il fallait.

1.83. Théorème. *Supposons que, en plus des hypothèses 1.81b, la fonction* $\Phi (t, y, \lambda)$ *satisfasse dans* $T_h \times Y_r \times \Lambda_\rho$ *à la condition de Lipschitz renforcée*

$$| \Phi (t, y_1, \lambda_1) - \Phi (t, y_2, \lambda_2) | \leqslant L | y_1 - y_2 | + B\rho (\lambda_1, \lambda_2),$$

et que la fonction $\varphi (\lambda)$ *satisfasse dans* Λ_ρ *à la condition de Lipschitz*

$$| \varphi (\lambda_1) - \varphi (\lambda_2) | \leqslant A\rho (\lambda_1, \lambda_2).$$

Dans ce cas, la solution $y (t, \lambda)$ *satisfait, pour* $\{t, \lambda\} \in T_\delta \times \Lambda_\sigma$, *à la condition de Lipschitz par rapport à* λ:

$$| y (t, \lambda_1) - y (t, \lambda_2) | \leqslant C\rho (\lambda_1, \lambda_2), \quad C \leqslant \frac{A + B\delta}{1 - L\delta}.$$

Démonstration. En posant dans *1.81* (6) $\lambda = \lambda_1$ et $\lambda = \lambda_2$ et soustrayant les égalités obtenues l'une de l'autre on a:

$$|y(t, \lambda_1) - y(t, \lambda_2)| \leqslant |\varphi(\lambda_1) - \varphi(\lambda_2)| +$$

$$+ \left| \int_{t_0}^{t} [\Phi(\tau, y(\tau, \lambda_1), \lambda_1) - \Phi(\tau, y(\tau, \lambda_2), \lambda_2)]\, d\tau \right| \leqslant$$

$$\leqslant A\rho(\lambda_1, \lambda_2) + \int_{t_0}^{t} [L|y(\tau, \lambda_1) - y(\tau, \lambda_2)| + B\rho(\lambda_1, \lambda_2)]\, d\tau \leqslant$$

$$\leqslant (A + B\delta)\rho(\lambda_1, \lambda_2) + L\delta \sup_{T_\delta} |y(t, \lambda_1) - y(t, \lambda_2)|,$$

d'où

$$\sup_{T_\delta} |y(t, \lambda_1) - y(t, \lambda_2)| \leqslant$$

$$\leqslant (A + B\delta)\rho(\lambda_1, \lambda_2) + L\delta \sup_{T_\delta} |y(t, \lambda_1) - y(t, \lambda_2)|;$$

par conséquent, si $L\delta < 1$, alors

$$\sup_{T_\delta} |y(t, \lambda_1) - y(t, \lambda_2)| \leqslant \frac{A + B\delta}{1 - L\delta} \rho(\lambda_1, \lambda_2),$$

ce qu'il fallait démontrer.

1.84. Dans ce qui suit on s'intéresse à la dérivabilité de la solution $y(t, \lambda)$ par rapport au paramètre λ; bien entendu, il faut d'abord supposer que l'espace Λ où le paramètre λ varie soit non seulement métrique mais aussi normé, et que la fonction $\Phi(t, y, \lambda)$ soit dérivable par rapport à λ. Nous commençons par renforcer de façon convenable le théorème sur le point fixe.

a. Soient X et Λ des espaces normés complets, $M \subset X$ et $\Delta \subset \Lambda$ des boules fermées. Soit $A_\lambda(u): M \times \Delta \to M$ une application qui, pour tout $\lambda \in \Delta$ admet un seul point fixe $u = y(\lambda) \in M$, de sorte que $A_\lambda(y) = y$. Supposons que l'application $A_\lambda(u)$ soit dérivable en u au point $u = y(\lambda)$ et que l'opérateur $E - \frac{\partial A_\lambda(y)}{\partial u}$ soit inversible.

Théorème. *Dans l'hypothèse formulée, la dérivabilité de la fonction $A_\lambda(u)$ par rapport à l'espace $X \times \Lambda$ implique la dérivabilité de la fonction $y(\lambda)$.*

Démonstration. Fixons $\lambda = \lambda_1 \in \Delta$. L'équation

$$\Phi(u, \lambda) \equiv u - A_\lambda(u) = 0$$

est satisfaite par hypothèse pour $\lambda = \lambda_1$, $u = y(\lambda_1)$. La fonction $\Phi(u, \lambda)$ a la dérivée en u:

$$\frac{\partial \Phi(u, \lambda)}{\partial u} = E - \frac{\partial A_\lambda(u)}{\partial u}$$

qui est un opérateur inversible. Nous nous trouvons dans l'hypothèse du théorème sur la fonction implicite *1.53*; en l'appliquant nous voyons qu'il existe une solution $u = u(\lambda)$ de l'équation (2) qui coïncide avec $y(\lambda_1)$ pour $\lambda = \lambda_1$. Le point fixe étant unique, nous avons $y(\lambda) \equiv u(\lambda)$. La fonction $y(\lambda)$ est continue d'après *1.53*; la dérivabilité de $A_\lambda(u)$ implique celle de $y(\lambda)$ (*1.55*), ce qu'il nous fallait.

Si les dérivées $\frac{\partial A_\lambda(u)}{\partial u}$ et $\frac{\partial A_\lambda(u)}{\partial \lambda}$ sont continues, $y'(\lambda)$ l'est aussi d'après *1.55b*.

b. Signalons que l'inégalité $\left\|\frac{\partial A_\lambda(u)}{\partial u}\right\| < 1$ représente une condition suffisante de l'inversibilité de l'opérateur $E - \frac{\partial A_\lambda(y)}{\partial u}$ (*U12.82*).

1.85. Appliquons la proposition démontrée à l'équation différentielle ordinaire *1.81* (3) contenant un paramètre λ:

$$\frac{dy(t, \lambda)}{dt} = \Phi(t, y(t, \lambda), \lambda) \tag{1}$$

avec la condition initiale

$$y(t_0, \lambda) = \varphi(\lambda). \tag{2}$$

En ce qui concerne la fonction $\Phi(t, y, \lambda)$, on suppose qu'elle soit définie et continue dans un domaine $T_h \times Y_r \times \Lambda_\rho$ (comme dans *1.81b*), qu'elle satisfasse à la condition de Lipschitz en y

$$|\Phi(t, y_1, \lambda) - \Phi(t, y_2, \lambda)| \leqslant L |y_1 - y_2|$$

et que ses valeurs appartiennent à la boule Y_r de l'espace normé Y; la fonction $\varphi(\lambda)$ est supposée continue dans Λ_ρ.

Nous avons vu dans *1.81b* que, dans l'espace $M = M_\delta$ de toutes les fonctions continues $u = u(t)$: $T_\delta \to Y_r$ (qui est une boule fermée de l'espace $Y(T_\delta)$ de *toutes* les fonctions continues $u(t)$: $T_\delta \to Y$), l'opérateur

$$A_\lambda(u) \equiv \varphi(\lambda) + \int_{t_0}^{t} \Phi(\tau, u(\tau), \lambda)\, d\tau \tag{3}$$

possède un seul point fixe $y = y(t, \lambda)$, du moment que δ soit suffisamment petit; ce point fixe représente la solution de l'équation (1) sous la condition (2). Si l'opérateur est une fonction dérivable en u et λ, alors, d'après *1.84a*, la solution $y(t, \lambda)$ sera une fonction dérivable en λ. Il reste donc d'établir les conditions sous lesquelles l'opérateur (3) soit dérivable en u et λ. Montrons que l'on peut choisir ces conditions, par exemple, comme suit:

(1) La fonction $\Phi(t, y, \lambda): T_h \times Y_r \times \Lambda_\rho \to Y$ a la dérivée continue en u;

(2) La fonction $\Phi(t, y, \lambda)$ a la dérivée continue en λ;

(3) La fonction $\varphi(\lambda): \Lambda_\rho \to Y$ a la dérivée continue en λ.

Théorème. *Dans les hypothèses 1.81b* et *(1)-(3)*, *la solution* $y(t, \lambda)$ *a la dérivée continue en* λ.

Démonstration. En vertu de *1.48a-b* et de notre hypothèse, l'opérateur $A_\lambda(u)$ a la dérivée continue en u, notamment

$$\frac{\partial A_\lambda(u)}{\partial u} = \int_{t_0}^{t} \frac{\partial \Phi(\tau, u(\tau), \lambda)}{\partial u}\, d\tau.$$

Il est clair que, pour δ suffisamment petit, $\left\| \frac{\partial A_\lambda(u)}{\partial u} \right\| \leqslant$

$$\leqslant \sup_{\tau, u, \lambda} \left\| \frac{\partial \Phi(\tau, u, \lambda}{\partial u} \right\| \cdot \delta < 1.$$

En vertu de *1.48d*, l'opérateur possède la dérivée continue en λ:

$$\frac{\partial A_\lambda(u)}{\partial \lambda} = \varphi'(\lambda) + \int_{t_0}^{t} \frac{\partial \Phi(\tau, u(\tau), \lambda)}{\partial \lambda}\, d\tau.$$

D'après *1.47a-b*, l'opérateur $A_\lambda(u)$ est dérivable dans $\Lambda_\rho \times M_\delta$ par rapport à tout l'espace $\Lambda \times Y(T_\delta)$. En appliquant *1.84a* nous voyons que la solution $y(t, \lambda)$ a la dérivée, même continue, en λ. Le théorème est démontré.

1.86. a. La dérivée en λ de la solution $y(t, \lambda)$ de l'équation *1.85* (1) avec la condition *1.85* (2) peut être trouvée en dérivant directement l'équation opératorielle *1.81* (6):

$$y(t, \lambda) = \varphi(\lambda) + \int_{t_0}^{t} \Phi(\tau, y(\tau, \lambda), \lambda)\, d\tau.$$

On obtient alors:

$$\frac{\partial y(t, \lambda)}{\partial \lambda} = \varphi'(\lambda) + \int_{t_0}^{t} \left[\frac{\partial \Phi(\tau, y(\tau, \lambda), \lambda)}{\partial y} \frac{\partial y(\tau, \lambda)}{\partial \lambda} + \frac{\partial \Phi(\tau, y(\tau, \lambda), \lambda)}{\partial \lambda} \right] d\tau. \tag{1}$$

Il en résulte que *la fonction* $z(t, \lambda) = \frac{\partial y(t, \lambda)}{\partial \lambda}$ *satisfait à l'équation linéaire différentielle avec le paramètre* λ

$$\frac{dz(t, \lambda)}{dt} = \frac{\partial \Phi(t, y(t, \lambda), \lambda)}{\partial y} z(t, \lambda) + \frac{\partial \Phi(t, y(t, \lambda), \lambda)}{\partial \lambda} \tag{2}$$

pour la condition initiale

$$z(t_0, \lambda) = \varphi'(\lambda). \tag{3}$$

b. Considérons des cas particuliers les plus simples.

Supposons que $\Lambda = Y$, $\varphi(\lambda) = \lambda$ et que $\Phi(t, y, \lambda) \equiv \Phi(t, y)$ ne dépende pas de λ. Il s'agit dans ce cas de la dérivabilité de la solution $y(t, \lambda)$ par rapport à sa valeur initiale λ. Nous voyons que, pour une fonction $\Phi(t, y)$ ayant les dérivées continues en t et y, *la solution* $y(t, \lambda)$ *a la dérivée continue par rapport à* λ. La formule *1.86* (1) donne:

$$\frac{\partial y(t, \lambda)}{\partial \lambda} = E + \int_{t_0}^{t} \frac{\partial \Phi(\tau, y(\tau, \lambda))}{\partial y} \cdot \frac{\partial y(\tau, \lambda)}{\partial \lambda} d\tau.$$

En particulier, pour δ assez petits, $|t - t_0| < \delta$, l'opérateur $\frac{\partial y(t, \lambda)}{\partial \lambda}$ s'avère inversible. Il résulte de ce fait et de *1.56a-b* que, pour $t - t_0$ petits, *l'application* $y(t, \lambda)$ *est un difféomorphisme d'un voisinage* $U \subset Y$ *du point* λ *sur un voisinage* V *du point* $y(t, \lambda)$; nous voyons, par exemple, que le point $y(t, \lambda)$ admet un voisinage, notamment V, par tout point duquel passe une courbe intégrale partant du voisinage U. A un déplacement $\Delta\lambda$ du point initial λ dans le voisinage U, il correspond le déplacement

$$\frac{\partial y(t, \lambda)}{\partial \lambda} \Delta\lambda + o(\Delta\lambda) = \Delta\lambda + \int_{t_0}^{t} \frac{\partial \Phi(\tau, y(\tau, \lambda))}{\partial y} \frac{\partial y(\tau, \lambda)}{\partial \lambda} d\tau + o(\Delta\lambda) =$$

$$= \Delta\lambda + \frac{\partial \Phi(t_0, \lambda)}{\partial y} \Delta\lambda \Delta t + o(\Delta t)\, \Delta\lambda + o(\Delta\lambda)$$

de la solution $y(t, \lambda)$. Ce dernier déplacement, à des infiniment petits d'ordre supérieur près, coïncide avec $\Delta\lambda$ augmenté de $\frac{\partial \Phi(t_0, \lambda)}{\partial y} \Delta\lambda\, \Delta t$. La dérivée $\frac{\partial y(t, \lambda)}{\partial \lambda} = z(t, \lambda)$ satisfait à l'équation linéaire homogène qui se déduit de (2):

$$\frac{dz(t, \lambda)}{dt} = \frac{\partial \Phi(t, y(t, \lambda))}{\partial y} z(t, \lambda) \tag{4}$$

avec la condition initiale

$$z(t_0, \lambda) = E. \tag{5}$$

Si $z(t, \lambda)$ est une fonction numérique, le problème (4)-(5) admet la solution immédiate:

$$z(t, \lambda) = e^{\int_{t_0}^{t} \frac{\partial \Phi(\tau, y(\tau, \lambda))}{\partial y} d\tau}. \tag{3}$$

c. Supposons que $\varphi(\lambda) = y_0$ ne dépende pas de λ, de sorte que le paramètre λ ne figure que dans le second membre de l'équation *1.85* (1). Ceci étant, la fonction $z(t, \lambda) = \frac{\partial y(t, \lambda)}{\partial \lambda}$ satisfait à l'équation

1.86 (2) avec la condition initiale

$$z(t_0, \lambda) = 0. \tag{4}$$

1.87. Système dynamique local. Toute solution $y = y(t)$ d'une équation $\frac{dy}{dt} = \Phi(t, y)$ est naturellement interprétée comme courbe dans l'espace $T \times Y$. Mais on peut également lui faire correspondre une courbe dans l'espace Y en considérant t comme paramètre. Cette dernière courbe $y(t)$ se déduit de la courbe $\{t, y(t)\} \in T \times Y$ en la projetant dans l'espace Y. Si l'on interprète la variable t comme temps, alors la solution $y = y(t)$ représente la loi du mouvement d'un point mobile dans l'espace Y, à la vitesse $\Phi(t, y)$. Pour cette raison, l'ensemble des solutions $y = y(t)$ s'appelle *système dynamique*, ou bien, si l'on considère la situation locale, *système dynamique local*. Le problème devient beaucoup plus simple si $\Phi(t, y)$ ne dépend pas de t, de sorte que l'équation initiale a la forme

$$\frac{dy}{dt} = \Phi(y) \quad (G \subset Y \to Y). \tag{1}$$

Dans ce qui suit on suppose que la fonction $\Phi(y)$ ait la dérivée continue. Dans le cas considéré, on peut concevoir le système dynamique local comme mouvement d'un milieu continu dont la vitesse à tout point est donnée par le vecteur $\Phi(y)$ et ne dépend pas du temps. Les courbes dans l'espace Y qui correspondent aux solutions $y = y(t)$ s'appellent *trajectoires* du système. Il découle de *1.81-1.82* que deux trajectoires ou bien n'ont pas de point commun ou bien coïncident entièrement *. Si, pour un y_0, nous avons $\Phi(y_0) = 0$, alors $y \equiv y_0$ (const) est une solution immédiate de l'équation (1); la trajectoire correspondante dégénère en un seul point. Soit $\Phi(y_0) = z_0 \neq 0$; pour la raison de continuité on a $\Phi(y) \neq 0$ dans un voisinage du point y_0; il n'y a aucun point immobile dans ce voisinage, tous les points se déplacent réellement avec le temps suivant leurs trajectoires. Un cas particulier du système de ce genre s'interprète comme mouvement des points d'un solide en translation uniforme. Il s'avère que *le cas général se réduit à celui-ci par un difféomorphisme dans l'espace Y qui transforme le système dynamique au voisinage du point y_0 en une famille de segments parallèles parcourus à une vitesse constante.* En effet, supposons que $y_0 = 0$ (ce qu'on peut toujours obtenir par une translation) et qu'il existe une fonctionnelle linéaire continue $f(y): Y \to R_1$ telle que $f(z_0) \neq 0$ (lorsque Y est un espace hilbertien, on peut poser $f(y) = (y, z_0)$; dans un espace de Banach l'existence d'une telle fonctionnelle est garantie par le théorème

* Si $y_1(t_1) = y_2(t_2) = p \in Y$, alors $y_1(t_1 + t) \equiv y_2(t_2 + t)$ puisque les deux membres en tant que fonctions de t satisfont à une même équation $\frac{dy}{dt} = \Phi(y)$ avec la condition initiale commune $y(0) = p$.

de Hahn-Banach, voir, par exemple, Г. Е. Шилов, Математический анализ, специальный курс, 2 изд., М., 1961, Дополнение, § 2, стр. 426). Considérons le sous-espace $H = \{h \in Y : f(h) = 0\}$; il est fermé et ne contient pas z_0. Soit, ensuite, L le sous-espace unidimensionnel engendré par le vecteur z_0; nous affirmons que tout l'espace Y est la somme directe de L et H. En effet, pour tout $y \in Y$ et pour $h = y - \frac{f(y)z_0}{f(z_0)}$, nous avons évidemment $f(h) = 0$, de sorte que $h \in H$ et $y = h + z$, où $z = \frac{f(y)z_0}{f(z_0)} \in L$, cette décomposition étant unique puisque L et H ont un seul point commun 0. Dans le sous-espace H considérons la boule $H_\rho = \{h \in H : \| h \| \leqslant \rho\}$, et dans le sous-espace L le segment $T_\delta = \{tz_0 : | t | \leqslant \delta\}$. Faisons correspondre à tout couple $(t, h) \in T_\delta \times H_\rho$ le point $y(t, h)$ qui est la solution de l'équation (1) pour la valeur initiale $y(0, h) = h$. Si δ et ρ sont suffisamment petits, la quantité $y(t, h)$ est définie d'après *1.81b*. Prouvons que l'application $(t, h) \to y(t, h)$ est le difféomorphisme cherché. Nous avons $y(t, h) = \eta(t, h) + \lambda(t, h) z_0$ où $\eta(t, h) \in H$, $\lambda(t, h) \in R_1$. Remarquons que la fonction $y(t, h)$ est continue par rapport à l'ensemble (t, h) (puisque, dans *1.81b*, on a démontré la continuité de la fonction $y(t, h)$ par rapport à h dans l'espace des fonctions $y(t, h)$ avec la sup-métrique en t); il en résulte la continuité par rapport à (t, h) des composantes $\eta(t, h)$ et $\lambda(t, h)$. La fonction $y(t, h)$ admet la dérivée en t qui vaut $\Phi(y)$ en vertu de l'équation (1); il en résulte que les fonctions $\eta(t, h)$ et $\lambda(t, h)$ ont elles aussi les dérivées par rapport à t qui sont égales aux composantes respectives de $\Phi(y)$; bien entendu, ces composantes sont continues avec Φ. Pour $t = 0$, $h = 0$, elles prennent, en tant que composantes du vecteur $\Phi(0) = z_0$, les valeurs 0 et 1 respectivement. La fonction $y(t, h)$ a, à son tour, la dérivée en h (*1.86b*) qui est continue par rapport à (t, h). Il découle de l'égalité $y(0, h) = h$ que $\frac{\partial \eta(0, h)}{\partial h} = \frac{\partial y(0, h)}{\partial h} = E_H$ (opérateur identique dans H). En vertu de *1.48*, la fonction $y(t, h)$ est dérivable par rapport à (t, h); sa dérivée s'écrit naturellement comme matrice:

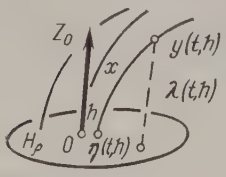


Fig. 1.8-1

$$\left\| \begin{matrix} \frac{\partial \lambda(t, h)}{\partial t} & \frac{\partial \lambda(t, h)}{\partial h} \\ \frac{\partial \eta(t, h)}{\partial t} & \frac{\partial \eta(t, h)}{\partial h} \end{matrix} \right\|.$$

Pour $t = 0$, $h = 0$, la matrice a la forme

$$\left\| \begin{matrix} 1 & 0 \\ \frac{\partial \eta(0, 0)}{\partial t} & E_H \end{matrix} \right\|$$

et s'avère inversible d'après *1.14i*. Il suit maintenant de *1.56b* que l'application $(t, h) \to y(t, h)$ représente un difféomorphisme d'un voisinage du zéro de l'espace $T \times H$ sur un voisinage du zéro de l'espace Y. Le difféomorphisme inverse π^{-1} transforme les trajectoires du système dynamique local construit d'après l'équation (1) en des segments « verticaux » auxquels t sert de coordonnée principale, de sorte que, du point de vue dynamique, ils sont parcourus à une vitesse constante (qui vaut 1); c'est bien cela qu'on affirmait.

1.88. Intégrales premières. **a.** Considérons l'ensemble des solutions de l'équation

$$\frac{dy}{dt} = \Phi(y) \qquad (1)$$

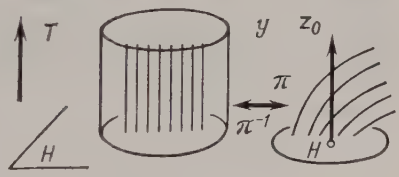


Fig. 1.8-2

dans un voisinage V d'un point $y_0 \in Y$ tel que $\Phi(y_0) \neq 0$. Une fonction numérique $z(y)$ définie, continue et possédant la dérivée continue $z'(y)$ dans V, s'appelle *intégrale première de l'équation* (1) (plus exactement, *intégrale première locale*) si elle est constante sur chaque trajectoire de cette équation. En utilisant le difféomorphisme π construit dans *1.87* il est possible d'obtenir une description avantageuse de toutes les intégrales premières de l'équation (1). Notamment, le difféomorphisme π^{-1} transforme chaque intégrale première $z(y)$ en une fonction $z(t, h): T_\delta \times H_\rho \to R_1$, qui est constante sur les segments verticaux de l'ensemble $T_\delta \times H_\rho$ et qui possède elle aussi la dérivée continue (par rapport à la variable (t, h)). Une telle fonction est déterminée d'une façon univoque par ses valeurs pour $t = 0$, $z(0, h) = \psi(h)$, où $\psi(h): H_\rho \to R_1$ est une fonction continûment dérivable par rapport à $h \in H$. Inversement, étant donnée une fonction continûment dérivable $\psi(h): H_\rho \to R_1$, la fonction $z(t, h) = \psi(h): T_\delta \times H_\rho \to R_1$ a la dérivée continue par rapport à la variable (t, h), et le difféomorphisme π la transforme en une fonction $z(y)$ continûment dérivable en y qui est constante sur les trajectoires, i.e. est une intégrale première de l'équation (1). En remarquant que le difféomorphisme π transforme l'ensemble H_ρ identiquement en lui-même, nous voyons que *toute intégrale première de l'équation* (1) *est bien déterminée par ses valeurs sur* H_ρ *qui forment une fonction continûment dérivable*; *ses valeurs sur les autres points y du voisinage V résultent du fait qu'une intégrale première est constante sur chaque trajectoire.*

b. Si $Y = R_n$, de sorte que l'on peut poser $H = R_{n-1}$, la situation s'éclaircit d'autant plus. Il s'avère, notamment, que, *dans un voisinage du point* $y_0 \in Y$ *avec* $\Phi(y_0) \neq 0$, *il existe* $n - 1$ *intégrales premières indépendantes* (*1.66b*) *de l'équation* (1); *d'autre part, toute intégrale première de cette équation représente, dans un voisinage du*

point y_0, une fonction (dérivable) de n'importe quelles $n-1$ intégrales premières indépendantes fixes. Pour démontrer la première partie de la proposition, choisissons dans H_ρ les coordonnées quelconques $h_1, \ldots, h_{n-1}$. Dans $T_\delta \times H_\rho$, les fonctions $z_1(t, h) = h_1, \ldots, z_{n-1}(t, h) = h_{n-1}$ sont constantes sur les segments verticaux et, évidemment, indépendantes; par conséquent, leurs images $z_1(y), \ldots, z_{n-1}(y)$ par le difféomorphisme π représentent $n-1$ intégrales premières indépendantes de l'équation (1). D'autre part, s'il y a les intégrales premières quelconques $z_1(y), \ldots, z_{n-1}(y)$ (indépendantes), le difféomorphisme π^{-1} les transforme en certaines fonctions $z_1(t, h), \ldots, z_{n-1}(t, h)$ qui sont constantes sur les segments verticaux et indépendantes elles aussi, de sorte que le rang de la matrice

$$\left\| \begin{array}{ccc} \frac{\partial z_1(0,0)}{\partial t} & \frac{\partial z_1(0,0)}{\partial h_1} & \cdots \quad \frac{\partial z_1(0,0)}{\partial h_{n-1}} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \frac{\partial z_{n-1}(0,0)}{\partial t} & \frac{\partial z_{n-1}(0,0)}{\partial h_1} & \cdots \quad \frac{\partial z_{n-1}(0,0)}{\partial h_{n-1}} \end{array} \right\|$$

vaut $n-1$. La première colonne de cette matrice étant composée des zéros, nous avons

$$\left| \begin{array}{ccc} \frac{\partial z_1(0,0)}{\partial h_1} & \cdots & \frac{\partial z_1(0,0)}{\partial h_{n-1}} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \frac{\partial z_{n-1}(0,0)}{\partial h_1} & \cdots & \frac{\partial z_{n-1}(0,0)}{\partial h_{n-1}} \end{array} \right| \neq 0.$$

Ainsi, les fonctions $z_1(0, h), \ldots, z_{n-1}(0, h)$ réalisent un difféomorphisme d'un voisinage $U(0) \subset H_\rho$ dans un domaine $V \subset R_{n-1}$. Par conséquent, toute fonction $\psi(h)$ dérivable dans $U(0)$ peut être mise sous forme de $F(z_1(0, h), \ldots, z_{n-1}(0, h))$ (*1.56c*), avec la fonction $F(z_1, \ldots, z_{n-1})$ dérivable. Soit maintenant $z(y)$ n'importe quelle intégrale première de l'équation (1); le difféomorphisme π^{-1} la transforme en une fonction $z(t, h)$ ne dépendant que de h, donc, la met sous la forme $F(z_1(0, h), \ldots, z_{n-1}(0, h)) = F(z_1(t, h), \ldots, z_{n-1}(t, h)) = F(z_1(y), \ldots, z_{n-1}(y))$, ce qu'il fallait démontrer.

c. Même une collection incomplète de $k \leqslant n-1$ premières intégrales indépendantes nous fournit une information essentielle sur les trajectoires du système. Notamment, s'il y a $k \leqslant n-1$ intégrales premières indépendantes, par exemple $z_1(y), \ldots, z_k(y)$, alors, pour tout choix de constantes $c_1, \ldots, c_k$ (d'un voisinage des valeurs $c_1^0 = z_1(y_0), \ldots, c_k^0 = z_k(y_0)$), les équations

$$z_1(y) = c_1, \ldots, z_k(y) = c_k \qquad (2)$$

déterminent, d'après *1.66b*, une variété $(n-k)$-dimensionnelle $M(c_1, \ldots, c_k) \subset Y$. De plus, si une trajectoire de l'équation (1)

a des points communs avec une variété $M(c_1, \ldots, c_k)$, elle y est incluse entièrement, puisque les fonctions $z_1(y), \ldots, z_k(y)$ conservent leurs valeurs sur cette trajectoire. Lorsque k devient plus grand, la dimension des variétés $M(c_1, \ldots, c_k)$ diminue et l'information sur la disposition des trajectoires devient plus exacte. Enfin, si $k = n - 1$, les équations (2) déterminent les variétés unidimensionnelles, c'est-à-dire les trajectoires mêmes.

d. Dans certains cas, on peut trouver $n - 1$ intégrales premières indépendantes au voisinage d'un point donné y_0, sans connaître les trajectoires. Alors ces dernières seront déterminées d'une façon implicite par les équations (2).

Exemple. Soit $Y = R_3$, $y = (y_1, y_2, y_3)$. Considérons l'équation

$$\frac{dy}{dt} = \{y_2 - y_3, y_3 - y_1, y_1 - y_2\}, \tag{3}$$

ou, sous la forme scalaire, le système d'équations

$$\frac{dy_1}{dt} = y_2 - y_3, \quad \frac{dy_2}{dt} = y_3 - y_1, \quad \frac{dy_3}{dt} = y_1 - y_2. \tag{4}$$

Nous cherchons les trajectoires de l'équation (3) ou, ce qui est le même, du système (4). En additionnant les équations (4) nous avons

$$\frac{d(y_1 + y_2 + y_3)}{dt} = 0.$$

Il en résulte que l'égalité $y_1 + y_2 + y_3 = \text{const}$ est valable sur chaque trajectoire du système (4) et en représente donc une intégrale première. Puis, en multipliant les équations (4) par y_1, y_2, y_3 respectivement et additionnant les résultats, nous obtenons

$$\frac{y_1\,dy_1 + y_2\,dy_2 + y_3\,dy_3}{dt} = y_1(y_2 - y_3) + y_2(y_3 - y_1) + y_3(y_1 - y_2) = 0,$$

d'où une autre intégrale première $z_2(y) = y_1^2 + y_2^2 + y_3^2$.

La matrice

$$\left\| \begin{matrix} \frac{\partial z_1}{\partial y_1} & \frac{\partial z_1}{\partial y_2} & \frac{\partial z_1}{\partial y_3} \\ \frac{\partial z_2}{\partial y_1} & \frac{\partial z_2}{\partial y_2} & \frac{\partial z_2}{\partial y_3} \end{matrix} \right\| = \left\| \begin{matrix} 1 & 1 & 1 \\ 2y_1 & 2y_2 & 2y_3 \end{matrix} \right\|$$

a le rang 2 partout, sauf les points de la droite $\gamma = \{y \in R_3 : y_1 = y_2 = y_3\}$. Notons que ces points sont, évidemment, des points immobiles du système (4). Au voisinage de chaque point $y \notin \gamma$, les trajectoires sont localement déterminées par les équations

$$y_1 + y_2 + y_3 = \text{const}, \quad y_1^2 + y_2^2 + y_3^2 = \text{const},$$

et nous voyons qu'elles représentent les circonférences orthogonales à la droite γ et de centres sur cette droite.

1.89. Equations linéaires aux dérivées partielles. Supposons qu'en tout point y d'un domaine G d'un espace de Banach Y soit donnée une fonction $\Phi(y)$ continue et continûment dérivable, à valeurs dans le même espace Y; autrement dit, $\Phi(y)$ représente un champ vectoriel. Nous cherchons une fonction $z(y): G \to Z$, continue et continûment dérivable, à partir de l'équation

$$z'(y)\,\Phi(y) = 0. \tag{5}$$

(Rappelons que $z'(y)$ est un opérateur linéaire de Y dans Z, donc le premier membre de (5) est l'image par l'opérateur $z'(y)$ du vecteur $\Phi(y)$.) Conformément à *1.27a*, l'équation (5) veut dire qu'à tout point $y \in G$ la dérivée de la fonction $z(y)$ suivant la direction du vecteur $\Phi(y)$ est nulle. Considérons, dans le domaine G, l'équation

$$\frac{dy}{dt} = \Phi(y). \tag{6}$$

Sur chaque trajectoire de l'équation (6) la fonction cherchée $z(y)$ devient fonction du paramètre t, et l'équation (5) signifie que la dérivée de cette dernière est nulle. Ainsi, la fonction cherchée $z(y)$ doit être constante sur chaque trajectoire de l'équation (6), c'est-à-dire doit être intégrale première de cette équation. La réciproque est évidente: toute intégrale première de l'équation (6) est une solution de l'équation (5). Par conséquent, le problème de résolution de l'équation (5) est réduit au problème des intégrales premières de l'équation (6). Par exemple, comme nous le savons déjà de *1.88b*, au voisinage de chaque point non fixe de l'équation (6), dans l'espace $Y = R_n$ il y a $n-1$ intégrales premières indépendantes; posons $z_1(y), \ldots, z_{n-1}(y)$, et toute intégrale première s'exprime en fonction d'elles selon la formule

$$z(y) = \psi(z_1(y), \ldots, z_{n-1}(y)),$$

où $\psi(z_1, \ldots, z_{n-1})$ est une fonction (arbitrairement choisie) continûment dérivable.

Ainsi, étant donnée l'équation

$$\frac{\partial z}{\partial y_1}(y_2 - y_3) + \frac{\partial z}{\partial y_2}(y_3 - y_1) + \frac{\partial z}{\partial y_3}(y_1 - y_2) = 0 \tag{7}$$

dans R_3, le système dynamique correspondant est déterminé par l'équation

$$\frac{dy}{dt} = \{y_2 - y_3,\ y_3 - y_1,\ y_1 - y_2\}$$

vue dans *1.88d*, et l'on connaît, en dehors de la droite $\gamma = \{y \in R_3 : y_1 = y_2 = y_3\}$, ses deux intégrales premières

$$z_1(y) = y_1 + y_2 + y_3 \qquad z_2(y) = y_1^2 + y_2^2 + y_3^2;$$

par conséquent, au voisinage de chaque point $y \notin \gamma$, toute solution de l'équation (7) est décrite par la formule

$$z(y) = \psi(y_1 + y_2 + y_3,\ y_1^2 + y_2^2 + y_3^2)$$

où $\psi(z_1, z_2)$ est une fonction continûment dérivable.

L'équation non homogène

$$z'(y)\,\Phi(y, z) = \psi(y, z) \qquad (Y \times R_1 \to R_1)$$

se ramène aisément à une équation homogène dans l'espace $Y \times R_1$ (voir l'exercice 28).

§ 1. 9. Equations différentielles (théorèmes non locaux)

1.91. Dans le paragraphe précédent, on étudiait les propriétés des solutions d'une équation différentielle au voisinage d'un point; ici nous considérons les propriétés des solutions dans de grands domaines. Soit Y un espace de Banach ; dans un domaine $G \subset R_1 \times Y$ (éventuellement non borné), il est donné une équation différentielle

$$\frac{dy}{dt} = \Phi(t, y). \tag{1}$$

Un sous-domaine $Q \subset G$ est dit *régulier* si, pour un $r > 0$, toute boule de rayon r et de centre à un point $a \in Q$ se trouve entièrement dans G. On suppose que la fonction $\Phi(t, y)$ dans (1) soit continue dans G, bornée dans tout sous-domaine régulier $Q \subset G$:

$$|\Phi(t, y)| \leqslant A_Q \tag{2}$$

et satisfasse à la condition de Lipschitz par rapport à y:

$$|\Phi(t, y_1) - \Phi(t, y_2)| \leqslant B_Q\,|y_1 - y_2|. \tag{3}$$

En vertu de *1.82b* et des hypothèses ci-formulées, par chaque point $(t_0, y_0) \in G$ passe une seule solution de l'équation (1):

$$y(t) \equiv y(t;\ t_0,\ y_0), \qquad y(t_0) = y_0. \tag{4}$$

Cette solution définie dans un intervalle $|t - t_0| < \delta$ peut s'avérer non prolongeable sur un intervalle plus ou moins important de l'axe des t, pour la raison, par exemple, qu'elle atteint, à une valeur finie $t = t_1$, la frontière du domaine G. Montrons que c'est *la seule raison possible.*

Théorème. *Toute solution (4) peut être prolongée dans les deux sens de variation de t jusqu'à ce qu'elle sorte de chaque sous domaine régulier $Q \subset G$.*

Démonstration. Pour une solution donnée (4), trouvons le plus grand demi-intervalle $[t_0, \beta)$ sur lequel elle soit définie; il peut être défini comme réunion de tous les demi-intervalles $[t_0, \beta')$

sur lesquels la solution (4) existe. (Remarquons ici que la quantité $y(t; t_0, y_0)$, si elle existe, est définie d'une façon univoque, car le théorème d'unicité *1.62* interdit la bifurcation de solutions dans le domaine G.) Supposons que $\beta < \infty$; considérons n'importe quelle suite $t_n \in [t_0, \beta)$, $t_n \nearrow \beta$ et la suite correspondante $(t_n, y(t_n))$. Supposons que l'arc $(t, y(t))$, pour $t \in [t_0, \beta)$, reste dans un sous-domaine régulier $Q \subset G$, de sorte que les valeurs $\Phi(t, y(t))$ soient majorées en module par la constante A_Q de (2). Alors, pour $m < n$, nous avons :

$$|y(t_n) - y(t_m)| \leqslant \sup_{t_m < t < t_n} \left| \frac{dy(t)}{dt} \right| (t_n - t_m) \leqslant A_Q (t_n - t_m).$$

Ainsi, la suite $y(t_n)$ est de Cauchy dans l'espace Y ; désignons sa limite par $\bar{y}$. La fonction $\Phi(t, y)$ étant continue, nous avons $\Phi(\beta, \bar{y}) = \lim_{n \to \infty} \Phi(t_n, y_n)$; par conséquent, en complétant l'arc semi-ouvert $(t, y(t))$, $t \in [t_0, \beta)$ par le point $(\beta, \bar{y})$ nous avons un arc fermé $(t, y(t))$, $t \in [t_0, \beta]$ avec la tangente continue, c'est-à-dire une solution de l'équation (1). En vertu du théorème *1.81b*, la solution peut être prolongée sur des valeurs de t supérieures à $t = \beta$, ce qui contredit la supposition. Nous voyons donc que, ou bien $\beta = \infty$, ou, si $\beta < \infty$, les points $(t, y(t))$ sortent, pour $t \in [t_0, \beta)$, de n'importe quel sous-domaine régulier $Q \subset G$, ce qu'on affirmait. Dans l'autre sens $(t \to -\infty)$ le théorème se démontre de façon analogue.

1.92. Dans ce qui suit nous étendrons les théorèmes sur la continuité et la dérivabilité de solution par rapport à la condition initiale au cas de grands intervalles de variation de t. Nous aurons besoin du lemme suivant :

Lemme. *Si une fonction* $\varphi(t)$, *continue et dérivable dans un intervalle* $[0, T]$, *satisfait à la condition*

$$|\varphi(t)| \leqslant M \left(1 + k \int_0^t |\varphi(\tau)| \, d\tau\right), \tag{1}$$

elle satisfait également, dans $[0, T]$, *à la condition*

$$|\varphi(t)| \leqslant M e^{kMt}. \tag{2}$$

Démonstration. Posons

$$\int_0^t |\varphi(\tau)| \, d\tau = v(t).$$

L'inégalité (1) prend alors la forme :

$$v'(t) \leqslant M(1 + kv(t)), \tag{3}$$

ou bien

$$\frac{v'(t)}{1+kv(t)} \leqslant M.$$

En l'intégrant et compte tenu de ce que $v(0) = 0$, nous avons

$$\ln(1 + kv(t)) \leqslant kMt$$

ou

$$1 + kv(t) \leqslant e^{kMt}.$$

En appliquant (3) trouvons

$$v'(t) \leqslant Me^{kMt},$$

d'où l'inégalité cherchée (2).

1.93. Pour $t_0 = 0$, nous écrirons $y(t, y_0)$ au lieu de $y(t; t_0, y_0)$.

Lemme. *Soit* $\{t, y(t, y_1)\}$, $0 \leqslant t \leqslant T$ *la courbe qui représente une solution de l'équation 1.91* (1) *et se trouve entièrement dans un sous-domaine régulier* Q *du domaine* G ($\subset R_1 \times Y$). *Nous affirmons qu'il existe* $\delta > 0$ *tel que toute solution* $y(t, y_2)$, $|y_2 - y_1| < \delta$ *est aussi bien définie pour* $0 \leqslant t \leqslant T$.

Démonstration. Choisissons un $\varepsilon > 0$ tel que le 2ε-voisinage du domaine Q est inclus dans le domaine G. Alors le ε-voisinage H du domaine Q représente, avec Q, un sous-domaine régulier; en particulier, la condition de Lipschitz

$$|\Phi(t, y) - \Phi(t, z)| \leqslant B_H |y - z|,$$

avec une constante fixe B_H, est remplie pour n'importe quels (t, y), (t, z) de H. Posons $\delta = \varepsilon e^{-B_H T}$ et considérons une solution arbitraire $y(t, y_2)$ avec $|y_2 - y_1| < \delta$. Soit $[0, \beta)$ le plus grand demi-intervalle pour lequel le point $(t, y(t, y_2))$ reste encore dans le domaine H; montrons que $\beta \geqslant T$. Soit $\beta < T$. Les solutions $y(t, y_1)$ et $y(t, y_2)$ vérifient les relations

$$y(t, y_1) = y_1 + \int_0^t \Phi(\tau, y(\tau, y_1))\, d\tau,$$

$$y(t, y_2) = y_2 + \int_0^t \Phi(\tau, y(\tau, y_2))\, d\tau.$$

On en tire, pour $0 \leqslant t < \beta$,

$$|y(t, y_1) - y(t, y_2)| \leqslant |y_1 - y_2| + \int_0^t |\Phi(\tau, y(\tau, y_1)) -$$

$$- \Phi(\tau, y(\tau, y_2))|\, d\tau \leqslant \delta + B_H \int_0^t |y(\tau, y_1) - y(\tau, y_2)|\, d\tau.$$

D'après le lemme *1.92*, nous avons

$$|y(t, y_1) - y(t, y_2)| \leqslant \delta e^{B_H \beta} \leqslant \varepsilon. \quad (1)$$

Ainsi, pour $0 \leqslant t \leqslant \beta$, la courbe $\{t, y(t, y_2)\}$ se trouve dans le sous-domaine régulier H. Alors, d'après *1.91*, la solution $y(t, y_2)$ peut être prolongée en dehors de $t = \beta$, ce qui contredit la supposition. Le lemme est donc démontré.

1.94. Théorème sur la continuité globale. *Considérons un arc* $\{t, y(t, y_1)\}$, $0 \leqslant t \leqslant T$ *représentant une solution de l'équation 1.91* (1) *qui est inclus dans un sous-domaine régulier* Q *du domaine* G. *D'après 1.93, il existe, pour un* $\delta > 0$ *et pour tous les* y_2, $|y_2 - y_1| < \delta$, *une solution* $y(t, y_2)$, $t \leqslant T$. *Nous affirmons que le point* $y(T, y_2)$ *dépend continûment de* y_2.

Démonstration. Il suffit de considérer une suite de points $y_2^{(1)}, \ldots, y_2^{(m)}, \ldots \to y_1$ et de montrer que $y(T, y_2^{(m)}) \to y(T, y_1)$. L'estimation *1.93* (1), pour $y_2 = y_2^{(m)}$, $\delta = \delta_m = |y_2^{(m)} - y_1|$, $t = T$, fournit

$$|y(T, y_1) - y(T, y_2^m)| \leqslant \delta_m e^{BT} \to 0,$$

ce qu'il fallait.

1.95. Théorème sur la dérivabilité globale. *Si, dans les hypothèses 1.91, la fonction* $\Phi(t, y)$ *a la dérivée continue* $\frac{\partial \Phi(t, y)}{\partial y}$ *alors, dans les hypothèses 1.94, le point* $y(T, y_0)$ *dépend de* y_0 *d'une façon dérivable.*

Démonstration. En vertu de *1.45*, la quantité $\frac{\partial \Phi(t, y)}{\partial y}$ est estimée, dans tout sous-domaine régulier Q, avec la constante figurant dans la condition de Lipschitz :

$$\left\| \frac{\partial \Phi(t, y)}{\partial y} \right\| \leqslant B_Q.$$

Les constructions que nous allons réaliser se déroulent dans les limites du domaine V qui représente la réunion de toutes les boules de rayon δ (du lemme *1.93*) et de centres aux points de la courbe $y(t, y_0)$, $0 \leqslant t \leqslant T$. Soit $(\bar{t}, \bar{y})$ un point sur cette courbe. Nous avons vu dans *1.85* que la solution $y(t; \bar{t}, \bar{y})$ (qui, en réalité, coïncide avec la solution $y(t, y_0)$ d'après le théorème d'unicité) est dérivable par rapport à $\bar{y}$, quel que soit t éloigné de $\bar{t}$ moins que de $\delta_0 < \frac{1}{B} \leqslant \left(\sup_V \left\| \frac{\partial \Phi(t, y)}{\partial y} \right\| \right)^{-1}$. Partageons l'intervalle $[0, T]$ avec les points $0 = t_0 < t_1 < \ldots < t_m = T$ de façon à avoir $t_{j+1} - t_j < \delta_0$. Nous aurons :

$y_1 = y(t_1, y_0)$ une fonction dérivable de y_0 ;

$y_2 = y(t_2;\ t_1,\ y_1)$ une fonction dérivable de y_1;

. . . . . .

$y_m = y(t_m;\ t_{m-1},\ y_{m-1})$ une fonction dérivable de y_{m-1}.

En appliquant un nombre convenable de fois la règle de dérivation d'une fonction composée nous voyons que $y_m = y(T,\ y_0)$ est une fonction dérivable de y_0, ce qu'il fallait démontrer.

1.96. Equation dans tout l'espace. Supposons que le domaine $G \subset R_1 \times Y$ de *1.91* coïncide avec tout l'espace $R_1 \times Y$. Il n'y a alors que deux possibilités, d'après *1.91* : ou bien la solution $y(t,\ y_0)$ soit définie *pour tout* t, $0 \leqslant t < \infty$, ou bien, pour une valeur finie $t = t_0$, la solution $y(t,\ y_0)$ devienne *non bornée*. (La deuxième possibilité est réalisée, par exemple, pour l'équation $\frac{dy}{dt} = y^2$ dans $R_1 \times R_1$: la solution correspondant à la valeur initiale $y_0 = \frac{1}{a} > 0$, $t = 0$ a la forme $y = \frac{1}{a-t}$ et ne peut être prolongée que jusqu'à la valeur $t = a$ dans le sens positif.)

La question se pose naturellement, quelles sont les conditions à imposer à la fonction $\Phi(t,\ y)$ pour que l'existence de la solution $y(t,\ y_0)$ soit garantie pour *tous* les t, $0 \leqslant t < \infty$. Montrons qu'une telle condition est donnée par l'inégalité

$$|\Phi(t,\ y)| \leqslant A_t + B_t |y|, \tag{1}$$

où les constantes A_t et B_t sont uniformément bornées dans chaque domaine borné de variation de t. Considérons, comme plus haut, le plus grand demi-intervalle $[0,\ \beta)$ sur lequel la solution $y(t,\ y_0)$ soit définie; nous avons à montrer que $\beta = \infty$. Soit $\beta < \infty$. Il résulte de l'équation

$$y(t) = y_1 + \int_0^t \Phi(\tau,\ y)\, d\tau, \quad 0 \leqslant t < \beta$$

et de (1) que

$$|y(t)| \leqslant |y_1| + \int_0^t (A_\tau + B_\tau |y|)\, d\tau \leqslant |y_1| + \bar{A}_\beta \cdot \beta + \bar{B}_\beta \int_0^t |y|\, d\tau,$$

où $\bar{A}_\beta = \sup\limits_{0 \leqslant t \leqslant \beta} A_t$, $\bar{B}_\beta = \sup\limits_{0 \leqslant t \leqslant \beta} \bar{B}_t$. En appliquant le lemme *1.92* nous voyons que

$$|y(t)| \leqslant (|y_1| + \bar{A}_\beta \beta)\, e^{\bar{B}_\beta \beta}.$$

Ainsi, la quantité $y(t)$ est bornée dans l'intervalle $[0,\ \beta)$, donc le point $\{t,\ y(t)\}$ reste dans un sous-domaine régulier du domaine G; mais d'après *1.91*, la solution $y(t,\ y_0)$ peut être prolongée sur les

valeurs de t supérieures à $t = \beta$, ce qui contredit l'hypothèse. Définitivement, pour que la solution $y(t, y_0)$ soit indéfiniment prolongeable par rapport à t, *il suffit* que (1) soit remplie.

En particulier, l'inégalité (1) a lieu pour l'équation linéaire

$$\frac{dy}{dt} = A(t)\, y + B(t) \qquad (2)$$

avec les coefficients $A(t): R_1 \to L(Y)$ et $B(t): R_1 \to Y$ continus (pour tout t). Ainsi, *chaque solution de l'équation linéaire (2) à coefficients $A(t)$ et $B(t)$ continus peut être prolongée sur tout l'axe des t*, $-\infty < t < \infty$.

1.97. Ici, comme dans *1.88*, nous interprétons les solutions $y = y(t, y_0)$ de l'équation $\frac{dy}{dt} = \Phi(t, y)$, comme loi de mouvement, dans l'espace Y, chaque du mobile qui se trouve à l'instant $t = 0$ au point y_0, $\Phi(t, y)$ étant la vitesse du mouvement. Toujours comme dans *1.88*, supposons que $\Phi(t, y)$ ne dépende pas de t, de sorte que l'équation donnée ait la forme

$$\frac{dy}{dt} = \Phi(y). \qquad (1)$$

Montrons qu'*a lieu la relation*

$$y(t,\ y(t_1,\ y_0)) = y(t + t_1,\ y_0), \qquad (2)$$

quels que soient t_1 et t admissibles pour le premier membre. En effet, si $y(t_1, y_0)$ a un sens, alors les deux membres dans (2) coïncident pour $t = 0$ et sont définis pour les $t \neq 0$ petits, au moins. Pour ces valeurs de t, les deux membres satisfont, évidemment, à l'équation différentielle (1). Dans ce cas, ils coïncident forcément pour tous les t où ils sont définis (puisque le théorème d'unicité *1.82* interdit leur séparation). Si les valeurs mentionnées de t n'épuisent pas le domaine de définition du second (premier) membre, alors, aux points qui restent, on peut définir le premier (second) membre en le posant égal aux valeurs correspondantes du second (premier) membre, et le premier (second) membre reste solution de l'équation (1) égale à $y(t_1, y_0)$ pour $t = 0$.

On a montré dans *1.88* que tout point $y_0 \in Y$ tel que $\Phi(y_0) \neq 0$ admet un voisinage $V(y_0)$ et un difféomorphisme π qui transforme les parties se trouvant dans $V(y_0)$ des trajectoires de l'équation (1) en des segments parallèles parcourus (lorsque t varie) à une vitesse constante. Nous allons étendre cette proposition au cas de grandes parties (en t) des trajectoires.

Considérons, au voisinage d'un point $y_0 \in Y$ avec $\Phi(y_0) \neq 0$, la décomposition $Y = L + H$ construite dans *1.88*, L étant un sous-espace unidimensionnel et H un sous-espace fermé de Y, et soit $H_\rho = \{h \in H,\ |h| \leqslant \rho\}$ une boule dans H qui, en produit avec

le segment T_δ, forme le domaine de valeurs du difféomorphisme π. Est-ce qu'on peut prolonger le difféomorphisme π (construit dans *I.88* pour $|t| \leqslant \delta$) sur de grandes valeurs de t? Les trajectoires de l'équation (1) prises en entier peuvent entrer dans la boule H_ρ encore une fois, c'est pourquoi la réponse à la question posée sera négative sans suppositions supplémentaires. Imposons aux solutions de l'équation (1) la condition suivante.

La condition du non-retour des trajectoires sur l'intervalle $[-t_1, t_2]$ $(t_1, t_2 > 0)$: *pour un* $\rho > 0$, *chaque trajectoire* $y(t, h)$, $h \in H_\rho$, $t \in [-t_1, t_2]$ *n'a plus* (*i.e. lorsque* $t \neq 0$) *de point commun avec la boule* H_ρ.

Dans cette condition, le prolongement du difféomorphisme s'avère possible:

Théorème. *Si les trajectoires* $y(t, h)$ *sont définies pour* $h \in H_\rho$, $t \in [-t_1, t_2]$ *et si la condition du non-retour est remplie sur l'intervalle* $[-t_1, t_2]$, *alors il existe un difféomorphisme de l'ensemble* $[-t_1, t_2] \times H_\rho$ *qui transforme tout point* (t, h) *en le point* $y(t, h)$.

Démonstration. Montrons que l'application $\pi: (t, h) \to y(t, h)$ est biunivoque. Si l'on avait $y(t_0, h_0) = y(t_1, h_1)$ pour certains $t_0 \leqslant t_1$, $h_0 \in H_\rho$, $h_1 \in H_\rho$, alors, en appliquant (2), on obtiendrait $y(t_1 - t_0, h_1) = y(-t_0, y(t_1, h_1)) = y(-t_0, y(t_0, h_0)) = y(0, h_0) = h_0$, ce qui contredit, pour $h_1 \neq h_0$, $t_1 \neq t_0$, la condition du non-retour. Par conséquent, l'application π est biunivoque. Notons ensuite qu'il n'y a pas de point fixe parmi les points $y(t, h)$ $(t \in [-t_1, t_2], h \in H_\rho)$ (ce qui découle du théorème d'unicité); donc, en vertu de *I.88*, l'application π représente un difféomorphisme d'un voisinage de chaque point (t, h) sur le voisinage correspondant du point $y(t, h)$. Par conséquent, l'application $\pi: (t, h) \to y(t, h)$ est biunivoque et bidérivable, donc difféomorphe, ce qu'il fallait démontrer.

En particulier, si la condition du non-retour est remplie, avec un $\rho > 0$, pour toute la droite numérique $-\infty < t < \infty$, nous voyons que l'ensemble des points de toutes les trajectoires $y(t, h)$, $h \in H_\rho$, $-\infty < t < \infty$ s'applique d'une façon difféomorphe sur l'ensemble des points des droites parallèles $-\infty < t < \infty$, $h \in H_\rho$, d'après la formule $y(t, h) \to (t, h)$.

Dans cette situation on dit que l'ensemble des trajectoires $y(t, h)$ $(-\infty < t < \infty, h \in H_\rho)$ est rectifiable.

EXERCICES

1. Montrer que la fonction $f(x)$ $(R_2 \to R_1)$ donnée en coordonnées polaires φ, r par les conditions

$$f(x) = \begin{cases} [r(1-r)]^\varphi \text{ pour } 0 < \varphi \leqslant 2\pi, 0 \leqslant r \leqslant 1, \\ 0 \text{ pour les autres valeurs de } \varphi \text{ et } r \end{cases}$$

est continue sur toute demi-droite issue de l'origine des coordonnées, sans être continue en l'origine même.

2. Montrer que toute fonction $f(x)$ $(R_2 \to R_1)$ continue sur toute courbe dérivable issue de l'origine des coordonnées est continue en l'origine aussi.

3. Soit $\lambda(\varphi)$ une fonction de φ, dérivable et périodique de période 2π. Montrer que la fonction $f(x)$ $(R_2 \to R_1)$ définie en coordonnées polaires φ, r par la formule

$$f(x) = \lambda(\varphi)\, r$$

a) est dérivable suivant toute demi-droite issue de l'origine des coordonnées, la dérivée valant $\lambda(\varphi)$;

b) est dérivable suivant la droite $\varphi = \varphi_0$ dans un seul cas où $\lambda(\varphi_0 + \pi) = -\lambda(\varphi_0)$;

c) est dérivable en l'origine des coordonnées si, et seulement si, $\lambda(\varphi) = \alpha \sin(\varphi + \beta)$, α et β étant des constantes.

4. Montrer que, pour qu'une application dérivable $w = f(z)$ $(G \subset R_2 \to R_2)$ définisse une fonction analytique $f(z)$, il faut et il suffit que, à tout point $z \in G$ l'opérateur $f'(z)$ soit celui de multiplication par un nombre complexe (dépendant de z).

5. Ecrire une équation différentielle pour la ligne de la plus rapide ascension d'une fonction $y = f(x_1, x_2)$ $(R_2 \to R_1)$.

6. Parmi les lignes de niveau de la fonction $\rho(x, -a)\, \rho(x, a)$ $(R_2 \to R_1)$, indiquer celles qui sont convexes; qui admettent des points d'inflexion; qui se composent de deux ovales.

7. Si une fonction numérique dérivable $f(x)$ $(G \subset X \to Y)$ prend une même valeur en deux points a et b, alors, pour tout contour L joignant ces points, la dérivée de la fonction $f(x)$ suivant L s'annule en un point au moins de ce contour. Donner un exemple de fonction vectorielle $(R_1 \to R_2$ au moins) pour laquelle le fait analogue n'a pas lieu.

8. Soit Q le plus petit ensemble convexe fermé contenant toutes les valeurs de la fonction $f'(t)$, $a \leqslant t \leqslant b$, $f\colon [a, b] \to Y$, $f(a) = f(b)$. Montrer que $0 \in Q$.

9. Considérons la surface d'un ellipsoïde de révolution en tant que surface de niveau de la fonction $\rho(x, a) + \rho(x, b)$. Montrer que les rayons lumineux qui partent du foyer a et sont réfléchis par la surface de l'ellipsoïde se réunissent au second foyer b.

10. Supposons qu'une fonction $F(t, \xi)$ $(R_2 \to R_1)$ ait la dérivée continue par rapport à ξ. Trouver les points stationnaires pour la fonctionnelle

$$f(x) = \int_a^b F(t, x(t))\, dt$$

définie dans l'espace $R(a, b)$ des fonctions réelles continues $x = x(t)$, $a \leqslant t \leqslant b$.

11. Faire l'analyse des points stationnaires pour la fonctionnelle de l'exercice 10 dans le cas où $a = 0$, $b = 1$, $F(t, \xi) = \xi^3 - (t+1)^2 \xi$.

12. Trouver les points stationnaires liés pour la fonctionnelle

$$f(x) = \int_0^1 x^3(t)\, dt, \quad x(t) \in C(0, 1)$$

sous la condition

$$g(x) \equiv \int_0^1 x^2(t)\, dt = 1$$

et étudier leurs types.

13. La méthode itérative de la résolution de l'équation $f(x) = 0$ ($f: X \to Y$) consiste en ce qui suit. Etant donnée une valeur initiale $x_1 = a$ où $f'(a)$ est un opérateur inversible, on écrit l'équation

$$f(a) + f'(a)(x_2 - a) = 0 \tag{1}$$

par rapport à x_2. Cette équation fournirait une racine exacte de l'équation $f(x) = 0$ si $f(x)$ était une fonction du premier degré. Dans le cas général, la solution x_2 de l'équation (1) est un nouveau point où la valeur $f(x_2)$ est, éventuellement, plus proche de 0 que $f(a)$. Nous tirons de (1):

$$x_2 = a - [f'(a)]^{-1} f(a). \tag{2}$$

La formule (2) nous suggère qu'il est bon de considérer la suite

$$x_{n+1} = x_n - [f'(x_n)]^{-1} f(x_n),$$

ou même la suite plus simple

$$x_{n+1} = x_n - [f'(a)]^{-1} f(x_n). \tag{3}$$

Montrer que, si les conditions

$$\| f'(a) \|^{-1} \sup_{|x-a| \leqslant r} \| f'(x) - f'(a) \| < \frac{1}{2},$$

$$| [f'(a)]^{-1} f(a) | < \frac{r}{2}$$

sont satisfaites, la suite (3) converge vers une solution de l'équation $f(x) = 0$, cette solution appartenant à la boule de rayon r centrée en a.

14. Montrer que la fonction $y = |x|$ ($X \to R_1$) n'est pas dérivable pour $x = 0$.

15. Montrer que, dans un espace hilbertien, la fonction $y = |x|$ est dérivable pour $x \neq 0$.

16. Montrer que, dans l'espace l_1 des suites numériques $x = \{\xi_1, \ldots, \xi_n, \ldots\}$ avec $|x| = \sum_1^\infty |\xi_k| < \infty$, la fonction $y = |x|$ ($l_1 \to R_1$) n'est dérivable en aucun point.

17. Pour voir comment sont disposées dans R_2 les branches d'une courbe

$$f(x, y) \equiv \sum_{0 \leqslant k,\, m \leqslant N} a_{km} x^k y^m = 0 \tag{1}$$

lorsque $x \searrow 0$ *, on utilise la règle suivante. Il faut poser dans l'équation (1) $y = Ax^r E$ où A et r sont des constantes inconnues, $E \to 1$ pour $x \searrow 0$; puis il faut déterminer $r = r_0$ dans l'équation obtenue

$$\sum_{0 \leqslant k,m \leqslant N} a_{km} A^m x^{k+rm} E = 0 \tag{2}$$

à partir de la condition que le plus petit des exposants $k + r_0 m$, désignons-le p, s'y rencontre deux fois au moins (à cette fin on peut se servir d'un diagramme des exposants dont un exemple est donné fig. 1.*E*-1 **; ensuite, on divise

---

* Le symbole $x \searrow 0$ veut dire: x tend vers zéro en décroissant.

** Pour plus de détails, voir Г. Е. Ш и л о в, Особые точки алгебраических кривых на плоскости, УМН 5, вып. 5 (1950), стр. 180-192.

l'équation (2) par x^p et on y pose $x = 0$; il est affirmé qu'à toute racine réelle simple $A_0 \neq 0$ de l'équation obtenue

$$\sum_{k+r_0 m=p} a_{km} A_0^m = 0 \tag{3}$$

il correspond une branche réelle de la courbe (1) avec pour équation $y = A_0 x^{r_0} E$.

Prouver la dernière affirmation.

18. (Suite.) Si A_0 est une racine multiple de l'équation (3), alors la courbe 1) n'a pas forcément une branche avec l'équation $y = A_0 x^{r_0} E$. Tout de même, si l'on pose dans (1) $y = A_0 x^{r_0} (1 + Bx^s E)$ et que l'on trouve de façon analogue les valeurs B_0 et s_0, alors, dès que B_0 soit une racine simple de l'équation correspondante, la courbe (1) admet une branche de la forme

$$y = A_0 x^{r_0} (1 + B_0 x^{s_0} E).$$

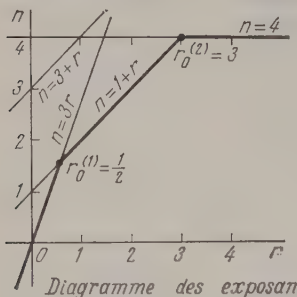


Diagramme des exposants de l'équation $x^4 + x^3y - y^3 - xy = 0$

Fig. 1.E-1

19. Montrer que les fonctions $y = f(x)$ $(G \subset X \to R_1)$ admettant la dérivée continue $f'(x)$ $(G \to L(X))$ forment une algèbre $\mathfrak{A}(G)$ qui est fermée par rapport à la norme X

$$\| f(x) \| = \sup_G \{ | f(x) | + \| f'(x) \| \}.$$

20. (Suite.) L'ensemble $J(a)$ de toutes les fonctions $f(x) \in \mathfrak{A}(G)$ pour lesquelles, en un point donné $a \in G$, sont remplies les conditions $f(a) = 0$, $f'(a) = 0$ représente un idéal fermé dans l'algèbre $\mathfrak{A}(G)$. Montrer que l'algèbre quotient $\mathfrak{A}/J(a)$ est isomorphe à l'espace X^* de toutes les fonctionnelles linéaires continues sur X, avec la multiplication nulle et avec une unité formellement introduite.

21. (Suite.) On appelle *dérivation formelle* pour $x = a \in G$ une fonctionnelle linéaire D dans l'algèbre $\mathfrak{A}(G)$ qui est continue sur la norme de l'algèbre et qui vérifie la condition

$$D[fg] = f(a) \cdot Dg + Df \cdot g(a).$$

Dans le cas où X est un espace hilbertien H, montrer que $f(x) \in J(a)$ implique $Df = 0$.

22. (Suite.) Montrer que, dans un espace hilbertien complet $X = H$, n'importe quelle dérivation formelle pour $x = a$ dans l'algèbre $\mathfrak{A}(G)$ est la dérivation au point a suivant un vecteur y:

$$Df = \lim_{t \to 0} \frac{f(a + ty) - f'(a)}{t}.$$

23. Donner un exemple de fonction $f(x)$ $(G \subset X \to L(X, Y))$ telle que, pour tout $h \in X$, la fonction $f(x)\,h$ $(G \to Y)$ soit dérivable et la fonction $f(x)$ ne le soit pas.

24. Donner un exemple de fonction $f(x, y)$ $(R_2 \to R_1)$ qui ne soit pas dérivable pour $x = 0$, mais qui ait la dérivée nulle suivant chaque direction partant du point 0. Montrer que l'existence de la dérivée nulle d'une fonction $f(x)$ $(R_n \to R_1)$ suivant chaque ligne (dérivable) qui part du point 0 implique déjà la dérivabilité de $f(x)$ en ce point.

25. Donner un exemple de fonction $f(x)$ $(H \to R_1)$ (H est un espace hilbertien de dimension infinie) qui ne soit pas dérivable pour $x = 0$, mais qui ait la dérivée nulle suivant chaque ligne dérivable issue du point 0.

26. Montrer que la dérivée de l'opérateur d'inversion $I(x) = \dfrac{x - x_0}{|x - x_0|^2}$ (dans un espace hilbertien H) a la forme $T(x)/|x - x_0|^2$ où $T(x)$ est un opérateur orthogonal (i.e. $(T(x)p,\ T(x)q) = (p, q)$, quels que soient p et q de H).

27. Les fonctions $y_1 = x_1$, $y_2 = x_2^2$ $(R_2 \to R_2)$, sont-elles indépendantes (dépendantes) dans un voisinage du point (0, 0)?

28. Soient Y un espace de Banach, $z(y)$: $Y \to R_1$, $\Phi(y, z)$: $Y \times R_1 \to Y$, $\Psi(y, z)$: $Y \times R_1 \to R_1$. On a l'équation

$$z'(y)\,\Phi(y, z) = \Psi(y, z). \qquad (1)$$

Montrer que toute solution $z(y)$ de l'équation (1) se déduit d'une solution $w(y, z)$ de l'équation homogène

$$\frac{\partial w}{\partial y}\,\Phi(y, z) + \frac{\partial w}{\partial z}\,\Psi(y, z) = 0 \qquad (2)$$

dans $Y \times R_1$ en imposant la condition $w(x, z) = 0$ (et réciproquement).

29. Soient M un espace métrique non compact et $x_1, \ldots, x_n, \ldots$ une suite de points de M qui n'admet aucune sous-suite de Cauchy. Montrer qu'il existe une telle suite de nombres positifs $p_1, \ldots, p_n, \ldots$ que les boules $S(x_n, p_n) = \{x \in M : \rho(x, x_n) \leqslant p_n\}$ sont deux à deux disjointes.

30. (Suite.) Construire, dans un espace métrique non compact M, une fonction numérique continue $f(x)$ telle que $\sup f(x) = \infty$.

31. (Suite.) Construire, dans un espace métrique non compact M, une fonction continue mais non uniformément continue.

Historique

Déjà les fondateurs de l'analyse infinitésimale comprennent que la dérivation de fonctions amène à des expressions linéaires par rapport aux accroissements des coordonnées; ils ne négligent pas cette possibilité de simplification de problèmes compliqués. C'est bien dans la linéarisation, dans la résolution du problème au niveau linéaire suivie par le retour à des quantités finies avec l'intégration, que réside l'essentiel de la découverte de Newton et Leibniz. La technique des différentielles totales est développée par Euler (années 1730). Pourtant, sans la théorie des limites systématisée, sinon créée, par Cauchy au début du XIX^e siècle, la linéarisation de XVII^e et XVIII^e siècles semble contradictoire; les mathématiciens de l'époque croient, probablement, que les fonctions considérées soient elles-mêmes linéaires à petite échelle et que leurs différentielles ne soient rien que leurs accroissements correspondant à des accroissements « extrêmement petits » des variables indépendantes. Il est, naturellement, impossible de développer ce point de vue d'une façon plus ou moins plausible, ce qui rend très difficile la construction des fondements de l'analyse et provoque au même temps la critique meurtrière de la part des philosophes. Berkeley, par exemple, après avoir critiqué les « sophismes frappants » des fluxions (dérivées) newtonniennes, dit: « celui qui est capable de digérer la deuxième ou la troisième fluxion... n'a pas à argoter, il me semble, sur quoi que ce soit dans la théologie ». Hegel, le philosophe d'une autre orientation, lie les méthodes des infiniment petits avec les lois dialectiques de la pensée, qu'il a découvertes, et traite la dérivation comme négation (d'une quantité finie) et l'intégration, comme négation de la négation; de ce point de vue, toute l'analyse apparaît comme résultat de l'application de la dialectique aux mathématiques; il devient alors clair pourquoi toutes les « démonstrations » qu'il y a dans l'analyse de l'époque sont inadmissibles du point de vue de la logique

formelle: il ne peut être autrement si l'on part des propositions non formelles. La justesse des raisonnements de Hegel ne contribue pourtant rien au progrès de l'analyse infinitésimale; à une certaine étape, la formalisation concrète des constructions de base devient nécessaire pour son développement fructueux. Les grands mathématiciens du XVIII^{e} siècle, tels que Euler, d'Alembert, Lagrange, réfléchissent sur la possibilité d'une telle formalisation; la construction d'un système logiquement irréprochable de l'analyse fait l'objet de concours académiques. Et pourtant, ce n'est qu'au XIX^{e} siècle, après avoir accumulé une quantité suffisante de faits, que Cauchy et Weierstrass réalisent la formalisation voulue. L'analyse des infiniment petits, débarrassée de contradictions formelles, devient tout de suite accessible à un cercle beaucoup plus vaste de chercheurs, son développement avance avec une rapidité et une efficacité extraordinaires. Le mérite principal de Cauchy est de considérer la différentielle d'une fonction non pas comme son accroissement mais comme partie linéaire principale de cet accroissement; la définition exacte de cette notion est basée chez Cauchy sur la notion de limite qui sert de fondement à tout le calcul infinitésimal. Weierstrass y ajoute la technique des raisonnements avec ε et δ, ce qui permet, en particulier, de corriger certaines conclusions un peu précipitées de Cauchy. Durant tout le XIX^{e} siècle plusieurs auteurs, de Cauchy à Goursat, forment de différentes idées du calcul différentiel des fonctions de plusieurs variables, y compris la théorie de fonctions implicites, de fonctions indépendantes et dépendantes, de jacobiens (ces derniers sont introduits par Jacobi en 1833, pour obtenir la règle générale de changement des variables dans une intégrale multiple, cf. chap. 3). En 1911-1913 on propose plusieurs variantes de définition de la différentielle d'une fonctionnelle (Fréchet, Gateaux; cf. L é v i P., Problèmes concrets d'analyse fonctionnelle, Gauthier-Villars, Paris, 1951); lorsque l'idée d'espace normé pénètre en mathématiques (1920-1922), c'est la définition de Fréchet qui reste principale (*1.23*). Elle ouvre la voie d'extension du calcul différentiel sur les fonctions définies dans des espaces de dimension infinie. La généralisation correspondante du théorème sur la fonction implicite est donnée par Hildebrandt et Graves (1927).

C'est encore du temps de Newton et Leibniz qu'on applique le calcul différentiel à la recherche d'extrémum (même avant: Fermat, 1629). La méthode des multiplicateurs dans les problèmes d'extrémum lié est décrite par Lagrange en 1797. Pour une fonctionnelle sur un espace normé et pour une liaison de dimension finie, elle est établie par Lusternik en 1934.

Dérivées d'ordre supérieur

On peut construire la théorie des dérivées d'ordre supérieur d'une fonction $y = f(x)$ $(G \subset X \to Y)$ soit en partant de la définition par récurrence $f^{(n)}(x) = [f^{(n-1)}(x)]'$, soit en se basant sur les parties linéaires principales du second, du troisième, etc., degré de l'accroissement de la fonction. Nous verrons dans § 2.4, que, dans les suppositions convenables de continuité, les deux méthodes sont équivalentes.

Au début nous étudions les fonctions numériques de n variables réelles (§ 2.1) ; les résultats classiques concrets que nous y obtenons nous servent de base pour la formation de certaines notions et propositions générales. Dans la théorie générale où les variables indépendante et dépendante prennent leurs valeurs dans des espaces à plusieurs dimensions, une nouvelle circonstance intervient : les valeurs des dérivées d'ordre supérieur appartiennent à des espaces de plus en plus éloignés ; en même temps, les différentielles d'ordre supérieur se trouvent naturellement liées aux formes multilinéaires symétriques par rapport aux différentielles des variables indépendantes, plutôt qu'aux formes polynomiales. Ces considérations conduisent au théorème de Frobenius (§ 2.5) qui est un résultat bien important : la condition de résolubilité de l'équation du premier ordre $y' = F(x, y)$ est une condition imposée à la fonction $F(x, y)$ étroitement liée à la propriété de symétrie de la différentielle seconde en tant que forme bilinéaire. Puisque, en langage classique, le théorème de Frobenius fournit les conditions de compatibilité d'un système d'équations du premier ordre à dérivées partielles, il représente, de même que le théorème sur la fonction implicite, un des plus puissants moyens de l'analyse.

§ 2. 1. Dérivées d'ordre supérieur d'une fonction numérique de n variables

2.11. a. Si les dérivées partielles $\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}$ d'une fonction numérique $f(x_1, \ldots, x_n) : G \subset R_n \to R_1$ qui est dérivable en tout point d'un domaine G sont elles aussi des fonctions dérivables $(G \subset R_n \to R_1)$, il est possible de considérer les dérivées partielles secondes que l'on désigne comme suit :

$$\frac{\partial}{\partial x_1}\left(\frac{\partial f(x)}{\partial x_1}\right) \equiv \frac{\partial^2 f(x)}{\partial x_1^2}, \quad \frac{\partial}{\partial x_2}\left(\frac{\partial f(x)}{\partial x_1}\right) \equiv$$

$$\equiv \frac{\partial^2 f(x)}{\partial x_2 \partial x_1}, \ldots, \frac{\partial}{\partial x_i}\left(\frac{\partial f(x)}{\partial x_j}\right) \equiv \frac{\partial^2 f(x)}{\partial x_i \partial x_j}, \ldots, \frac{\partial}{\partial x_n}\left(\frac{\partial f(x)}{\partial x_n}\right) \equiv \frac{\partial^2 f(x)}{\partial x_n^2}.$$

En continuant la dérivation (si possible), on obtient les dérivées partielles du troisième ordre $\frac{\partial}{\partial x_k}\left(\frac{\partial^2 f(x)}{\partial x_i\,\partial x_j}\right) \equiv \frac{\partial^3 f(x)}{\partial x_k\,\partial x_i\,\partial x_j}$, du quatrième ordre $\frac{\partial}{\partial x_p}\left(\frac{\partial^3 f(x)}{\partial x_k\,\partial x_i\,\partial x_j}\right) \equiv \frac{\partial^4 f(x)}{\partial x_p\,\partial x_k\,\partial x_i\,\partial x_j}$, etc.

b. Conformément à la définition ci-dessus, une fonction de n variables peut posséder n dérivées du premier ordre, n^2 dérivées du second ordre, n^3 du troisième, etc. En fait, le nombre de dérivées différentes (i.e. dont les valeurs en un point donné ne coïncident pas) de tout ordre fixe est inférieur; il s'avère, pour d'assez bonnes fonctions, que l'on peut permuter la succession des variables par rapport auxquelles on dérive, sans changer le résultat, par exemple, $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$. Le théorème suivant a lieu:

Théorème. *Si les fonctions* $\frac{\partial^2 f(x)}{\partial x_i \partial x_j}$ *et* $\frac{\partial^2 f(x)}{\partial x_j \partial x_i}$ *existent dans un voisinage d'un point* $a = \{a_1, \ldots, a_n\}$ *et sont continues en ce point, alors*

$$\frac{\partial^2 f(a)}{\partial x_i\,\partial x_j} = \frac{\partial^2 f(a)}{\partial x_j\,\partial x_i}.$$

Démonstration. Sans restreindre la généralité, on peut poser $i = 1$, $j = 2$. En n'explicitant que la dépendance de x_1 et x_2 nous écrivons dans ce qui suit $f(x) = f(x_1, x_2)$. Considérons l'expression

$$w = f(a_1 + h_1,\ a_2 + h_2) - f(a_1 + h_1,\ a_2) - \\ - f(a_1,\ a_2 + h_2) + f(a_1,\ a_2). \qquad (1)$$

C'est l'accroissement de la fonction

$$\Phi(x_1) = f(x_1,\ a_2 + h_2) - f(x_1,\ a_2),$$

lorsque x_1 varie de a_1 à $a_1 + h_1$. D'après le théorème de Lagrange, on a

$$w \equiv \Phi(a_1 + h_1) - \Phi(a_1) = \Phi'(a_1 + \theta_1 h_1)\,h_1 = \\ = \left[\frac{\partial f(a_1 + \theta_1 h_1,\ a_2 + h_2)}{\partial x_1} - \frac{\partial f(a_1 + \theta_1 h_1,\ a_2)}{\partial x_1}\right] h_1$$

pour un θ_1, $0 < \theta_1 < 1$. En réappliquant le théorème de Lagrange, par rapport à x_2 cette fois, on trouve

$$w = \frac{\partial^2 f(a_1 + \theta_1 h_1,\ a_2 + \theta_2 h_2)}{\partial x_2\,\partial x_1}\, h_2 h_1 \qquad (2)$$

avec un θ_2, $0 < \theta_2 < 1$. Or, la même quantité w peut être considérée comme accroissement de la fonction

$$\psi(x_2) = f(a_1 + h_1,\ x_2) - f(a_1,\ x_2),$$

lorsque x_2 varie de a_2 à $a_2 + h_2$. Toujours avec le théorème de Lagrange, on trouve

$$w = \frac{\partial^2 f\,(a_1+\tau_1 h_1,\ a_2+\tau_2 h_2)}{\partial x_1\,\partial x_2}\,h_1 h_2 \qquad (3)$$

avec certains τ_1 et τ_2, $0 < \tau_1 < 1$, $0 < \tau_2 < 1$.

En égalant (2) et (3) et divisant par $h_1 h_2$, on obtient

$$\frac{\partial^2 f\,(a_1+\theta_1 h_1,\ a_2+\theta_2 h_2)}{\partial x_2\,\partial x_1} = \frac{\partial^2 f\,(a_1+\tau_1 h_1,\ a_2+\tau_2 h_2)}{\partial x_1\,\partial x_2}.$$

Faisons tendre vers zéro h_1 et h_2; par hypothèse de continuité des fonctions $\frac{\partial^2 f(x)}{\partial x_1 \partial x_2}$ et $\frac{\partial^2 f(x)}{\partial x_2 \partial x_1}$ pour $x = a$ nous avons, à la limite,

$$\frac{\partial^2 f\,(a)}{\partial x_1\,\partial x_2} = \frac{\partial^2 f\,(a)}{\partial x_2\,\partial x_1},$$

ce qu'il nous fallait.

c. Pour les dérivées d'ordre supérieur à deux, la possibilité de permutation des dérivations se démontre (sous l'hypothèse d'existence et de continuité des dérivées en question au point $x = a$) en appliquant plusieurs fois le théorème *b*. Ainsi, par exemple,

$$\frac{\partial^3 f\,(x)}{\partial x_1^2\,\partial x_2} = \frac{\partial}{\partial x_1}\left(\frac{\partial^2 f}{\partial x_1\,\partial x_2}\right) = \frac{\partial}{\partial x_1}\left(\frac{\partial^2 f}{\partial x_2\,\partial x_1}\right) = \frac{\partial^3 f}{\partial x_1\,\partial x_2\,\partial x_1} =$$

$$= \frac{\partial^2}{\partial x_1\,\partial x_2}\left(\frac{\partial f}{\partial x_1}\right) = \frac{\partial^2}{\partial x_2\,\partial x_1}\left(\frac{\partial f}{\partial x_1}\right) = \frac{\partial^3 f}{\partial x_2\,\partial x_1^2}.$$

d. Lorsque la fonction $f\,(x_1, \ldots, x_n)$ est un polynôme en $x_1, \ldots, x_n$ d'un degré inférieur à k, il est évident que toutes les dérivées d'ordre supérieur ou égal à k de la fonction $f\,(x)$ sont nulles. En particulier, pour une fonction linéaire $f\,(x) = \sum_{i=1}^{n} c_i x_i$, les dérivées secondes sont déjà nulles.

2.12. Différentielles d'ordre supérieur.

a. D'après *1.22* (3), la différentielle première d'une fonction $y = f\,(x_1, \ldots, x_n)$: $G \subset R_n \to R_1$ vaut

$$dy = \sum_{i=1}^{n} \frac{\partial f\,(x)}{\partial x_i}\,dx_i \qquad (1)$$

et représente une fonction des variables $x = \{x_1, \ldots, x_n\}$ et $dx = \{dx_1, \ldots, dx_n\}$, linéaire par rapport à dx. En considérant (1) comme fonction de x seul nous pouvons, d'une manière analogue,

calculer sa *différentielle totale seconde* $d(dy)$:

$$d^2y \equiv d(dy) = \sum_{j=1}^{n} \frac{\partial (dy)}{\partial x_j} dx_j = \sum_{j=1}^{n} \frac{\partial}{\partial x_j} \left(\sum_{i=1}^{n} \frac{\partial f(x)}{\partial x_i} dx_i \right) dx_j =$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f(x)}{\partial x_j \partial x_i} dx_i dx_j.$$

C'est une fonction des variables x_i et dx_i qui est déjà du second degré en dx_i.

En continuant de même nous obtenons la différentielle totale troisième

$$d^3y \equiv d(d^2y) = d \left(\sum_{i,j=1}^{n} \frac{d^2 f(x)}{\partial x_j \partial x_i} dx_i dx_j \right) =$$

$$= \sum_{i,j,k=1}^{n} \frac{\partial^3 f(x)}{\partial x_k \partial x_j \partial x_i} dx_i dx_j dx_k,$$

qui représente déjà une forme cubique de la variable dx, et ainsi de suite ; la *différentielle totale s-ème de la fonction* $y = f(x)$ est définie par récurrence :

$$d^s y = d(d^{s-1} y)$$

et représente une forme de s-ème degré par rapport aux coordonnées du vecteur dx. A côté des différentielles totales ci-définies dy, d^2y, d^3y, . . ., on considère les différentielles partielles d'ordre supérieur. Ainsi, l'expression

$$\frac{\partial^2 y}{\partial x_1 \partial x_2} dx_2 dx_1$$

s'appelle *différentielle partielle seconde de la fonction* $y = f(x)$ *correspondant aux différentielles* dx_2 et dx_1 des variables indépendantes. La définition générale des différentielles partielles est construite de façon analogue.

b. Si la fonction $f(x_1, \ldots, x_n)$ est un polynôme en $x_1, \ldots, x_n$ d'un degré inférieur à k, alors, conformément à *2.11d*, toutes les différentielles d'ordre supérieur ou égal à k de la fonction $f(x)$ sont nulles. En particulier, pour la fonction linéaire $f(x) = \sum_{i=1}^{n} c_i x_i$, c'est déjà la différentielle totale seconde qui s'annule. Les relations $d^2x_i = 0$ sont souvent utilisées. Ce qu'on a dit ici est valable lorsque x_i sont des variables indépendantes ; si x_i sont des fonctions d'autres variables, ces formules n'ont en général pas lieu.

2.13. Formule de Taylor. On se sert des différentielles d'ordre supérieur pour préciser le comportement d'une fonction au voisinage d'un point donné. Notamment, si les différentielles correspondantes existent, on a la série suivante des formules de plus en plus précises:

$$\Delta y \equiv f(a+dx) - f(a) = \sum_{i=1}^{n} \frac{\partial f(a)}{\partial x_i} dx_i + o(dx) \equiv dy(a) + o(dx),$$

$$\Delta y = dy(a) + \frac{1}{2} d^2 y(a) + o(|dx|^2),$$

. . . . . .

$$\Delta y = dy(a) + \frac{1}{2} d^2 y(a) + \ldots + \frac{1}{m!} d^m y(a) + o(|dx|^m),$$

. . . . . .

où $\lim\limits_{dx \to 0} o(|dx|^m)/|dx|^m = 0$. Toutes ces formules résultent de la *formule* générale *de Taylor*

$$y(a+dx) = y(a) + dy(a) + \frac{1}{2} d^2 y(a) + \ldots$$

$$\ldots + \frac{1}{m!} d^m y(a) + \frac{1}{(m+1)!} d^{m+1} y(a + \theta\, dx). \qquad (1)$$

La déduction de la formule de Taylor sera donnée dans *2.41*. Dans *2.43* on démontrera la proposition suivante: si, dans un domaine $G \subset R_n$ et pour un m, la fonction $y(x)$ vérifie l'égalité

$$y(x+dx) = y(x) + \sum_{i=1}^{n} \varphi_i(x)\, dx_i + \frac{1}{2} \sum_{i,\, j=1}^{n} \varphi_{ij}(x)\, dx_i\, dx_j + \ldots$$

$$\ldots + \frac{1}{m!} \sum_{i_1, \ldots, i_m = 1}^{n} \varphi_{i_1, \ldots, i_m}(x)\, dx_{i_1} \ldots dx_{i_m} + o(|dx|^m) \qquad (2)$$

avec les coéfficients continus $\varphi_i(x)$, $\varphi_{ij}(x)$, . . ., $o(|dx|^m)$ étant un infiniment petit uniforme, alors la fonction $y = f(x)$ est dérivable dans le domaine G, et la décomposition (2) représente sa décomposition de Taylor.

Cette proposition établit l'équivalence de la définition directe des dérivées d'ordre supérieur à la définition qui se base sur la séparation des termes du premier, du second, etc., ordres de petitesse dans l'accroissement de la fonction.

2.14. Comportement d'une fonction numérique au voisinage d'un point donné aux infiniment petits d'ordre supérieur à deux près.

a. Pour $m = 1$, la formule de Taylor représente la définition d'une fonction dérivable et détermine la partie linéaire principale de l'accroissement de la fonction.

Pour $m = 2$, la formule de Taylor sous la forme

$$y(a+dx) - y(a) - dy(a) = \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 y(a)}{\partial x_i \partial x_j} dx_i dx_j + o(|dx|^2) \quad (1)$$

détermine la partie quadratique principale de ce qui reste après avoir séparé dans l'accroissement de la fonction sa partie linéaire principale. Cette partie quadratique principale est donnée par la forme quadratique

$$Q(dx) = \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 y(a)}{\partial x_i \partial x_j} dx_i dx_j. \quad (2)$$

b. La forme (2) peut être positive pour tous les $dx \neq 0$, comme, par exemple, dans le cas

$$Q(dx) = \sum_{i=1}^{n} (dx_i)^2.$$

Si la forme quadratique $Q(dx) = \sum_{i,j=1}^{n} q_{ij} dx_i dx_j$ est positive, soit $C > 0$ sa valeur minimale sur la sphère unité $|dx| = 1$; alors, pour tout dx, nous avons

$$\left| \frac{dx}{|dx|} \right| = 1, \quad Q\left(\frac{dx}{|dx|}\right) = \sum_{i,j=1}^{n} q_{ij} \frac{dx_i dx_j}{|dx|^2} \geqslant C,$$

d'où

$$Q(dx) = \sum_{i,j=1}^{n} q_{ij} dx_i dx_j \geqslant C |dx|^2.$$

Par conséquent, du moment que la différentielle seconde $d^2y(a)$ est une forme quadratique positive en dx, nous avons, pour $\varepsilon > 0$ donné et pour tout $dx \neq 0$ suffisamment petit :

$$\frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 y(a)}{\partial x_i \partial x_j} dx_i dx_j + o(|dx|^2) \geqslant (C - \varepsilon) |dx|^2 \geqslant C_1 |dx|^2 > 0,$$

donc

$$\Delta y - dy > 0, \qquad \Delta y > dy.$$

Cela veut dire que le graphique de la fonction $y = f(x_1, \ldots, x_n)$, au voisinage du point $x = a$, est disposé au-dessus du plan tangent

$$dy = \sum_{i=1}^{n} \frac{\partial y(a)}{\partial x_i} dx_i$$

(*1.26b*) (fig. 2.1-1).

c. D'une manière analogue, si la forme $d^2y\ (a)$ est négative pour tous les $dx \neq 0$ (par exemple, $d^2y = -\sum_{i=1}^{n} (dx_i)^2$), alors, $dx \neq 0$ étant suffisamment petit, nous avons $\Delta y < dy$, le graphique de la

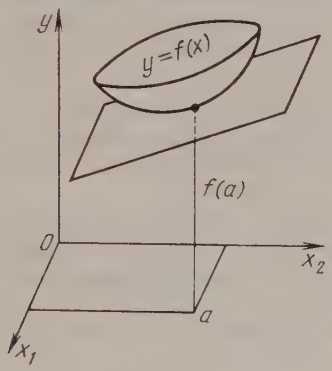


Fig. 2.1-1

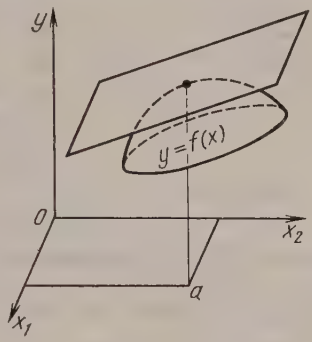


Fig. 2.1-2

fonction $y = f\ (x_1, \ldots, x_n)$ se trouvant donc au-dessous du plan tangent (au voisinage du point $x = a$) (fig. 2.1-2).

d. La forme quadratique $d^2y\ (a)$ peut être positive pour certaines valeurs de dx et négative pour d'autres (par exemple, la forme $d^2y\ (a) = \sum_{i=1}^{m} dx_i^2 - \sum_{j=m+1}^{n} dx_j^2$); cela veut dire que l'on a $\Delta y > dy$ pour certaines dx et $\Delta y < dy$, pour d'autres. Géométriquement, le graphique de la fonction $y = f\ (x)$ au voisinage du point $x = a$ sera en forme de selle, en se disposant en partie au-dessus du plan tangent, en partie, au-dessous (fig. 2.1-3).

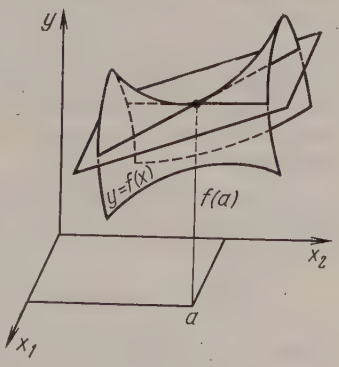


Fig. 2.1-3

e. Il peut se présenter un cas dégénéré où la forme $d^2y\ (a)$, tout en étant non négative (non positive), s'annule sur une droite (ou sur un ensemble de droites); alors on ne peut plus étudier le comportement du graphique de la fonction $y\ (x)$ suivant cette droite (ou ces droites) d'après la différentielle seconde, il faut attirer les différentielles supérieures.

Supposons, par exemple que $y = y\ (x) \equiv y\ (x_1, x_2)$, $d^2y \geqslant 0$ et que, pour un point donné $a = (a_1, a_2)$, il n'y ait dans le plan (x_1, x_2) qu'une droite, mettons l, sur laquelle $d^2y = 0$.

Alors la surface $y = y(x)$ a, au voisinage du point a, la forme d'un chéneau allant suivant la droite l. Si l'on tient compte des infiniment petits d'ordre supérieur, le chéneau peut présenter une courbure dans la direction de la droite l; ainsi, lorsque $d^3y \not\equiv 0$ sur la droite l, le graphique de la fonction $y(x)$ passe, le long de la droite l, d'un côté du plan tangent à l'autre (puisque la forme impaire d^3y change de signe le long de l). C'est cela qui conditionne la forme courbée du chéneau (fig. 2.1-4).



Fig. 2.1-4

f. Il y a une règle algébrique (*règle de Sylvester*) qui permet de reconnaître directement d'après les coefficients de la forme quadratique $Q(\xi) = \sum_{i,j=1}^{n} q_{ij}\xi_i\xi_j$ lequel des cas ci-mentionnés se réalise. A savoir, il faut calculer n déterminants

$$\delta_1 = q_{11}, \quad \delta_2 = \begin{vmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{vmatrix}, \quad \ldots, \quad \delta_n = \begin{vmatrix} q_{11} & \ldots & q_{1n} \\ \cdot & \cdot & \cdot \\ q_{n1} & \ldots & q_{nn} \end{vmatrix}. \tag{3}$$

La règle de Sylvester dit: si tous les déterminants $\delta_1, \delta_2, \ldots, \delta_n$ sont positifs, il en est de même de la forme $Q(\xi)$; si $\delta_1 < 0$, $\delta_2 > 0$, $\delta_3 < 0$, . . ., la forme $Q(\xi)$ est négative; si tous les nombres δ_j sont non nuls, et leurs signes sont répartis d'une autre façon, la forme prend les valeurs de deux signes. Si l'un des déterminants δ_k est nul, le critère de Sylvester ne donne pas de réponse; il faut une étude plus détaillée. On peut trouver la démonstration de la règle de Sylvester dans *L7.96*.

On peut également montrer que la forme $Q(\xi)$ prend les valeurs de deux signes dès que l'un au moins des déterminants $\delta_2, \delta_4, \delta_6, \ldots$ soit négatif (voir l'exercice 1).

g. Exemple. Soit $y = f(x_1, x_2) = 3x_1x_2 - x_1^3 - x_2^3$. Dans ce cas:

$$\frac{\partial f}{\partial x_1} = 3x_2 - 3x_1^2, \quad \frac{\partial f}{\partial x_2} = 3x_1 - 3x_2^2,$$

$$\frac{\partial^2 f}{\partial x_1^2} = -6x_1, \quad \frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1} = 3, \quad \frac{\partial^2 f}{\partial x_2^2} = -6x_2,$$

$$d^2y = -6x_1(dx_1)^2 + 2 \cdot 3\, dx_1\, dx_2 - 6x_2(dx)^2, \tag{4}$$

$$\delta_1 = -6x_1, \quad \delta_2 = \begin{vmatrix} -6x_1 & 3 \\ 3 & -6x_2 \end{vmatrix} = 9(4x_1x_2 - 1).$$

Sur la fig. 2.1-5 sont montrés les domaines dans le plan x_1, x_2 où les quantités δ_1 et δ_2 ne s'annulent pas; donc d'après la règle de Sylvester, on peut déduire certaines assertions sur le comportement de la fonction $f(x_1, x_2)$.

Considérons les points auxquels la règle ne s'applique pas. Avant tout, sur l'axe des x_2 nous avons $\delta_1 = -6x_1 = 0$. Si maintenant on fait changer de rôle les coordonnées et l'on considère x_2 comme première coordonnée, alors puisque dans ce cas $\delta_1 = -6x_2 \neq 0$, à tout point $(0, x_2)$, $x_2 \neq 0$ il correspond un point-selle du graphique. Au point (0, 0) la permutation des axes ne donne plus de

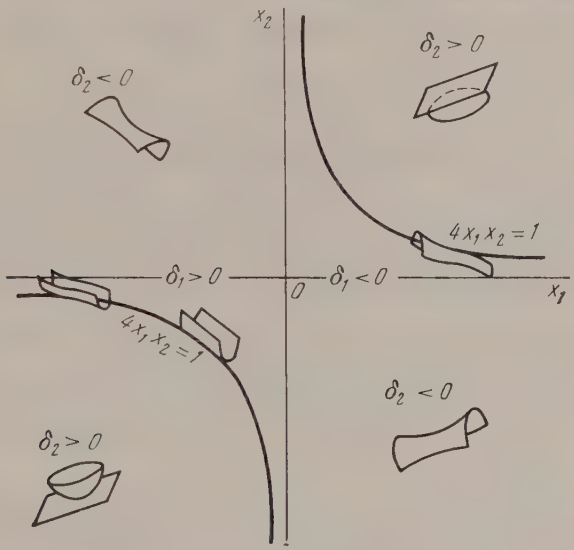


Fig. 2.1-5

résultat ; mais ici $d^2y = 6dx_1\,dx_2$; cette forme change de signe, donc le graphique a la forme de selle à ce point aussi. Aux points de l'hyperbole $4x_1x_2 = 1$ où $\delta_2 = 0$, la différentielle seconde devient proportionnelle au carré parfait d'une forme linéaire :

$$d^2y = -6x_1\,(dx_1)^2 + 2\cdot 3\,dx_1\,dx_2 - 6x_2\,(dx_2)^2 =$$

$$= \begin{cases} -6\left(\sqrt{x_1}\,dx_1 - \dfrac{1}{2\sqrt{x_1}}\,dx_2\right)^2 & \text{pour } x_1 > 0, \\ 6\left(\sqrt{-x_1}\,dx_1 + \dfrac{1}{2\sqrt{-x_1}}\,dx_2\right)^2 & \text{pour } x_1 < 0. \end{cases}$$

Sur les droites le long desquelles la forme linéaire correspondante s'annule, l'analyse du comportement de la fonction exige la considération des termes du troisième ordre. Ces droites sont

$$\left.\begin{aligned} &\sqrt{x_1}\,dx_1 - \frac{1}{2\sqrt{x_1}}\,dx_2 = 0, \quad \text{soit} \quad dx_2 = 2x_1\,dx_1 \quad \text{pour} \quad x_1 > 0; \\ &\sqrt{-x_1}\,dx_1 + \frac{1}{2\sqrt{-x_1}}\,dx_2 = 0, \quad \text{soit} \quad dx_2 = 2x_1\,dx_1 \quad \text{pour} \quad x_1 < 0. \end{aligned}\right\} \quad (5)$$

Ici dx_1 et dx_2 désignent les accroissements des coordonnées le long de la droite (5). On peut noter $dx_1 = X_1 - x_1$, $dx_2 = X_2 - x_2$ où (x_1, x_2) est un point de

l'hyperbole, (X_1, X_2) le point courant de la droite, de sorte que les équations (5) prennent la forme

$$X_2 - x_2 = 2x_1 (X_1 - x_1).$$

Pour savoir comment la fonction $y(x_1, x_2)$ se comporte le long de ces droites, dérivons encore une fois l'égalité (4):

$$d^3y = -12 (dx_1)^3 - 12 (dx_2)^3 = -12 ((dx_1)^3 + (dx_2)^3).$$

Si $dx_2 = 2x_1 dx_1$, alors

$$d^3y = -12 (dx_1)^3 (1 + 8x_1^3).$$

Pour $x_1 \neq -1/2$ la forme d^3y est non dégénérée, et le graphique de la fonction $y(x)$ au voisinage de ces points représente un « chéneau courbé » du type de la fig. 2.1-4. Et si $x_1 = -1/2$ (donc aussi $x_2 = -1/2$), alors on a le long de la ligne $dx_2 = -dx_1$: $dy = 0$, $d^2y = 0$, $d^3y = 0$, donc le polynôme du troisième degré $y(x)$ est constant sur cette droite; le graphique de la fonction $y(x)$ au voisinage du point $(-1/2, -1/2)$ est encore un chéneau mais qui n'a plus de courbure dans la direction longitudinale.

2.15. Les résultats obtenus fournissent certains critères permettant de classer les points stationnaires (*1.81a*).

a. Soit $a \in G$ un point stationnaire d'une fonction numérique $y = f(x)$: $R_n \to R_1$. Supposons que $f(x)$ soit deux fois dérivable

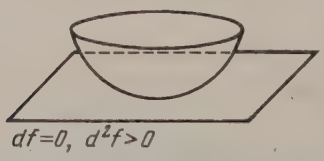


Fig. 2.1-6

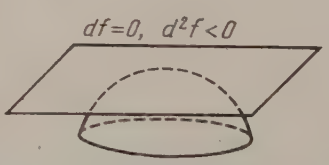
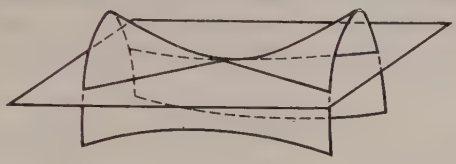


Fig. 2.1-7

dans le domaine G. Alors on a $f'(a) = 0$, et le plan tangent à la surface $y = f(x)$ au point $(a, f(a))$ est horizontal. Envisageons au point a la différentielle seconde $d^2f(a) = \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 f(a)}{\partial x_i \partial x_j} dx_i\, dx_j$ de la fonction $f(x)$. Si la forme quadratique $d^2f(a)$ est positive, alors, en vertu de *2.14b*, le graphique de la fonction $f(x)$ se trouve *au-dessus* du plan tangent, au voisinage du point a; en d'autres termes $f(a + dx) \geqslant f(a)$ pour les valeurs suffisamment petites de $|dx|$, ou bien, ce qui est le même, le point a est un point de *minimum local* (fig. 2.1-6) de la fonction $f(x)$. Si la forme quadratique $d^2f(a)$ est négative, alors, par analogie, le point a est un point de *maximum local* de la fonction $f(x)$ (fig. 2.1-7). Si la forme $d^2f(a)$ n'est pas définie, c'est-à-dire qu'elle prend les valeurs de deux signes dans tout voisinage du point a, alors le point stationnaire *a n'est pas un point extrémal* (fig. 2.1-8).

Et si la forme $d^2f(a)$ dégénère (par exemple, s'annule identiquement), on peut toujours essayer les différentielles d'ordre supérieur. Pourtant, pour ces dernières, il n'y a plus de critères assez simples



$df=0$, d^2f de deux signes

Fig. 2.1-8

comme celui de Sylvester dans le cas de la différentielle seconde, on est forcé de considérer ces différentielles directement, comme nous l'avons fait dans l'exemple précédent.

b. Exemple. Les points stationnaires de la fonction $f(x) = 3x_1x_2 - x_1^3 - x_2^3$ se déterminent d'après les équations

$$\frac{\partial f(x)}{\partial x_1} = 3x_2 - 3x_1^2 = 0, \qquad \frac{\partial f(x)}{\partial x_2} = 3x_1 - 3x_2^2 = 0$$

de sorte qu'il y a deux points stationnaires : $a_1^{(1)} = a_2^{(1)} = 0$ et $a_1^{(2)} = a_2^{(2)} = 1$. Nous les avons classés dans *1.81c*, nous y avons vu que le second est un point de maximum, le premier n'étant pas extrémal. Là, on a obtenu ces résultats par les calculs plus ou moins simples. Ici, avec la différentielle seconde et la règle de Sylvester, ils sont immédiats : le point $a_1^{(2)} = a_2^{(2)} = 1$ appartient au domaine où $\delta_1 < 0$, $\delta_2 > 0$, c'est-à-dire où la forme $d^2f(a)$ est négative, donc représente un point de maximum ; le point $a_1^{(1)} = a_2^{(1)} = 0$ se trouve dans le domaine où $\delta_2 < 0$, donc n'est pas extrémal.

Ainsi, l'utilisation de la différentielle seconde peut essentiellement simplifier l'analyse du comportement d'une fonction au voisinage d'un point stationnaire donné.

2.16. Les mêmes méthodes peuvent être appliquées aux fonctions implicites.

a. Soit $y = y(x)$ une fonction donnée par l'équation

$$\Phi(x_1, \ldots, x_n, y) = 0, \qquad (1)$$

l'hypothèse du théorème sur la fonction implicite *1.53* étant remplie à un point $\{a, b\} = \{a_1, \ldots, a_n, b\}$ et la fonction $\Phi(x_1, \ldots, x_n, y)$ étant m fois continûment dérivable par rapport à toutes les variables au voisinage du point $\{a, b\}$. On montrera dans *2.36* (pour un cas plus général) que la fonction implicite $y = y(x)$ possède elle aussi les dérivées continues, y compris les m-èmes, au voisinage du point a. En appliquant ce théorème on peut trouver toutes les différentielles de la fonction $y(x)$ au point $x = a$, y compris la m-ème, sans résoudre l'équation (1) par rapport à y. Raisonnons de la façon suivante. Si

l'on met dans l'équation (1) sa solution $y(x)$ à la place de y, on obtient une fonction de n variables $x_1, \ldots, x_n$ qui est identiquement nulle dans un voisinage du point a. Toutes les différentielles de cette fonction sont donc nulles elles aussi. Nous les calculons tour à tour; en calculant la première nous nous rappelons qu'il est sans importance si y est une variable indépendante ou fonction d'une autre variable:

$$\frac{\partial \Phi(x, y)}{\partial x_1} dx_1 + \ldots + \frac{\partial \Phi(x, y)}{\partial x_n} dx_n + \frac{\partial \Phi(x, y)}{\partial y} dy = 0,$$

en particulier,

$$\frac{\partial \Phi(a, b)}{\partial x_1} dx_1 + \ldots + \frac{\partial \Phi(a, b)}{\partial x_n} dx_n + \frac{\partial \Phi(a, b)}{\partial y} dy = 0. \tag{2}$$

D'où

$$dy(a) = -\frac{1}{\dfrac{\partial \Phi(a, b)}{\partial y}} \sum_{i=1}^{n} \frac{\partial \Phi(a, b)}{\partial x_i} dx_i, \tag{3}$$

ce qu'on aurait pu écrire encore dans *1.55* lorsqu'on cherchait les dérivées partielles d'une fonction implicite.

b. En particulier, pour trouver les points stationnaires de la fonction $y(x)$, nous avons à considérer le sytsème suivant de $n+1$ équations à $n+1$ inconnues $x_1 = a_1, \ldots, x_n = a_n, y = b$:

$$\left.\begin{aligned} &\Phi(x_1, \ldots, x_n, y) = 0, \\ &\frac{\partial \Phi(x_1, \ldots, x_n, y)}{\partial x_1} = 0, \\ &\ldots\ldots\ldots\ldots \\ &\frac{\partial \Phi(x_1, \ldots, x_n, y)}{\partial x_n} = 0. \end{aligned}\right\} \tag{4}$$

En dérivant l'équation (2) encore une fois et tenant compte du fait que les variables $x_1, \ldots, x_n$ sont indépendantes, de sorte que $d^2 x_i = 0$ (*2.12b*), et que y est une fonction, nous trouvons

$$\begin{aligned} &d\left(\frac{\partial \Phi(x, y)}{\partial x_1}\right) dx_1 + \ldots + d\left(\frac{\partial \Phi(x, y)}{\partial x_n}\right) dx_n + \\ &\quad + d\left(\frac{\partial \Phi(x, y)}{\partial y}\right) dy + \frac{\partial \Phi(x, y)}{\partial y} d^2 y = \\ &\quad = \frac{\partial^2 \Phi(x, y)}{\partial x_1^2} dx_1^2 + \frac{\partial^2 \Phi(x, y)}{\partial x_2 \partial x_1} dx_2 dx_1 + \ldots \\ &\ldots + \frac{\partial^2 \Phi(x, y)}{\partial x_n \partial x_1} dx_n dx_1 + \frac{\partial^2 \Phi(x, y)}{\partial y \partial x_1} dy dx_1 + \ldots \\ &\ldots + \frac{\partial^2 \Phi(x, y)}{\partial x_1 \partial x_n} dx_1 dx_n + \ldots + \frac{\partial^2 \Phi(x, y)}{\partial x_n^2} dx_n^2 + \end{aligned}$$

$$+\frac{\partial^2\Phi(x,\ y)}{\partial y\,\partial x_n}\,dy\,dx_n+\frac{\partial^2\Phi(x,\ y)}{\partial x_1\,\partial y}\,dx_1\,dy+\ldots$$

$$\ldots+\frac{\partial^2\Phi(x,\ y)}{\partial x_n\,\partial y}\,dx_n\,dy+\frac{\partial^2\Phi(x,\ y)}{\partial y^2}\,dy^2+\frac{\partial\Phi(x,\ y)}{\partial y}\,d^2y=0. \qquad (5)$$

En posant ici $x=a$, $y=b$ et compte tenu de (3), on a

$$d^2y(a)=-\frac{1}{\dfrac{\partial\Phi(a,\ b)}{\partial y}}\Big[\sum_{i,\ j=1}^{n}\frac{\partial^2\Phi(a,\ b)}{\partial x_i\,\partial x_j}\,dx_i\,dx_j+$$

$$+2\sum_{i=1}^{n}\frac{\partial^2\Phi(a,\ b)}{\partial x_i\,\partial y}\,dx_i\,dy+\frac{\partial^2\Phi(a,\ b)}{\partial y^2}\,dy^2\Big]=$$

$$=-\frac{1}{\dfrac{\partial\Phi(a,\ b)}{\partial y}}\Big[\sum_{i,\ j=1}^{n}\frac{\partial^2\Phi(a,\ b)}{\partial x_i\,dx_j}\,dx_i\,dx_j-$$

$$-\frac{2}{\dfrac{\partial\Phi(a,\ b)}{\partial y}}\sum_{i,\ j=1}^{n}\frac{\partial^2\Phi(a,\ b)}{\partial x_i\,\partial y}\,\frac{\partial\Phi(a,\ b)}{\partial x_j}\,dx_i\,dx_j+$$

$$+\frac{\dfrac{\partial^2\Phi(a,\ b)}{\partial y^2}}{\left(\dfrac{\partial\Phi(a,\ b)}{\partial y}\right)^2}\sum_{i,\ j=1}^{n}\frac{\partial\Phi(a,\ b)}{\partial x_i}\,\frac{\partial\Phi(a,\ b)}{\partial x_j}\,dx_i\,dx_j\Big]. \qquad (6)$$

Si $(a,\ b)$ est un point stationnaire, alors $dy(a)=0$, et la relation (5) prend la forme simple que voici :

$$\sum_{i,\ j-1}^{n}\frac{\partial^2\Phi(a,\ b)}{\partial x_i\,\partial x_j}\,dx_i\,dx_j+\frac{\partial\Phi(a,\ b)}{\partial y}\,d^2y(a)=0.$$

On en tire

$$d^2y(a)=-\frac{1}{\dfrac{\partial\Phi(a,\ b)}{\partial y}}\sum_{i,\ j=1}^{n}\frac{\partial^2\Phi(a,\ b)}{\partial x_i\,\partial x_j}\,dx_i\,dx_j, \qquad (7)$$

c'est bien la forme quadratique qu'il faut analyser pour reconnaître le type du point stationnaire donné.

En continuant la dérivation de l'égalité (5) on aboutira aux formules pour les différentielles d'ordre supérieur, mais nous n'insistons pas là-dessus.

c. Exemple. Trouvons les points stationnaires de la fonction $y=y(x)$ $(R_1\to R_1)$ définie par l'équation

$$\Phi(x,\ y)=x^3+y^3-3xy=0. \qquad (8)$$

Pour le faire, nous avons à résoudre, conformément à b, le système composé de l'équation (8) et de l'équation

$$\frac{\partial \Phi(x, y)}{\partial x} \equiv 3x^2 - 3y = 0. \qquad (9)$$

En éliminant y nous arrivons à l'équation bicubique

$$x^3 + x^6 - 3x^3 = 0,$$

d'où les points stationnaires

$$x_1 = y_1 = 0, \quad x_2 = \sqrt[3]{2}, \quad y_2 = \sqrt[3]{4}.$$

Le point (0, 0) ne vérifie pas l'hypothèse du théorème sur la fonction implicite et on ne le considère plus. Ayant en vue le deuxième point $\{\sqrt[3]{2}, \sqrt[3]{4}\}$, cherchons la différentielle seconde de la fonction $y(x)$ en le faisant non pas d'après (6), mais directement. Nous avons

$$3x^2\,dx + 3y^2\,dy - 3x\,dy - 3y\,dx = 0;$$

en divisant par 3 et dérivant encore une fois, trouvons

$$2x\,dx^2 + 2y\,dy^2 + y^2\,d^2y - dx\,dy - x\,d^2y - dy\,dx = 0.$$

Puisque $dy = 0$ en tout point stationnaire, nous avons

$$2x\,dx^2 + y^2\,d^2y - x\,d^2y = 0,$$

d'où

$$d^2y = \frac{2x}{x - y^2}\,dx^2, \quad d^2y(\sqrt[3]{2}) = -2\,dx^2 < 0;$$

par conséquent, la fonction $y(x)$ a un maximum local au point $x = \sqrt[3]{2}$.

2.17. L'utilisation des différentielles secondes permet également, en principe, de classer les points stationnaires dans les problèmes sur l'extrémum lié. Nous n'envisageons ici que le problème sur l'extrémum lié d'une fonction $y = f(x) \equiv f(x_1, \ldots, x_n)$ avec une seule condition : $\varphi(x) \equiv \varphi(x_1, \ldots, x_n) = c$ ($\varphi : R_n \to R_1$) dans un domaine $G \subset R_n$. Soit $x = a = (a_1, \ldots, a_n)$ un point stationnaire de ce problème ; supposons qu'il soit un point ordinaire de deux surfaces $\varphi(x) = c$ et $f(x) = b$ ($= f(a)$) ; ces hypothèses sont remplies, par exemple, lorsque $\frac{\partial \varphi(a)}{\partial x_n} \neq 0$, $\frac{\partial f(a)}{\partial x_n} \neq 0$, ce que nous supposons donné. Les égalités *1.83* (2) que l'on peut mettre sous la forme

$$\operatorname{grad} f(a) = \lambda \operatorname{grad} \varphi(a)$$

montrent que les surfaces de niveau $\varphi = c$ et $f = b$ ont un plan tangent commun au point $x = a$. En résolvant les équations $f = b$ et $\varphi = c$ par rapport à x_n nous aboutissons aux équations respectives de ces surfaces :

$$x_n = g(x'), \quad x_n = \psi(x'), \quad x' = (x_1, \ldots, x_{n-1});$$

de plus, nous avons $g(a') = \psi(a') = a_n$ pour $a' = (a_1, \ldots, a_{n-1})$.

Supposons encore que $\frac{\partial f(a)}{\partial x_n} > 0$, sinon on peut inverser le sens de l'axe des x_n. Nous avons donc $\frac{\partial f(x)}{\partial x_n} > 0$ pour tous les x assez proches de a.

Supposons ensuite que les fonctions $f(x)$ et $\varphi(x)$, donc aussi $g(x')$ et $\psi(x')$ (*2.16*) soient deux fois dérivables. Les relations suivantes sont alors valables:

$$\left.\begin{aligned} g(x') \equiv g(a'+h') &= g(a') + (g'(a'), h') + \\ &+ \frac{1}{2} \sum_{i,j=1}^{n-1} \frac{\partial^2 g(a')}{\partial x_i \partial x_j} h'_i h'_j + o(|h'|^2), \\ \psi(x') \equiv \psi(a'+h') &= \psi(a') + (\psi'(a'), h') + \\ &+ \frac{1}{2} \sum_{i,j=1}^{n-1} \frac{\partial^2 \psi(a')}{\partial x_i \partial x_j} h'_i h'_j + o(|h'|^2). \end{aligned}\right\} \quad (1)$$

Théorème. *Dans les hypothèses formulées, considérons la forme quadratique*

$$Q(h', h') = \sum_{i,j=1}^{n-1} \left(\frac{\partial^2 g(a')}{\partial x_i \partial x_j} - \frac{\partial^2 \psi(a')}{\partial x_i \partial x_j} \right) h'_i h'_j. \quad (2)$$

Du moment que cette forme est définie positive, le point a est un point de maximum lié de la fonction f (x) sous la condition φ (x) = c; si la forme (2) est définie négative, le point a est un point de minimum lié de la fonction f (x); si la forme (2) n'est pas définie, alors le point a n'est pas un point d'extrémum lié de la fonction f (x).

Démonstration. Supposons que la forme (2) soit définie positive. Alors, compte tenu du fait que $g(a') = \psi(a')$, $g'(a') = \psi'(a')$, pour $h' = \{h_1, \ldots, h_{n-1}\} \neq 0$ assez petits, nous avons $g(x') > \psi(x')$ (cf. *2.14b*). Ensuite, il résulte de la condition $\frac{\partial f(x)}{\partial x_n} > 0$ (qui est valable pour tous les x suffisamment proches de a) que $f(x', \psi(x')) < f(x', g(x'))$ $(x' \neq a')$.

Or, le point $(x', \psi(x'))$ appartient à la surface $\varphi(x) = c$ et le point $(x', g(x'))$ à la surface $f(x) = b$, de sorte que le second membre de la dernière inégalité vaut b. Nous voyons que l'inégalité $f(x) < b$ a lieu partout dans l'intersection de la surface $\varphi(x) = c$ avec un voisinage assez petit du point a (sauf au point a même). Il s'ensuit que le point a est un point de *maximum lié* de la fonction $f(x)$ sous la condition $\varphi(x) = c$. Les deux autres possibilités s'épuisent de façon analogue. Le théorème est démontré.

Dans le cas où la liaison $\varphi(x) = c$ est linéaire (même pour φ: $R_n \to R_k$) il existe un critère qui détermine le type de l'extrémum lié d'après les mineurs diagonaux d'un déterminant formé des quantités $\frac{\partial^2 f(a)}{\partial x_i \partial x_j}$ et des coefficients dans les équations de la liaison (*L7.98*).

§ 2. 2. Définition générale des dérivées supérieures

2.21. a. Comme nous avons déjà vu, pour une fonction $f(x)$: $G \subset X \to Y$, l'opérateur linéaire $f'(a): X \to Y$ est défini par l'égalité

$$f(a+h) - f(a) = f'(a)\,h + o(h), \quad h \in X.$$

Supposons que $f'(x)$ existe en tout point $x \in G$. Alors $f'(x)$ est une fonction qui fait correspondre à tout point $x \in G$ un opérateur $X \to Y$. Nous désignons par Y_1 l'espace $L(X, Y)$ de tous les opérateurs linéaires continus de X dans Y (et par Y_0 l'espace même Y). De la sorte, $f(x)$ est une fonction de x à valeurs dans Y_0 et $f'(x)$ une fonction de x à valeurs dans Y_1.

A présent, définissons l'opérateur $f''(a)$ par l'égalité (si elle est possible)

$$f'(a+h) - f'(a) = f''(a)\,h + o(h), \quad h \in X.$$

Alors, $f''(a)$ est un opérateur linéaire continu de X dans Y_1. Désignons par Y_2 l'espace $L(X, Y_1)$ de tous les opérateurs de ce type. Si $f''(x)$ existe pour tout $x \in G$, elle est une fonction de x à valeurs dans Y_2. En continuant de même nous arrivons à la définition générale suivante.

Définition. Soit $Y_p = L(X, Y_{p-1})$ $(p = 1, 2, \ldots)$; un opérateur linéaire $f^{(p)}(a): X \to Y_{p-1}$ s'appelle *dérivée p-ème de la fonction* $f(x)$ $(G \subset X \to Y_0)$ *à un point* $x = a$, si la relation suivante a lieu

$$f^{(p-1)}(a+h) - f^{(p-1)}(a) = f^{(p)}(a)\,h + o(h) \quad (h \in X). \tag{1}$$

Tous les termes dans cette égalité appartiennent à l'espace Y_{p-1}. Ainsi $f^{(p)}(a) \in Y_p$.

Pour abréger on peut noter

$$f^{(p)}(a) = [f^{(p-1)}(x)]'|_{x=a} \tag{2}$$

en réduisant ainsi la définition de la dérivée p-ème, en fin des comptes, à celle de la dérivée première.

b. Une fonction $y = f(x)$ $(G \subset X \to Y)$ qui possède, pour $x = a \in G$, les dérivées successives y compris la p-ème est dite *p fois dérivable* ou, en abrégé, *p-dérivable au point a*; lorsque cette situation a lieu en tout point du domaine $G \in X$, la fonction est dite *p fois dérivable* (*p-dérivable*) *dans le domaine G*. Une fonction $f(x)$ qui possède au point $x = a$ (à tout point $x \in G$) les dérivées de tous

les ordres $p = 1, 2, \ldots$ se dit *indéfiniment dérivable au point a (dans le domaine G).*

c. *Si une fonction $f(x)$ est k fois dérivable en un point a et que sa dérivée k-ème est m fois dérivable au même point, alors la fonction $f(x)$ est $k + m$ fois dérivable au point $x = a$, et l'on a*

$$f^{(k+m)}(a) = [f^{(k)}(x)]^{(m)}|_{x=a}. \tag{3}$$

En effet, si $m = 1$, alors la proposition se réduit à la définition d'une fonction $k + 1$ fois dérivable; dans le cas général, le résultat énoncé s'établit aisément par récurrence.

L'inversion de la proposition ci-dessus est immédiate: *si une fonction $f(x)$ est $k + m$ fois dérivable pour $x = a$, alors $f^{(k)}(x)$ est m fois dérivable et la formule (3) a lieu.*

2.22. Si $X = R_1$, alors tout opérateur linéaire $X \to Y$ s'identifie naturellement à un élément de l'espace Y, de sorte que $Y \equiv Y_0 = Y_1 = Y_2 = \ldots$ (*1.14c*). Par conséquent, étant donnée une fonction $f(x)$ d'une variable réelle x à valeurs dans l'espace Y, toutes les dérivées $f'(x)$, $f''(x)$, ... prennent elles aussi leurs valeurs dans l'espace Y. (Nous l'avons déjà vu dans *U12.64.*)

2.23. Envisageons le cas $X = R_n$, $Y = R_1$. Ici $Y_1 = L(R_n, R_1) = R_n$, $Y_2 = L(R_n, R_n) = R_{n^2}$, etc. L'opérateur $f'(x)$ peut être déterminé par n composantes:

$$\frac{\partial f(x)}{\partial x_1}, \ldots, \frac{\partial f(x)}{\partial x_n};$$

cet opérateur agit sur le vecteur de déplacement $h = dx = \{dx_1, \ldots, dx_n\}$ d'après la formule *1.22* (5):

$$f'(x)\,dx = \sum_{i=1}^{n} \frac{\partial f(x)}{\partial x_i}\,dx_i.$$

L'existence de $f'(x)$ implique celle des dérivées mentionnées (mais pas inversement; voir l'exercice 3, chap. 1).

L'existence et la continuité de $f'(x)$ sont équivalentes à l'existence et la continuité des dérivées mentionnées (*1.47c*).

En appliquant les mêmes raisonnements à la dérivée $f''(x)$ on obtient:

L'existence de $f''(x)$ signifie la dérivabilité de toutes les fonctions $\frac{\partial f'(x)}{\partial x_1}, \ldots, \frac{\partial f'(x)}{\partial x_n}$ et implique, en particulier, l'existence de toutes les dérivées secondes $\frac{\partial^2 f(x)}{\partial x_i \partial x_j}$ (mais pas réciproquement!); *l'existence et la continuité de $f''(x)$ dans un domaine G sont équivalentes à l'existence et la continuité de toutes les dérivées partielles secondes $\frac{\partial^2 f(x)}{\partial x_i \partial x_j}$ dans le domaine G.* En continuant de même nous voyons que *l'existence de l'opérateur $f^{(p)}(x)$ signifie la dérivabilité de toutes les dérivées partiel-*

les d'ordre $p-1$ de la fonction $f(x)$ et, en particulier, implique l'existence de toutes les dérivées partielles d'ordre p (mais pas réciproquement!), *et l'existence et la continuité de l'opérateur $f^{(p)}(x)$ dans un domaine G sont équivalentes à l'existence et la continuité de toutes les dérivées partielles d'ordre p de la fonction $f(x)$ dans le domaine G.*

2.24. Formes multilinéaires.

a. Etant donnés un opérateur linéaire $A_1: X \to Y = Y_0$ et un vecteur $h_1 \in X$, on peut former le vecteur $A_1 h_1 \in Y_0$. Avec un opérateur linéaire $A_2: X \to Y_1 = L(X, Y_0)$ et un vecteur $h_2 \in X$, on forme l'opérateur $A_2 h_2 \in L(X, Y_0)$, puis, avec le vecteur $h_1 \in X$, l'expression $(A_2 h_2) h_1 \in Y$. C'est une forme bilinéaire des vecteurs h_1 et h_2 à valeurs dans Y_0. En continuant de la même manière, à l'aide d'un opérateur linéaire $A_p: X \to Y_{p-1} = L(X, Y_{p-2})$ et des vecteurs $h_1, \ldots, h_p$ de l'espace X, on établit l'expression

$$A_p h_p h_{p-1} \ldots h_1 \equiv (\ldots ((A_p h_p) h_{p-1}) \ldots h_1) \in Y_0$$

qui représente une forme p-linéaire des vecteurs $h_1, \ldots, h_p$ à valeurs dans Y_0.

b. Si $A_p: X \to Y_{p-1}$ est un opérateur linéaire borné, alors, conformément à la définition générale de la norme d'un opérateur (*U12.71b*), on a l'estimation

$$| A_p h_p h_{p-1} \ldots h_1 | \leqslant \| A_p h_p \ldots h_2 \| \, | h_1 | \leqslant \ldots$$

$$\ldots \leqslant \| A_p \| \, | h_p | \ldots | h_1 |,$$

d'où

$$\sup_{|h_1| \leqslant 1, \ldots, |h_p| \leqslant 1} | A_p h_p \ldots h_1 | \leqslant \| A_p \|. \qquad (1)$$

D'autre part, soit θ, $0 < \theta < 1$, un nombre fixe; étant donné un opérateur $A_p: X \to Y_{p-1}$, il existe un vecteur h_p, $| h_p | = 1$, tel que $\| A_p h_p \| \geqslant \theta \| A_p \|$; ensuite, il y a un vecteur h_{p-1}, $| h_{p-1} | = 1$ tel que $\| A_p h_p h_{p-1} \| \geqslant \theta \| A_p h_p \| \geqslant \theta^2 \| A_p \|$; en continuant cette construction nous aboutissons à une suite $h_p, \ldots, h_1$ des vecteurs normés pour lesquels on a l'inégalité

$$| A_p h_p h_{p-1} \ldots h_1 | \geqslant \theta^p \| A_p \|.$$

Il en découle

$$\sup_{|h_1| \leqslant 1, \ldots, |h_p| \leqslant 1} | A_p h_p \ldots h_1 | \geqslant \theta^p \| A_p \|$$

et, $\theta \in (0, 1)$ étant arbitraire,

$$\sup_{|h_1| \leqslant 1, \ldots, |h_p| \leqslant 1} | A_p h_p \ldots h_1 | \geqslant \| A_p \|. \qquad (2)$$

En comparant (1) et (2) nous voyons que

$$\| A_p \| = \sup_{|h_1| \leqslant 1, \ldots, |h_p| \leqslant 1} | A_p h_p \ldots h_1 |.$$

c. Si l'on pose $h_1 = \ldots = h_p = h$ on obtient une forme *de degré p*

$$A_p h \ldots h.$$

Une forme $A_p h_p \ldots h_1$ est appelée p-forme *symétrique* si aucune permutation des vecteurs $h_1, \ldots, h_p$ ne fait changer sa valeur.

Les valeurs d'une forme multilinéaire symétrique $A_p h_p \ldots h_1$ *peuvent être trouvées à partir des valeurs de la forme correspondante de degré p* $A_p h \ldots h$. Voici une méthode, bien qu'elle ne soit peut-être pas la meilleure. Ayant les vecteurs $h_1, \ldots, h_p$, considérons la valeur de la forme de degré p $A_p h \ldots h$ sur le vecteur $h = h_1 + \ldots + h_p \in X$:

$$(\mathrm{I}) = A_p (h_1 + \ldots + h_p) \ldots (h_1 + \ldots + h_p) =$$

$$= \sum_{(k)} c_{k_1 \ldots k_p} A_p \underbrace{h_p \ldots h_p}_{k_p \text{ fois}} \ldots \underbrace{h_1 \ldots h_1}_{k_1 \text{ fois}},$$

où $k_1 + \ldots + k_p = p$, (k) désigne un ensemble ordonné des p nombres $k_1 \geqslant 0, \ldots, k_p \geqslant 0$, les coefficients $c_{k_1 \ldots k_p}$, qui se sont formés en ordonnant les variables indépendantes (ce qui est possible d'après la symétrie supposée de la forme) et en réduisant les termes semblables, sont des nombres entiers positifs. Remplaçons dans la forme (I) la variable h_1 par $2h_1$; cela fait apparaître dans chaque terme de la forme (I) un facteur 2^{k_1} qui prend sa valeur maximale 2^p sur le terme $A_p h_1 \ldots h_1$.

Désignons cette forme par (II). La différence (III) $= 2^p$ (I) $-$ (II) n'a plus de terme qui contienne p fois la variable h_1; les autres termes sont toujours précédés par des coefficients entiers positifs. En remplaçant dans la forme (III), encore une fois, h_1 par $2h_1$ nous obtenons la forme (IV) dont les coefficients diffèrent de ceux de (III) par les facteurs puissances de deux, le plus grand des facteurs étant 2^{p-1}. La différence (V) $= 2^{p-1}$ (III) $-$ (IV) n'a donc pas de termes contenant $p - 1$ fois la variable h_1. Après avoir effectué la même opération $p - 1$ fois, nous aboutissons à la forme (VI) dont chaque terme a un coefficient entier positif et ne contient pas la variable h_1 plus qu'une fois. En procédant de la façon analogue on peut chasser de la forme (VI) les termes contenant plus d'une fois les variables $h_2, \ldots, h_p$. Définitivement, on arrive à une forme (VII) à coefficients entiers positifs dont chaque terme contient chacune des variables $h_1, \ldots, h_p$ pas plus qu'une fois. Cela veut dire que la forme (VII) est composée d'un seul terme:

$$(\mathrm{VII}) = c_p A_p h_p \ldots h_1,$$

où c_p est un entier positif. Or, par construction, la forme (VII) est une combinaison linéaire des valeurs de la forme $A_p h^{(i)} \ldots h^{(i)}$

de degré p, pour certains vecteurs $h^{(i)}$ spécialement choisis (notamment, $h^{(1)} = h_1 + \ldots + h_p$, $h^{(2)} = 2h_1 + \ldots + h_p$, $h^{(3)} = 2^2 h_1 + \ldots + h_p$, etc.).

Cela démontre la proposition.

d. Conséquence. *Pour une p-forme symétrique* $A_p h_p \ldots h_1$, *la condition* $A_p h \ldots h = 0$ (*quel que soit* $h \in X$) *implique* $A_p = 0$.

En effet, il résulte de $A_p h \ldots h = 0$ et de *c* que $A_p h_p \ldots h_1 = 0$ pour n'importe quels $h_1, \ldots, h_p$ de X ; d'après *b* on a $\| A_p \| = 0$.

e. Conséquence. *Quel que soit p, il existe une constante* $C_p > 0$ *telle que, pour toute p-forme, on a*:

$$\| A_p \| = \sup_{|h_1| \leqslant 1, \ldots, |h_p| \leqslant 1} | A_p h_p \ldots h_1 | \leqslant C_p \sup_{|h| \leqslant 1} | Ah \ldots h |.$$

En effet, toute *p*-forme $A_p h_p \ldots h_1$ peut être mise, d'après *c*, sous la forme d'une combinaison linéaire des formes $A_p h^{(i)} \ldots h^{(i)}$ de degré *p* de façon que les vecteurs $h^{(i)}$ appartiennent à une boule fixe, d'où l'estimation

$$| A_p h_p \ldots h_1 | \leqslant C_p \sup_{|h| \leqslant 1} | A_p h \ldots h |.$$

Si l'on suit attentivement la marche de la démonstration de *c*, on peut évaluer la constante C_p. Le nombre de termes de la combinaison linéaire construite n'est pas supérieur à p^2, les vecteurs $h^{(i)}$ ne dépassent pas, en norme, $2^p p$, et les coefficients sont au plus égaux à 2^{p^3}, d'où

$$C_p \leqslant 2^{p^3} \cdot 2^p \cdot p^3.$$

f. La conséquence *d* peut être renforcée de la manière suivante: *Etant donnée une p-forme symétrique* $A_p h_p \ldots h_1$, *supposons que*

$$A_p h \ldots h = o(|h|^p)$$

pour $h \to 0$, *autrement dit que, pour tout* $\varepsilon > 0$, *il existe* $\delta > 0$ *tel que l'inégalité*

$$| A_p h \ldots h | \leqslant \varepsilon | h |^p \qquad (3)$$

a lieu du moment que $| h | < \delta$. *Alors* $A_p = 0$.

En effet, si $A_p \neq 0$, alors il existe d'après *d* un vecteur $h_0 \in X$ pour lequel $A_p h_0 \ldots h_0 = l \neq 0$. On a donc sur la demi-droite $h = t h_0$ $(0 < t < \infty)$:

$$A_p h \ldots h = t^p A_p h_0 \ldots h_0 = t^p l = | h |^p \frac{l}{| h_0 |^p},$$

ce qui contredit (3) ; c'est pourquoi $A_p = 0$.

2.25. Différentielles d'ordre supérieur.

a. Supposons qu'une fonction $y = f(x)$ $(G \subset X \to Y)$ soit *p* fois dérivable pour $x = a$. A l'aide de l'opérateur $f'(a)$: $X \to Y$ et d'un vecteur $h \in X$, on peut composer la forme linéaire

$$df(a) = f'(a) h.$$

C'est *la différentielle première de la fonction* $f(x)$ *pour* $x = a$ *qui correspond au déplacement* h (*1.23*). Ensuite, avec l'opérateur $f''(a)$: $X \to Y_1$ et deux vecteurs h_1 et h_2, on peut composer la forme bilinéaire $f''(a)\, h_2 h_1$. La forme quadratique qui lui correspond

$$f''(a)\, hh = d^2 f(a)$$

s'appelle *différentielle seconde de la fonction* $f(x)$ *pour* $x = a$ *correspondant au déplacement* h. Nous construisons ainsi toutes les différentielles, y compris la différentielle d'ordre p

$$d^p f(a) = f^{(p)}(a)\, h \ldots h$$

qu'on obtien tde la forme p-linéaire $f^{(p)}(a)\, h_p \ldots h_1$ pour $h_1 = \ldots$ $\ldots = h_p = h$.

b. Trouvons les expressions de ces différentielles pour une fonction $y = f(x)$: $R_n \to R_1$. En posant $h_1 = (dx_1^{(1)}, \ldots, dx_n^{(1)})$ nous avons dans ce cas:

$$df(a) = f'(a)\, h_1 = \sum_{i=1}^{n} \frac{\partial f(a)}{\partial x_i}\, dx_i^{(1)}.$$

Soit maintenant $h_2 = (dx_1^{(2)}, \ldots, dx_n^{(2)})$. Alors

$$f''(a)\, h_2 h_1 = \left(\sum_{i=1}^{n} \frac{\partial f'(a)}{\partial x_j}\, dx_j^{(2)} \right) h_1 = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f(a)}{\partial x_i\, \partial x_j}\, dx_j^{(2)}\, dx_i^{(1)},$$

ce qui ramène, pour $h_1 = h_2 = h = (dx_1, \ldots, dx_n)$, à la même expression de la différentielle seconde de la fonction $f(x)$ au point a que nous avons introduite dans *2.12*:

$$d^2 f(a) = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f(a)}{\partial x_i\, \partial x_j}\, dx_i\, dx_j.$$

D'une manière analogue, en considérant la différentielle p-ème de la fonction $f(x)$ pour $x = a$ nous retrouvons l'expression vue dans *2.12*.

2.26. Dérivées partielles d'ordre supérieur et différentielles partielles. Supposons que l'espace X soit mis sous forme de somme droite de q sous-espaces fermés $X_1 + \ldots + X_q$. Naturellement, tout vecteur $x \in X$ se décompose en une somme $x = x_1 + \ldots + x_q$, $x_i \in X_i$, $i = 1, \ldots, q$. Supposons que la fonction $f(x)$: $G \subset X \to Y$ soit p fois dérivable dans le domaine G.

Nous savons que l'opérateur $f'(a)$ agit de l'espace X dans l'espace Y; sa restriction sur le sous-espace X_i coïncide avec la *déri-*

vée partielle $\frac{\partial f(x)}{\partial x_i}$ *de la fonction* $f(x)$ par rapport au sous-espace X_i (*1.47d*). Cet opérateur, par sa définition même, est appliqué aux vecteurs $h_i \in X_i$. La fonction $\frac{\partial f(x)}{\partial x_i}$ est définie dans le domaine G et, de même que la fonction $f'(x)$, admet encore une dérivation ; sa dérivée partielle par rapport au sous-espace X_j est désignée par $\frac{\partial^2 f(x)}{\partial x_j \partial x_i}$. Ce dernier opérateur agit, d'une façon naturelle, sur les vecteurs $h_j \in X_j$. En continuant nous avons les définitions des *dérivées partielles de la forme* $\frac{\partial^p f(x)}{\partial x_{i_1} \dots \partial x_{i_p}}$. Pour ces opérateurs, sont définies les expressions $\frac{\partial^p f(x)}{\partial x_{i_p} \dots \partial x_{i_1}} h_{i_1} \dots h_{i_p}$, $h_{i_1} \in X_{i_1}$, ..., $h_{i_p} \in X_{i_p}$ qui s'appellent *différentielles partielles de la fonction* $f(x)$ *par rapport aux sous-espaces* $X_{i_1}, \dots, X_{i_p}$ et pour déplacements $h_{i_1}, \dots, h_{i_p}$.

Ces définitions généralisent celles des dérivées partielles d'ordre supérieur (*2.11*) et des différentielles (*2.12*) de la fonction d'un nombre fini de variables réelles ($X = R_n$).

§ 2. 3. Propriétés des dérivées d'ordre supérieur

2.31. a. Soit A un opérateur linéaire qui applique l'espace X dans l'espace Y_{p-1} (*2.24*), de sorte que l'on peut considérer la forme multilinéaire

$$Ax_p \dots x_1, \quad x_1, \dots, x_p \in X \tag{1}$$

et la forme correspondante $Ax \dots x$, $x \in X$ de degré p.

Théorème. *Si la forme* (1) *est symétrique, alors la fonction* $f(x) = Ax \dots x$ *est indéfiniment dérivable et*

$$\left.\begin{aligned} &f'(x) = pA\underbrace{x \dots x}_{p-1 \text{ fois}} \in Y_1, \\ &\dots\dots\dots\dots\dots\dots\dots\dots \\ &f^{(k)}(x) = p(p-1)\dots(p-k+1)\underbrace{Ax \dots x}_{p-k \text{ fois}} \in Y_k, \\ &\dots\dots\dots\dots\dots\dots\dots\dots \\ &f^{(p)}(x) = p!A \in Y_p, \\ &f^{(q)}(x) = 0 \quad (q > p). \end{aligned}\right\} \tag{2}$$

Démonstration. La forme $Ax_p \dots x_1$ étant multilinéaire, on a

$$f(x+h) - f(x) = A(x+h) \dots (x+h) - Ax \dots x =$$
$$= Ahx \dots x + Axh \dots x + Ax \dots xh + o(h),$$

et, puisque la forme $Ax_p \ldots x_1$ est symétrique, on a

$$f(x+h) - f(x) = pA\underbrace{x \ldots x}_{p-1 \text{ fois}}h + o(h);$$

par conséquent,

$$f'(x) = pA\underbrace{x \ldots x}_{p-1 \text{ fois}}.$$

On peut continuer par récurrence en supposant que les formules (2) soient valables pour la forme $A\underbrace{x \ldots x}_{p-1 \text{ fois}}$ et en se rappelant que, dans le cas $p = 1$, le résultat cherché est démontré dans *1.31a-b.*

Les propositions *b*, *c* et *d* qui suivent sont démontrées pour $p = 1$ dans *1.32a*, *b* et *c* respectivement. Leurs démonstrations pour n'importe quel $p = 1, 2, \ldots$ se réalisent aisément par récurrence.

b. *Si deux fonctions* $f(x): V \subset X \to Y$ *et* $g(x): V \subset X \to Y$ *sont dérivables p fois pour* $x = a \in V$, *il en est de même de la fonction* $s(x) = f(x) + g(x)$, *et l'on a*

$$s^{(p)}(a) = f^{(p)}(a) + g^{(p)}(a);$$

en d'autres termes, quel que soit $h \in X$, *on a*

$$s^{(p)}(a)\underbrace{h \ldots h}_{p \text{ fois}} = f^{(p)}(a)\underbrace{h \ldots h}_{p \text{ fois}} + g^{(p)}(a)\underbrace{h \ldots h}_{p \text{ fois}}.$$

c. *Si une fonction* $y(x): V \to Y$ *est dérivable p fois pour* $x = a \in V$, *et si A est un opérateur linéaire continu de Y dans un espace Z, alors la fonction* $z(x) = Ay(x)$ *est elle aussi p fois dérivable pour* $x = a$, *et l'on a* $z^{(p)}(a) = Ay^{(p)}(a)$ *ou bien, ce qui revient au même,*

$$z^{(p)}(a)\underbrace{h \ldots h}_{p \text{ fois}} = Ay^{(p)}(a)\underbrace{h \ldots h}_{p \text{ fois}}$$

pour tout $h \in X$.

d. *Soit Y la somme droite des sous-espaces* $Y_{(1)}, \ldots, Y_{(n)}$ *de sorte que toute fonction* $y(x): V \to Y$ *admet les composantes*

$$y_{(1)}(x): V \to Y_{(1)}, \ldots, y_{(n)}(x): V \to Y_{(n)}.$$

Si Y est un espace complet et que les sous-espaces $Y_{(1)}, \ldots, Y_{(n)}$ *sont fermés, alors, dès que la fonction* $y(x)$ *soit p fois dérivable pour* $x = a$, *il en est de même de toute composante* $y_{(j)}(x)$ $(j = 1, \ldots, n)$; *de plus on a*

$$y^{(p)}(x) = \{y_{(1)}^{(p)}(x), \ldots, y_{(n)}^{(p)}(x)\},$$

où $\{\ \}$ *désigne l'ensemble ordonné des composantes de la fonction* $y^{(p)}(x)\colon V \to Y_p$ *(2.21a) dans la décomposition naturelle de l'espace* Y_p *en somme droite (1.14g).*

$$Y_p = Y_{p\,(1)} + \ldots + Y_{p\,(n)}.$$

Inversement, *du moment que toutes les composantes* $y_{(1)}(x), \ldots, y_{(n)}(x)$ *soient p fois dérivables pour* $x = a$, *il en est de même de la fonction* $y(x)$.

2.32. Symétrie de la dérivée seconde.

a. Soit $y = f(x)$ $(G \subset X \to Y)$ une fonction deux fois dérivable. Composons la forme bilinéaire $f''(a)\,hk$ par rapport aux vecteurs h et k de l'espace X. Montrons que *cette forme bilinéaire est symétrique, c'est-à-dire que, pour n'importe quels vecteurs h et k, a lieu la relation*

$$f''(a)\,hk = f''(a)\,kh. \tag{1}$$

A cette fin considérons l'expression

$$w = f(a + h + k) - f(a + h) - f(a + k) + f(a).$$

On peut la traiter comme accroissement de la fonction

$$\Phi(x) = f(x + k) - f(x),$$

lorsque x varie de a à $a + h$. D'après le théorème de la moyenne *1.42d*, on a

$$|\Phi(a+h) - \Phi(a) - \Phi'(a)\,h| \leqslant \sup_{0\leqslant\theta\leqslant 1} |\Phi'(a+\theta h) - \Phi'(a)|\,|h|. \tag{2}$$

Pour $\varepsilon > 0$ donné, trouvons $\delta > 0$ tel que l'inégalité $|h| < \delta$ implique

$$|f'(a + h) - f'(a) - f''(a)\,h| \leqslant \varepsilon\,|h|.$$

Dans les formules qui suivent, posons $|h| \leqslant \delta/2$, $|k| \leqslant \delta/2$ et désignons par $\varepsilon_1, \varepsilon_2, \ldots$ des quantités (vecteurs, opérateurs) dont la norme soit inférieure à ε $(|h| + |k|)$.

Nous avons $\Phi'(x) = f'(x + k) - f'(x)$, donc

$$\begin{aligned}\Phi'(a + \theta h) &= f'(a + k + \theta h) - f'(a + \theta h) = \\ &= [f'(a + k + \theta h) - f'(a)] - [f'(a + \theta h) - f'(a)] = \\ &= [f''(a)(k + \theta h) + \varepsilon_1] - [f''(a)\,\theta h + \varepsilon_2] = f''(a)\,k + 2\varepsilon_3.\end{aligned}$$

D'une façon analogue, on a

$$\Phi'(a) = f'(a + k) - f'(a) = f''(a)\,k + \varepsilon_4,$$

donc,

$$\Phi'(a + \theta h) - \Phi'(a) = 3\varepsilon_5,$$

ce qui a pour conséquence

$$\sup_{0\leqslant\theta\leqslant 1} |\Phi'(a+\theta h) - \Phi'(a)|\,|h| \leqslant 3\,|\varepsilon_5|\,|h|,$$

c'est dire que le second membre dans (2) est un infiniment petit d'ordre supérieur par rapport à $(|h| + |k|)^2$.

D'autre part,

$$\Phi'(a)\,h = f''(a)\,kh + \varepsilon_4 h,$$

d'où la relation

$$|f(a+h+k) - f(a+h) - f(a+k) + f(a) - f''(a)\,kh| \leqslant \leqslant 4|\varepsilon_6||h|.$$

En permutant h et k on obtient

$$\|f''(a)\,hk - f''(a)\,kh\| \leqslant 8\varepsilon(|h| + |k|)^2.$$

Ainsi, pour tous les h avec $|h| \leqslant \delta/2$ et tous les k avec $|k| \leqslant \leqslant \delta/2$, on a

$$\sup_{h,\,k} |f''(a)\,hk - f''(a)\,kh| \leqslant 8\varepsilon\delta^2.$$

Quels que soient $h_0 \in X$, $k_0 \in X$, on obtient donc

$$\left|f''(a)\,\frac{\delta}{2}\,\frac{h_0}{|h_0|}\,\frac{\delta}{2}\,\frac{k_0}{|k_0|} - f''(a)\,\frac{\delta}{2}\,\frac{k_0}{|k_0|}\,\frac{\delta}{2}\,\frac{h_0}{|h_0|}\right| \leqslant 8\varepsilon\delta^2,$$

ou, ce qui est le même,

$$|f''(a)\,h_0k_0 - f''(a)\,k_0h_0| \leqslant 32\varepsilon|h_0||k_0|.$$

Comme $\varepsilon > 0$ est arbitraire, on a

$$f''(a)\,h_0k_0 - f''(a)\,k_0h_0 = 0,$$

ce qu'il fallait.

b. Symétrie des dérivées mixtes. Pour une fonction deux fois dérivable $y = f(x)$ $(G \subset X \to Y)$, donnée dans un domaine G de la somme droite $X = X_1 + X_2$ de deux espaces X_1 et X_2, sont définies (*2.26*) les dérivées partielles secondes

$$\frac{\partial^2 f(x)}{\partial x_1\,\partial x_2} \text{ et } \frac{\partial^2 f(x)}{\partial x_2\,\partial x_1}$$

qui coïncident avec les restrictions respectives de l'opérateur $f''(x)$ (*2.26*).

Ces opérateurs étant donnés, on peut considérer les expressions

$$\frac{\partial^2 f(x)}{\partial x_1\,\partial x_2}\,h_2h_1 \text{ et } \frac{\partial^2 f(x)}{\partial x_2\,\partial x_1}\,h_1h_2, \quad h_1 \in X_1,\ h_2 \in X_2$$

qui sont des vecteurs de l'espace Y.

Comme nous avons d'après a

$$f''(x)\,h_2h_1 = f''(x)\,h_1h_2,$$

quels que soient h_1 et h_2 de X, alors, en particulier, nous avons l'égalité

$$\frac{\partial^2 f(x)}{\partial x_1 \partial x_2} h_2 h_1 = \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} h_1 h_2 \tag{3}$$

pour $h_1 \in X_1$ et $h_2 \in X_2$ considérés.

c. Pour une fonction $f(x_1, x_2): R_2 \to Y$ des variables numériques x_1 et x_2 ($X_1 = X_2 = R_1$), les opérateurs $\frac{\partial^2 f(x)}{\partial x_2 \partial x_1}$ et $\frac{\partial^2 f(x)}{\partial x_2 \partial x_1}$ prennent, à tout point $x = \{x_1, x_2\}$, leurs valeurs dans le même espace Y, et l'égalité (3) montre que ces valeurs coïncident:

$$\frac{\partial^2 f(x)}{\partial x_1 \partial x_2} = \frac{\partial^2 f(x)}{\partial x_2 \partial x_1}. \tag{4}$$

La relation (4) est nécessairement valable s'il existe $f''(x)$, i. e. (*2.23*) si les fonctions $\frac{\partial f(x)}{\partial x_1}$, $\frac{\partial f(x)}{\partial x_2}$ sont dérivables. Cette condition sous-entend, en particulier, l'existence des dérivées $\frac{\partial^2 f(x)}{\partial x_1^2}$ et $\frac{\partial^2 f(x)}{\partial x_2^2}$; par conséquent, le théorème obtenu ici est d'un autre caractère que celui *2.11a* où la relation (4) s'établit en n'utilisant que les propriétés de $\frac{\partial^2 f(x)}{\partial x_1 \partial x_2}$ et $\frac{\partial^2 f(x)}{\partial x_2 \partial x_1}$, sans toucher les autres dérivées secondes. (En récompense, on y exige l'existence de ces dérivées dans un voisinage du point a et leur continuité au point a même; dans le théorème général a, nous évitons les suppositions de ce genre.)

d. Symétrie des dérivées d'ordre supérieur. Soit $y = f(x)$ ($G \subset X \to Y$) une fonction p fois dérivable ($p \geqslant 2$) pour $x = a$; montrons que *la forme*

$$f^{(p)}(a) h_p \ldots h_1, \quad h_1, \ldots, h_p \in X$$

est symétrique. Supposons que la proposition analogue soit valable pour toute fonction $p-1$ fois dérivable. Nous pouvons alors écrire

$$f^{(p)}(a) h_p \ldots h_1 = g^{(p-1)}(a) h_{p-1} \ldots h_1 \tag{5}$$

où $g(x) = f'(x) h_p$ est une fonction $(p-1)$-dérivable pour $x = a$. Par conséquent, d'après l'hypothèse de récurrence, on peut permuter les variables $h_{p-1}, \ldots, h_1$ dans l'expression (5). Or, on peut également écrire

$$f^{(p)}(a) h_p \ldots h_1 = w'(a) h_1$$

où $w(x) = f^{(p-1)}(x) h_p \ldots h_2$ est une fonction une fois dérivable pour $x = a$; on peut permuter ici n'importe quelles variables $h_p, \ldots, h_2$. Comme $p > 2$, on voit que chaque permutation des vecteurs $h_1, \ldots, h_p$ est admissible, ce qu'il fallait.

2.33. Dérivées d'ordre supérieur d'un produit généralisé. Soient, comme dans *1.34b*, $x(t): G \to X$ et $y(t): G \to Y$ deux fonctions dérivables données dans un domaine G d'un espace T et $\zeta(t) = \langle x(t), y(t)\rangle: G \to Z$ leur produit généralisé. Nous supposons cette fois que ces fonctions soient p-dérivables et démontrons qu'il en est de même de la fonction $\zeta(t)$.

Pour la dérivée première du produit $\langle x(t), y(t)\rangle$, nous avons obtenu dans *1.34b* la formule

$$\zeta'(t) = \langle x'(t), y(t)\rangle + \langle x(t), y'(t)\rangle,$$

les opérateurs dans le second membre étant définis comme suit:

$$\zeta'(t)\,dt = \langle x'(t)\,dt, y(t)\rangle + \langle x(t), y'(t)\,dt\rangle.$$

On voit donc que ces opérateurs sont eux aussi des produits généralisés, dérivables en t dans le cas $p \geqslant 2$. En appliquant d'une façon formelle la notation *1.34b* nous trouvons:

$$\zeta''(t) = \langle x''(t), y(t)\rangle + \langle x'(t), y'(t)\rangle_1 + \\ + \langle x'(t), y'(t)\rangle_2 + \langle x(t), y''(t)\rangle. \qquad (1)$$

Les indices 1 et 2 signifient que les termes auxquels ils sont attribués ont les sens différents, ce qu'il est facile de voir en utilisant les différentielles:

$$\zeta''(t)\,h_2h_1 = \langle x''(t)\,h_2h_1, y(t)\rangle + \langle x'(t)\,h_2, y'(t)\,h_1\rangle + \\ + \langle x'(t)\,h_1, y'(t)\,h_2\rangle + \langle x(t), y''(t)\,h_2h_1\rangle. \qquad (2)$$

Tous les quatre termes dans le second membre de (1) sont encore des produits généralisés, ce qui permet de continuer la dérivation pour $p > 2$. Après p dérivations nous aboutissons à la formule:

$$\zeta^{(p)}(t) = \langle x^{(p)}(t), y(t)\rangle + \langle x^{(p-1)}(t), y'(t)\rangle_1 + \dots \\ \dots + \langle x^{(p-1)}(t), y'(t)\rangle_p + \langle x^{(p-2)}(t), y''(t)\rangle_1 + \dots \\ \dots + \langle x^{(p-2)}(t), y''(t)\rangle_{\frac{p(p-1)}{2}} + \dots + \langle x(t), y^{(p)}(t)\rangle. \qquad (3)$$

Ici les termes de forme identique munis des indices différents ont les sens différents que l'on explicite avec la notation en différentielles:

$$\zeta^{(p)}(t)\,h_p \dots h_1 = \langle x^{(p)}(t)\,h_p \dots h_1, y(t)\rangle + \\ + \langle x^{(p-1)}(t)\,h_p \dots h_2, y'(t)\,h_1\rangle + \dots \\ \dots + \langle x^{(p-1)}(t)\,h_{p-1} \dots h_1, y'(t)\,h_p\rangle + \\ + \langle x^{(p-2)}(t)\,h_p \dots h_3, y''(t)\,h_2h_1\rangle + \dots \\ \dots + \langle x^{(p-2)}(t)\,h_{p-2} \dots h_1, y''(t)\,h_p\,h_{p-1}\rangle + \dots \\ \dots + \langle x(t), y^{(p)}(t)\,h_p \dots h_1\rangle. \qquad (4)$$

Si l'on compose la forme correspondante de degré p en posant $h_1 = \ldots = h_p = h$, alors, vu la symétrie des dérivées mixtes, on obtient une formule plus simple :

$$\zeta^{(p)}(t)\,h \ldots h = \langle x^{(p)}(t) \underbrace{h \ldots h}_{p \text{ fois}},\ y(t)\rangle +$$
$$+ p\,\langle x^{(p-1)}(t) \underbrace{h \ldots h}_{p-1 \text{ fois}},\ y'(t)\,h\rangle +$$
$$+ \frac{p(p-1)}{1\cdot 2}\langle x^{(p-2)}(t) \underbrace{h \ldots h}_{p-2 \text{ fois}}, y'(t)\,hh\rangle + \ldots$$
$$\ldots + \langle x(t),\ y^{(p)}(t) \underbrace{h \ldots h}_{p \text{ fois}}\rangle, \qquad (5)$$

soit, en notation symbolique,

$$\zeta^{(p)}(t) = \langle x^{(p)}(t),\ y(t)\rangle + p\,\langle x^{(p-1)}(t),\ y'(t)\rangle +$$
$$+ \frac{p(p-1)}{1\cdot 2}\langle x^{(p-2)}(t),\ y''(t)\rangle + \ldots + \langle x(t),\ y^{(p)}(t)\rangle \qquad (6)$$

(formule de Leibniz) ; il ne faut tout de même pas oublier que cette formule n'a le sens que dans les réalisations (4) et (5).

2.34. Dérivées d'ordre supérieur d'une fonction composée.

Théorème. *Soient $y = y(x)$ $(G \subset X \to Y)$ une fonction p fois dérivable pour $x = a$ et $z = z(y)$ $(W \subset Y \to Z)$ une fonction p fois dérivable pour $y = b = f(a) \in W$; alors la fonction composée $z[y(x)] = \zeta(x)$ $(U \subset X \to Z)$ est p fois dérivable pour $x = a$.*

Démonstration. Le cas $p = 1$ a été épuisé dans *1.33* ; supposons que le théorème ait lieu pour le numéro $p - 1 \geqslant 0$ et démontrons-le pour le numéro p. Conformément à *1.33*, la dérivée première de la fonction $\zeta(x)$ a la forme

$$\zeta'(x) = z'[y(x)]\,y'(x).$$

Le premier facteur est la composée de la fonction p fois dérivable $y(x)$ et de la fonction $p - 1$ fois dérivable $z'(y)$; par hypothèse de récurrence, ce facteur représente une fonction $p - 1$ fois dérivable en x. Le deuxième facteur est, par hypothèse, $p - 1$ fois dérivable en x. En vertu de *2.33*, tout le produit est $p - 1$ fois dérivable en x. Il en résulte que $\zeta(x)$ est une fonction p fois dérivable en x, ce qu'il fallait.

2.35. Dérivées d'ordre supérieur d'un opérateur inverse. Soient, comme dans *1.35b*, $x \in L(U, V)$ un opérateur inversible linéaire d'un espace U dans un espace V, et $x^{-1} : V \to U$ son opérateur inverse. Montrons que *la fonction x^{-1} a la dérivée en x de n'importe quel ordre p.*

Cela a été démontré pour $p = 1$ dans *1.35c*, on y a obtenu la formule

$$d(x^{-1}) = -x^{-1}hx^{-1}. \tag{1}$$

La dérivée $(x^{-1})'$ peut être mise sous la forme de produit généralisé :

$$(x^{-1})' = -\langle x^{-1}, x^{-1}\rangle$$

(au sens de (1), bien entendu). Si l'on suppose que la fonction x^{-1} soit p fois dérivable, alors, d'après *2.33*, la fonction $(x^{-1})'$ elle aussi sera p fois dérivable, c'est-à-dire que x^{-1} sera dérivable $p + 1$ fois. Puisque la proposition est juste pour $p = 1$, notre raisonnement montre qu'elle l'est pour tout $p = 1, 2, \ldots$, ce qu'il nous fallait.

2.36. Dérivées d'ordre supérieur d'une fonction implicite.

a. Théorème. *Supposons que l'hypothèse du théorème sur la fonction implicite 1.53 soit remplie : on a une fonction* $z = \Phi(x, y)$ $(V = \{x \in X,\ y \in Y :\ |x - a| < r,\ |y - b| < \rho\} \to Z)$, *la relation* $\Phi(a, b) = 0$ *a lieu, l'opérateur* $\frac{\partial \Phi(x, y)}{\partial y}$ *est continu en* x *et* y *et inversible pour* $x = a$, $y = b$. *Alors, dès que la fonction* $\Phi(x, y)$ *soit* p *fois dérivable au voisinage* V, *il en sera de même de la fonction implicite* $y(x)$ *qui représente la solution de l'équation* $\Phi(x, y(x)) = 0$, $y(a) = b$ *et existe en vertu du théorème 1.53.*

Démonstration. Le cas $p = 1$ est épuisé dans *1.54* ; on y a établi la formule

$$y'(x) = -\left[\frac{\partial \Phi(x, y)}{\partial y}\right]^{-1} \frac{\partial \Phi(x, y)}{\partial x}.$$

En supposant le théorème juste pour le numéro $p - 1 \geqslant 0$, démontrons-le pour le numéro p.

Le premier facteur est la composée de la fonction $x \to \{x, y(x)\}$ $p - 1$ fois dérivable par hypothèse de récurrence et par *2.31d* et de la fonction d'inversion (*2.35*) qui est indéfiniment dérivable ; le deuxième facteur est la composée de la même fonction $x \to \{x, y(x)\}$ et de la fonction $p - 1$ fois dérivable $\frac{\partial \Phi(x, y)}{\partial x}$. D'après *2.34*, les deux facteurs sont $p - 1$ fois dérivables ; d'après *2.33*, il en est de même de leur produit.

Il en résulte que $y'(x)$ est $p - 1$ fois dérivable, donc $y(x)$ est dérivable p fois, ce qu'il fallait.

b. En particulier, *la fonction inverse* $y = f(x)$ *définie par l'équation* $x = \varphi(y)$, $a = \varphi(b)$ *avec l'opérateur* $\varphi'(b)$ (*1.56*) *inversible, est* p *fois dérivable dès que la fonction* $\varphi(y)$ *le soit.*

2.37. a. Dans *1.84*, nous avons ramené le théorème sur la dérivabilité du point fixe d'une application contractante $A(u, \lambda)$ $(U \times \Lambda \to U)$ par rapport au paramètre λ au théorème *1.55* sur la

dérivabilité d'une fonction implicite. En utilisant à la place de *1.55* le théorème *2.36a* nous arrivons à l'extension suivante du théorème *1.84* au cas de dérivées d'ordre supérieur :

Théorème. *Soit* $A(u, \lambda)$ *une application contractante d'un domaine fermé* $U \subset X$ *dans lui-même qui représente une fonction p fois dérivable de* λ ; *alors le point fixe* $u = y(\lambda)$ *de l'application A est une fonction p fois dérivable de* λ.

b. Ce fait, à son tour, permet de généraliser convenablement le théorème *1.85*, à savoir : *si le second membre de l'équation différentielle*

$$\frac{dy(t, \lambda)}{dt} = \Phi(t, y(t, \lambda), \lambda) \tag{1}$$

et la condition initiale

$$y(t_0, \lambda) = \varphi(\lambda) \tag{2}$$

sont des fonctions p-dérivables du paramètre λ, *la solution* $y = y(t, \lambda)$ *de l'équation (1) avec la condition (2) est une fonction p-dérivable de* λ.

c. Les dérivées de la solution $y(t, \lambda)$ possèdent des propriétés analogues en ce qui concerne leur dérivabilité en λ. Ainsi, *lorsque la fonction* $\Phi(t, y, \lambda)$ *est, dans le domaine considéré, p fois dérivable par rapport à la variable* (y, λ), *la fonction* $\frac{dy(t, \lambda)}{dt}$ *est elle aussi p fois dérivable en* λ (d'après *b* et en raison de l'équation (1) même). On peut dire autant sur les autres dérivées de la solution $y(t, \lambda)$ dont les équations se déduisent en dérivant par rapport à t l'équation (1) sous l'hypothèse de la dérivabilité convenable de la fonction $\Phi(t, y, \lambda)$.

§ 2. 4. Dérivées par rapport aux champs vectoriels

2.41. Supposons qu'une fonction $\xi(x) : G \to X$ soit donnée dans un domaine G d'un espace normé X, autrement dit, qu'on ait un *champ vectoriel* $\xi(x)$. Pour toute fonction dérivable $\Phi(x) : G \to Y$, en tout point $x \in G$, on peut calculer la dérivée suivant le vecteur correspondant $\xi(x)$ (*1.27*) :

$$\xi * \Phi(x) = \Phi'(x) \cdot \xi(x). \tag{1}$$

On aura pour résultat une fonction de x, à valeurs dans Y ; elle sera dérivable dès qu'on suppose la fonction $\Phi(x)$ deux fois dérivable et le champ $\xi(x)$ dérivable.

Soit $\eta(x)$ un autre champ vectoriel dérivable dans le domaine G. Dérivons la fonction (1) suivant le champ $\eta(x)$:

$$\eta * (\xi * \Phi(x)) = \eta * (\Phi'\xi) = (\Phi'\xi)'\,\eta(x) =$$

$$= \Phi''\xi\eta(x) + \Phi'\xi'\eta(x). \tag{2}$$

Le premier terme à droite est symétrique en ξ et η (*2.32a*), le deuxième en général non ; si l'on fait changer de rôles ξ et η et que l'on effectue la soustraction des égalités correspondantes, on aura

$$\eta * (\xi * \Phi(x)) - (\xi * (\eta * \Phi(x)) = \Phi'(x) \cdot [\xi'\eta - \eta'\xi](x) = \\ = [\xi'\eta - \eta'\xi] * \Phi(x), \qquad (3)$$

de sorte que les dérivations suivant les vecteurs ξ et η ne commutent en général pas. (Bien entendu, la commutativité a lieu si ξ et η ne dépendent pas de x, de sorte que $\xi'(x) \equiv \equiv \eta'(x) \equiv 0$.)

Le crochet dans (3) représente un vecteur (champ vectoriel, si l'on tient compte de la dépendance de x) ; nous le désignons par $[\xi, \eta](x)$ ou, tout court, $[\xi, \eta]$. Ce vecteur est même appelé *crochet* des vecteurs ξ et η. Lorsque ces derniers sont constants, on a $[\xi, \eta] = 0$. Evidemment,

$$[\xi, \eta] = -[\eta, \xi]. \qquad (4)$$

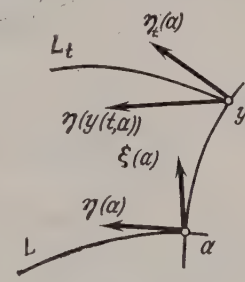


Fig. 2.4-1

2.42. Dans ce qui suit, nous considérons l'interprétation géométrique du crochet de deux vecteurs ainsi que certains problèmes géométriques. Soient $\xi(x)$ et $\eta(x)$ deux champs vectoriels dérivables donnés dans une boule $V = \{x \in X : |x - a| < r\}$. Envisageons l'équation différentielle

$$\frac{dy(t, x)}{dt} = \xi(y), \quad y(0, x) = x. \qquad (1)$$

Si la dérivée $\xi'(x)$ du champ $\xi(x)$ est continue, l'équation (1) admet une solution unique (*1.81b*) pour tout $x \in V$; cette solution $y(t, x)$, pour t, $|t| < \delta$, suffisamment petit, détermine le difféomorphisme π_t d'un voisinage du point a sur le voisinage correspondant du point $y(t, a)$ (*1.87a*). Soit $L = \{z(\tau), 0 \leqslant \tau \leqslant \tau_0\}$ une ligne dérivable issue, pour $\tau = 0$, du point a dans le sens du vecteur $\eta(a)$. Le difféomorphisme π_t la transforme en une ligne dérivable $L_t = \{y(t, z(\tau)), 0 \leqslant \tau \leqslant \tau_0\}$ issue (pour $\tau = 0$) du point $y(t, a)$. L'opérateur linéaire $\frac{\partial y(t, x)}{\partial x}$ transforme le vecteur $z'(0) = \eta(a)$ en un vecteur tangent à la courbe L_t pour $\tau = 0$ (*1.39d*) que nous désignons par $\eta_t(a)$; nous dirons que $\eta_t(a)$ est le *résultat du transport du vecteur* $\eta(a)$ *au point* $y(t, a)$ *par le champ* $\xi(x)$. Il y a maintenant deux vecteurs au point $y(t, a)$; le vecteur initial $\eta(y(t, a))$ et le vecteur $\eta_t(a)$, résultat du transport de $\eta(a)$ par le champ $\xi(x)$. La différence $\eta(y(t, a)) - \eta_t(a)$ montre l'écart entre le vecteur $\eta(y(t, a))$ et le vecteur $\eta_t(a)$. La vitesse de cet écart, i.e. la

quantité

$$\lim_{t\to 0}\frac{1}{t}\,[\eta\,(y\,(t,\,a))-\eta_t\,(a)],$$

s'appelle *dérivée de Lie du champ* $\eta\,(x)$ suivant le champ $\xi\,(x)$ au point $x=a$.

Calculons la dérivée de Lie. Nous avons:

$$\eta\,(y\,(t,\,a))=\eta\,(y\,(a)+ty'\,(a)+o\,(t))=\eta\,(a)+t\eta'\,(a)\,\xi\,(a)+o\,(t);$$

$$\eta_t\,(a)=\frac{\partial y\,(t,\,x)}{\partial x}\,\eta\,(a)=[E+t\xi'\,(a)+o\,(t)]\,\eta\,(a)=$$

$$=\eta\,(a)+t\xi'\,(a)\,\eta\,(a)+o\,(t),$$

d'où

$$\lim_{t\to 0}\frac{1}{t}\,[\eta\,(y\,(t,\,a))-\eta_t\,(a)]=\eta'\,(a)\,\xi\,(a)-\xi'\,(a)\,\eta\,(a).$$

Par conséquent, la dérivée de Lie du champ $\eta\,(x)$ suivant le champ $\xi\,(x)$ au point $x=a$ coïncide avec le crochet $[\eta,\ \xi]\,(a)$ défini dans *2.41*.

2.43. Ici nous donnons une construction directe de la courbe pour laquelle le vecteur $[\xi,\ \eta]\,(a)$ sera directeur. Pour le faire, effectuons la construction géométrique suivante (fig. 2.4-2). En fixant un s (suffisamment petit) partons du point a suivant la trajectoire du champ $\xi\,(x)$ définie par l'équation

$$\frac{dx}{dt}=\xi\,(x)$$

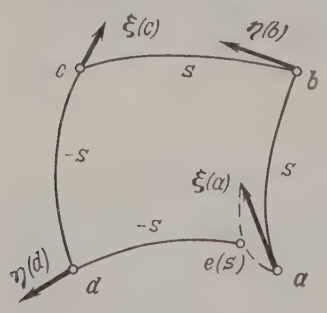


Fig. 2.4-2

et gagnons le point $b=x\,(s,\,a)$. Ensuite, partons-en et suivons la trajectoire du champ $\eta\,(y)$ définie par l'équation

$$\frac{dy\,(t)}{dt}=\eta\,(y(t))$$

jusqu'à gagner le point $c=y\,(s,\,b)$; puis suivons la trajectoire du champ $\xi\,(x)$ jusqu'au point $d=x\,(-s,\,c)$, enfin, suivons la trajectoire du champ $\eta\,(y)$ jusqu'au point $e=y\,(-s,\,d)$. Lorsque le paramètre s varie, le point $e=e\,(s)$ décrit dans le domaine V une courbe L partant du point a, pour $s=0$. Calculons la position du point $e\,(s)$ en tenant compte des infiniment petits de premier et deuxième ordres par rapport à s. Appliquons la formule *U16.17* (2):

$$x\,(t)-x\,(0)=tx'\,(0)+\frac{t^2}{2}x''\,(0)+o\,(t^2). \qquad (1)$$

Pour le premier de quatre arcs parcourus, on a :

$$x(0)=a,\ x'(0)=\xi(a),\ x''(0)=\frac{d}{dt}\,\xi[x(t)]\,|_{t=0}=$$

$$=\xi'(x)\cdot x'(t)\,|_{t=0}=\xi'(a)\,\xi(a), \qquad (2)$$

après quoi la formule (1) prend la forme

$$x(t)-a=t\xi(a)+\frac{t^2}{2}\,\xi'(a)\,\xi(a)+o(t^2).$$

En utilisant cette dernière formule et ses analogues, on trouve :

$$\left.\begin{aligned} b-a&=s\xi(a)+\frac{s^2}{2}\,\xi'(a)\,\xi(a)+o(s^2),\\ c-b&=s\eta(b)+\frac{s^2}{2}\,\eta'(b)\,\eta(b)+o(s^2),\\ d-c&=-s\xi(c)+\frac{s^2}{2}\,\xi'(c)\,\xi(c)+o(s^2),\\ e-d&=-s\eta(d)+\frac{s^2}{2}\,\eta'(d)\,\eta(d)+o(s^2). \end{aligned}\right\} \qquad (3)$$

Exprimons tous les coefficients de s et s^2 par $\xi(a)$, $\xi'(a)$, $\eta(a)$ et $\eta'(a)$. A cette fin, remarquons que

$$\left.\begin{aligned} &\eta(b)=\eta(a)+o(1),\ \eta'(b)=\eta'(a)+o(1),\\ &\xi(c)=\xi(a)+o(1),\ \xi'(c)=\xi'(a)+o(1),\\ &\eta(d)=\eta(a)+o(1),\ \eta'(d)=\eta'(a)+o(1); \end{aligned}\right\} \qquad (4)$$

donc, dans les termes du second degré dans (3), on peut remplacer b, c et d par a. En ce qui concerne les termes du premier degré, on doit y garder non seulement les termes constants de (4) mais aussi les termes proportionnels à s; composons donc les expressions

$$\left.\begin{aligned} &\eta(b)=\eta(a)+\eta'(a)\cdot s\xi(a)+o(s);\\ &\xi(c)=\xi(a)+\xi'(a)\,(s\xi(a)+s\eta(b))+o(s)=\\ &\quad=\xi(a)+\xi'(a)\,(s\xi(a)+s\eta(a))+o(s);\\ &\eta(d)=\eta(a)+\eta'(a)\,(s\xi(a)+s\eta(b)-s\xi(c))+o(s)=\\ &\quad=\eta(a)+\eta'(a)\,s\eta(a)+o(s). \end{aligned}\right\} \qquad (5)$$

En portant (5) et (4) dans (3) et additionnant, on obtient

$$e-a=s^2\Big[\frac{1}{2}\,\xi'(a)\,\xi(a)+\eta'(a)\,\xi(a)+\frac{1}{2}\,\eta'(a)\,\eta(a)-$$

$$-\xi'(a)\,\xi(a)-\xi'(a)\,\eta(a)+\frac{1}{2}\,\xi'(a)\,\xi(a)-\eta'(a)\,\eta(a)+$$

$$+\frac{1}{2}\,\eta'(a)\,\eta(a)\Big]+o(s^2)=s^2[\eta'(a)\,\xi(a)-\xi'(a)\,\eta(a)]+$$

$$+o(s^2)=s^2[\eta,\xi](a)+o(s^2). \qquad (6)$$

Si l'on choisit, sur la courbe L, le paramètre s^2 au lieu de s, le vecteur directeur de cette courbe au point a sera juste le vecteur $[\eta, \xi](a)$, en vertu de (6).

2.44. Exprimons le crochet de deux champs vectoriels ξ et η dans R_n par leurs composantes. Soit, dans une base fixe $e_1, \ldots, e_n$, $\xi(x) = \sum_{i=1}^{n} \xi_i(x)\, e_i$ et $\eta(x) = \sum_{i=1}^{n} \eta_i(x)\, e_i$. La dérivée $\xi'(x)$ est l'opérateur linéaire $R_n \to R_n$ donné dans la même base par la matrice $\left\| \frac{\partial \xi_i(x)}{\partial x_j} \right\|$ (*1.25a*) ; d'une façon analogue, la dérivée $\eta'(x)$ est l'opérateur linéaire $R_n \to R_n$ donné dans la même base par la matrice $\left\| \frac{\partial \eta_i(x)}{\partial x_j} \right\|$. Par conséquent, la i-ème coordonnée du vecteur $[\xi, \eta]$ a la forme

$$[\xi, \eta]_i = (\xi'(x)\,\eta(x) - \eta'(x)\,\xi(x))_i = \sum_{j=1}^{n} \left(\frac{\partial \xi_i}{\partial x_j}\,\eta_j - \frac{\partial \eta_i}{\partial x_j}\,\xi_j \right). \qquad (1)$$

Il est à noter que la formule (1) a lieu dans tout système de coordonnées curvilignes

$$y_i = y_i(x_1, \ldots, x_n), \qquad i = 1, \ldots, n,$$

où les fonctions $y_i(x)$ admettent les dérivées continues du premier et du deuxième ordres. Pour le démontrer, désignons, comme dans *1.56d*, $p_{ij} = \frac{\partial y_j(x)}{\partial x_i}$, $q_{ij} = \frac{\partial x_j(x)}{\partial y_i}$. Les matrices $\| p_{ij} \|$ et $\| q_{ij} \|$ sont mutuellement inverses, donc $\sum_{j=1}^{n} p_{ij} q_{jh} = \delta_{ih}$. Si ξ_i, η_i sont les composantes des vecteurs ξ et η dans la base initiale $e_1, \ldots, e_n$, et λ_i, μ_i leurs composantes dans la base locale $g_1(a), \ldots, g_n(a)$ au point a, alors, d'après ce qu'on a vu dans *1.56d*, pour tout vecteur $\theta = \sum \theta_i e_i = \sum \tau_j g_j$, on a

$$\theta_i = \sum_{s=1}^{n} q_{si} \tau_s, \quad \tau_j = \sum_{i=1}^{n} p_{ij} \theta_i.$$

Ensuite, quelle que soit une fonction $\Phi(x)$, sa dérivée par rapport à x_k s'écrit $\sum_{l=1}^{n} \frac{\partial \Phi}{\partial y_l} \frac{\partial y_l}{\partial x_k} = \sum_{l=1}^{n} p_{kl} \frac{\partial \Phi}{\partial y_l}$. Soit, enfin, $[\xi, \eta] = \sum_{i=1}^{n} [\xi, \eta]_i\, e_i = \sum_{j=1}^{n} \zeta_j g_j$. Avec toutes les formules obtenues, on a :

$$\zeta_j = \sum_{i=1}^{n} p_{ij} [\xi, \eta]_i = \sum_{i,\,k} p_{ij} \left(\frac{\partial \xi_i}{\partial x_k}\,\eta_k - \frac{\partial \eta_i}{\partial x_k}\,\xi_k \right).$$

Exprimons d'abord par les nouvelles coordonnées les premiers termes :

$$\sum_{i,\,k} p_{ij}\eta_k \frac{\partial \xi_i}{\partial x_k} = \sum_{i,\,k} p_{ij} \sum_r q_{rk}\mu_r \sum_l \frac{\partial}{\partial y_l}\Big(\sum_s q_{si}\lambda_s\Big)\cdot p_{kl} =$$

$$= \sum_{i,\,k,\,r,\,l,\,s} p_{ij}q_{rk}\mu_r p_{kl}q_{si}\frac{\partial \lambda_s}{\partial y_l} +$$

$$+ \sum_{i,\,k,\,r,\,l,\,s} p_{ij}q_{rk}\mu_r p_{kl}\lambda_s \frac{\partial q_{si}}{\partial y_l}.$$

Comme $\sum_i p_{ij}q_{si} = \delta_{js}$, on n'aura que le terme avec $s = j$ en effectuant la sommation sur s ; de même, en vertu de $\sum_k q_{rk}p_{kl} = \delta_{rl}$, il ne restera que le terme avec $r = l$. Définitivement, on a :

$$\sum_{i,\,k} p_{ij}\eta_k \frac{\partial \xi_i}{\partial x_k} = \sum_l \mu_l \frac{\partial \lambda_j}{\partial y_l} + \sum_{i,\,l,\,s} p_{ij}\mu_l\lambda_s \frac{\partial q_{si}}{\partial y_l}. \qquad (2)$$

Notons que $\frac{\partial q_{si}}{\partial y_l} = \frac{\partial^2 x_i}{\partial y_l \partial y_s}$. Par conséquent, étant donnée la symétrie des dérivées deuxièmes (*2.11b*), la dernière somme est symétrique en ξ et η. En remplaçant ξ, η par η, ξ, i.e. μ_l, λ_s par λ_l, μ_s, et soustrayant l'égalité obtenue de (2) nous aboutissons à la relation

$$\zeta_j = \sum_{i,\,k} p_{ij}\Big(\frac{\partial \xi_i}{\partial x_k}\eta_k - \frac{\partial \eta_i}{\partial x_k}\xi_k\Big) = \sum_{l=1}^{n}\Big(\frac{\partial \lambda_j}{\partial y_l}\mu_l - \frac{\partial \mu_j}{\partial y_l}\lambda_l\Big),$$

ce qu'il nous fallait.

2.45. Supposons que, dans un domaine $G \subset X = R_n$, il y ait m champs vectoriels $\xi^1, \ldots, \xi^m$ linéairement indépendants en tout point $x \in G$. Nous tenons à choisir, au voisinage d'un point $a \in G$, un nouveau système (local) de coordonnées tel que les composantes des champs $\xi^1, \ldots, \xi^m$ aient la plus simple forme possible. Posons $a = 0$; choisissons et fixons un sous-espace H dont la somme droite avec le sous-espace R_m engendré par les vecteurs $\xi^1(0), \ldots, \xi^m(0)$ représente tout l'espace X. Soit $x^j(t_0, p)$ la solution de l'équation $\frac{dx(t)}{dt} = \xi^j(x)$, $x(0) = p$, pour $t = t_0$. Considérons une boule $H_\rho = \{h \in H, \ |h| < \rho\}$ dans le sous-espace H, et le parallélépipède $R_{m\delta} = \{x \in R_m : \ x = \sum_{k=1}^{m} \alpha_k \xi^k(0), \ |\alpha_k| < \delta\}$ dans le sous-espace R_m. Pour $|t|$ et $|p|$ suffisamment petits, $x^j(t, p)$ est une fonction dérivable de ses variables. Faisons maintenant correspondre à tout ensemble $(t_1, \ldots, t_m, h_{m+1}, \ldots, h_n)$ $|t_j| < \delta$, $|h_r| < \rho$

un point $x \in G$ selon la règle suivante:

$$x^1 = x^1(t_1, h);\ x^2 = x^2(t_2, x_1);\ \ldots;\ x^{m-1} = x^{m-1}(t_{m-1}, x_{m-2});$$
$$x^m = x^m(t_m, x_{m-1}).$$

Autrement dit, pour déterminer le point x, nous partons du point h en suivant la trajectoire du champ ξ^1 jusqu'au point x^1 qui répond à la valeur t_1 du paramètre; puis nous suivons la trajectoire du champ ξ^2 jusqu'au point x^2 qui répond à la valeur t_2 du paramètre, et ainsi de suite jusqu'au point $x = x^m(t_m, x_{m-1})$.

L'application $x(t_1, \ldots, t_m, h_{m+1}, \ldots, h_n)$ ainsi obtenue est dérivable d'après ce qu'on a vu plus haut; désignons-la par $\pi(x)$. L'application $\pi(x)$ transforme chaque vecteur de H en lui-même, elle conserve également les vecteurs $\xi^1(0), \ldots, \xi^m(0)$, donc représente une application identique; il en résulte (*1.56c*) que l'application $\pi(x)$ est inversible au voisinage du point 0 et que les quantités $t_1, \ldots, t_m, h_{m+1}, \ldots, h_n$ peuvent servir de nouvelles coordonnées au voisinage de 0. Voyons comment sont les composantes des vecteurs $\xi^1, \ldots, \xi^m$ à ce nouveau système de coordonnées.

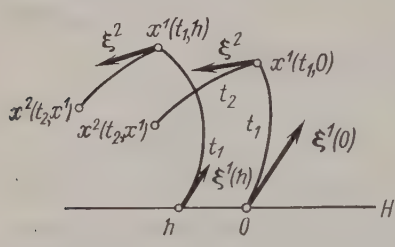


Fig. 2.4-3

Soit S_1 l'ensemble des points de la forme $x(t_1, h)$ dans le domaine de définition du difféomorphisme π. Cet ensemble est, évidemment, de dimension $n - m + 1$, et ses points sont décrits dans le nouveau système de coordonnées par l'expression $\{t_1, 0, \ldots, 0, h_{m+1}, \ldots, h_n\}$. Si l'on y fait varier seule la coordonnée t_1, on obtient, par construction, l'une des trajectoires du champ ξ^1. Le vecteur ξ^1, en tant que vecteur directeur de cette courbe, a donc, dans le nouveau système de coordonnées, la forme (*1.56d*):

$$\xi^1 = \{1, 0, \ldots, 0\} \qquad (\text{sur } S_1).$$

L'ensemble S_2 des points de la forme $x(t_2, x_1)$, $x_1 \in S_1$ est de dimension $n - m + 2$; dans le nouveau système ses points sont décrits par l'expression $\{t_1, t_2, 0, \ldots, 0, h_{m+1}, \ldots, h_n\}$. Si l'on y fait varier seule la coordonnée t_2 on obtient l'une des trajectoires du champ ξ^2. Par conséquent, le vecteur ξ^2, en tant que vecteur directeur de cette trajectoire, est:

$$\xi^2 = \{0, 1, 0, \ldots, 0\} \qquad (\text{sur } S_2).$$

En continuant de la même manière nous voyons que l'ensemble S_k des points de la forme $x(t_k, x_{k-1})$, où $x_{k-1} \in S_{k-1}$, est de dimension $n - m + k$; sur cet ensemble le vecteur ξ^k a, en nouvelles

coordonnées, la forme

$$\xi^k = \{0, 0, \ldots, \underset{k}{1}, \ldots, 0\} \qquad (\text{sur } S_k). \tag{1}$$

Notons que les composantes d'un vecteur ξ^k ne sont ici trouvées que sur S_k. Dans le seul cas $k = m$, où l'ensemble S_k est de dimension n et coïncide en réalité avec le voisinage considéré du point 0, les composantes du vecteur ξ^m sont connues partout dans ce voisinage :

$$\xi^m = \{0, \ldots, 0, 1, 0, \ldots, 0\}.$$

Il résulte de la formule (1) que, sur S_k :

$$\frac{\partial \xi^k}{\partial t_j} = 0 \ (j = 1, \ldots, k), \quad \frac{\partial \xi^k}{\partial h_r} = 0 \ (r = m+1, \ldots, n). \tag{2}$$

On en tire l'expression suivante pour $[\xi^j, \xi^k]$:

$$[\xi^j, \xi^k] \equiv \sum_{i=1}^{m} \left(\frac{\partial \xi^j}{\partial t_i} \xi_i^k - \frac{\partial \xi^k}{\partial t_i} \xi_i^j \right) + \sum_{r=m+1}^{n} \left(\frac{\partial \xi^j}{\partial h_r} \xi_r^k - \frac{\partial \xi^k}{\partial h_r} \xi_r^j \right) =$$

$$= \frac{\partial \xi^j}{\partial t_k} - \sum_{i=k+1}^{m} \frac{\partial \xi^k}{\partial t_i} \xi_i^j. \tag{3}$$

2.46. Existe-t-il des surfaces enveloppes de m champs vectoriels?

a. Soient $\xi^1(x), \ldots, \xi^m(x)$ m champs vectoriels donnés dans un domaine $G \subset X = R_n$ et linéairement indépendants en tout point $x \in G$. Désignons par $T(x)$ la variété linéaire engendrée par ces champs au point x. Une surface P de dimension m se trouvant dans le domaine G (*1.39e*) est dite *enveloppe pour la famille des champs* $\xi^1(x), \ldots, \xi^m(x)$ si la variété linéaire $\pi(x)$ tangente à la surface P (*1.39e*) à tout son point x coïncide avec la variété $T(x)$. On pose les questions suivantes : est-il possible de faire passer une surface enveloppe par un point donné $x \in G$? si oui, est-elle unique?

Soit $m = 1$, de sorte qu'il s'agisse d'un seul champ vectoriel $\xi^1(x) = \xi(x)$. Dans ce cas la notion de surface enveloppe se réduit à celle de trajectoire du champ $\xi(x)$; si le champ $\xi(x)$ possède la dérivée continue, les réponses aux deux questions ci-posées s'avèrent positives en vertu de théorèmes fondamentaux de la théorie des équations différentielles (*1.87*). Pour $m > 1$, la dérivabilité continue des champs $\xi^1(x), \ldots, \xi^m(x)$ n'est pas suffisante. Le théorème suivant a lieu :

Théorème. *Supposons que les champs* $\xi^1(x), \ldots, \xi^m(x)$ *possèdent leurs dérivées secondes continues dans le domaine G. Alors, pour qu'il existe une surface enveloppe passant par un point donné*

quelconque $a \in G$, *il faut et il suffit que* $[\xi^i(x), \xi^j(x)] \in T(x)$ *quel que soit* $x \in G$. *Cette condition étant remplie, la surface enveloppe passant par le point* $a \in G$ *est unique.*

La démonstration du théorème sera donnée dans *c* et *d*.

b. Lemme. *Soient* $P = \{x \in R_n : x = \Phi(u),\ u \in Q \subset R_m\}$ *une surface dans un domaine* $G \subset R_n$ *et* $\xi(x)$ *un champ vectoriel tangent à la surface* P *à tout point* $x \in P$. *Alors, dès que* $\Phi(u)$ *et* $\xi(x)$ *soient continûment dérivables, chaque trajectoire du champ* $\xi(x)$ *qui passe par un point* $a = \Phi(u_1^0, \ldots, u_m^0) \in P$ *se trouve entièrement sur* P.

Démonstration. Pour tout point x de la surface P, la variété tangente $\pi(x)$ admet une base composée des vecteurs $\frac{\partial x}{\partial u_1}, \ldots, \frac{\partial x}{\partial u_m}$ (*1.39e*). Le vecteur $\xi(x)$ se trouvant, par hypothèse, dans $\pi(x)$, on peut le développer suivant cette base:

$$\xi(x) = \sum_{j=1}^{m} \varphi_j(x) \frac{\partial x}{\partial u_j} = \sum_{j=1}^{m} \varphi_j(x(u)) \frac{\partial x}{\partial u_j} = \sum_{j=1}^{m} \psi_j(u) \frac{\partial x}{\partial u_j}.$$

Les fonctions $\varphi_j(x)$ sont continûment dérivables de même que les fonctions $\xi_i(x)$ composantes du champ $\xi(x)$ dans la base initiale $e_1, \ldots, e_n$ de l'espace X et que les fonctions $g_{ij}(x)$ composantes des vecteurs $\frac{\partial x}{\partial u_j}$ $(j = 1, \ldots, m)$ dans la même base; cela résulte des égalités

$$\xi = \sum_{i=1}^{n} \xi_i e_i,\ \frac{\partial x}{\partial u_j} = \sum_{i=1}^{n} g_{ij} e_i,\ \xi = \sum_{j=1}^{m} \varphi_j \frac{\partial x}{\partial u_j} = \sum_{i,j} g_{ij} \varphi_j e_i,\ \xi_i = \sum_{j=1}^{m} g_{ij} \varphi_j$$

après avoir choisi dans la dernière famille d'égalités m linéairement indépendantes et résolu le système obtenu selon la règle de Cramer. Il en est de même des fonctions $\psi_j(u)$.

Envisageons le système d'équations:

$$\frac{du_j(t)}{dt} = \psi_j(u_1, \ldots, u_m), \quad u_j(0) = u_j^0.$$

D'après ce qu'on a vu, il admet l'unique solution $u = u(t)$. Il lui répond une courbe $L \subset P$ qui est tangente à tout son point au vecteur correspondant $\xi(x)$. Or une telle courbe dans l'espace R_n représente une trajectoire du champ $\xi(x)$; d'après le théorème d'unicité *1.87*, la trajectoire passant par le point a est unique, donc coïncide avec la courbe $L \subset P$, ce qui achève la démonstration.

c. Démonstration du théorème *a* (nécessité). Soit P une surface enveloppe des champs $\xi^1(x), \ldots, \xi^m(x)$ passant par le point donné $a \in G$. On a vu dans *2.43* que le vecteur

$[\xi^i(a), \xi^j(a)]$ est un vecteur directeur pour une courbe L que l'on obtient en construisant une chaîne d'arcs de trajectoires des champs $\xi^i(x)$ et $\xi^j(x)$. Comme, par hypothèse, ces champs sont tangents à la surface P et que l'origine du premier arc est le point $a \in P$, toute la chaîne se trouve sur la surface P, d'après le lemme b ; il en découle que le vecteur $[\xi^i(a), \xi^j(a)]$ est lui aussi tangent à la surface P, donc appartient à $T(a)$, ce qu'il fallait.

d. Démonstration du théorème a (suffisance). Etant donnés les champs $\xi^1(x), \ldots, \xi^m(x)$, considérons au voisinage du point $a \in G$ le système spécial de coordonnées $t_1, \ldots, t_m, h_{m+1}, \ldots, h_n$ qu'on a introduit dans *2.45*. Dans ce système, comme dans n'importe quel système de coordonnées, les composantes $\xi_1^j(x), \ldots, \xi_n^j(x)$ des champs $\xi^j(x)$ $(j = 1, \ldots, m)$ vérifient les relations

$$[\xi^i, \xi^j] = \sum_{s=1}^{m} C_s^{ij}(x)\,\xi^s, \tag{1}$$

les coefficients $C_s^{ij}(x)$ étant continus et définis d'une façon univoque. Sur l'ensemble $S_2 = \{t_1, t_2, 0, \ldots, 0, h_{m+1}, \ldots, h_n\}$ on peut écrire, en utilisant la formule *2.45* (3) :

$$[\xi^1, \xi^2]_k \equiv \frac{\partial \xi_k^1}{\partial t_2} - \sum_{i=3}^{m} \frac{\partial \xi_k^2}{\partial t_i}\,\xi_i^1 = \sum_{r=1}^{m} C_r^{12}\xi_k^r. \tag{2}$$

Nous utilisons ces équations pour les valeurs $k = 3, \ldots, n$. De *2.45* (2), il résulte encore que, pour les mêmes valeurs $k = 3, \ldots, n$, sur S_2, on a les relations :

$$\frac{\partial \xi_k^2}{\partial t_2} = \frac{\partial \xi_k^3}{\partial t_2} = \ldots = \frac{\partial \xi_k^m}{\partial t_2} = 0. \tag{3}$$

On peut considérer les équations (2)-(3) comme système de $m(n-2)$ équations différentielles (avec la variable indépendante t_2) par rapport aux $m(n-2)$ inconnues $\xi_k^1, \ldots, \xi_k^m$ $(k = 3, \ldots, n)$, les coefficients $\frac{\partial \xi_k^2}{\partial t_i}$ et C_r^{12} étant continus. Pour $t_2 = 0$, i.e. sur S_1, toutes les fonctions inconnues s'annulent. Par conséquent, en vertu du théorème d'unicité, elles s'annulent pour toutes les valeurs admissibles de t_2, i.e. sur tout le S_2. Ainsi, sur S_2, les composantes du vecteur ξ^1 ont, dans l'hypothèse du théorème, la forme

$$\xi^1 = \{\xi_1^1, \xi_2^1, 0, \ldots, 0\}. \tag{4}$$

Par conséquent, l'expression $[\xi^1, \xi^2]$ a sur S_2 une forme plus simple que celle qui résulte de *2.45* (3), notamment :

$$[\xi^1, \xi^2] = \frac{\partial \xi^1}{\partial t_2}. \tag{5}$$

Ensuite, sur l'ensemble $S_3 = \{t_1, t_2, t_3, 0, \ldots, 0, h_{m+1}, \ldots, h_n\}$, considérons d'une manière analogue le système suivant de m $(n - 3)$ équations différentielles ordinaires avec la variable indépendante t_3 et m $(n - 3)$ fonctions inconnues $\xi_k^1, \ldots, \xi_k^m$ $(k = 4, \ldots, n)$:

$$\left.\begin{aligned} [\xi^1, \xi^3]_k &\equiv \frac{\partial \xi_k^1}{\partial t_3} - \sum_{i=4}^{m} \frac{\partial \xi_k^3}{\partial t_i} \xi_i^1 = \sum_{r=1}^{m} C_r^{13} \xi_k^r, \\ [\xi^2, \xi^3]_k &\equiv \frac{\partial \xi_k^2}{\partial t_3} - \sum_{i=4}^{m} \frac{\partial \xi_k^3}{\partial t_i} \xi_i^2 = \sum_{r=1}^{m} C_r^{23} \xi_k^r, \\ \frac{\partial \xi_k^3}{\partial t_3} &= \frac{\partial \xi_k^4}{\partial t_3} = \ldots = \frac{\partial \xi_k^m}{\partial t_3} = 0, \end{aligned}\right\} \tag{6}$$

les coefficients du système étant toujours continus. Pour $t_3 = 0$, i.e. sur S_2, toutes les fonctions inconnues s'annulent d'après *2.45* et les résultats précédents. En vertu du théorème d'unicité, elles sont nulles sur tout le S_3. Ainsi, sur S_3, les composantes des vecteurs ξ^1 et ξ^2 ont la forme

$$\xi^1 = \{\xi_1^1, \xi_2^1, \xi_3^1, 0, \ldots, 0\},$$
$$\xi^2 = \{\xi_1^2, \xi_2^2, \xi_3^2, 0, \ldots, 0\}.$$

Par conséquent, les expressions $[\xi^1, \xi^3]$ et $[\xi^2, \xi^3]$ prennent sur S_3 une forme plus simple:

$$[\xi^1, \xi^3] = \frac{\partial \xi^1}{\partial t_3}, \quad [\xi^2, \xi^3] = \frac{\partial \xi^2}{\partial t_3}. \tag{7}$$

En continuant par analogie on voit que, sur l'ensemble S_k $(k = 1, \ldots, m)$, toutes les composantes des vecteurs $\xi^1, \ldots, \xi^k$, à partir de $(k - 1)$-ème, s'annulent, et les expressions $[\xi^i, \xi^k]$ $(i < k)$ prennent la forme

$$[\xi^i, \xi^k] = \frac{\partial \xi^i}{\partial t_k} \quad (i = 1, \ldots, k-1). \tag{8}$$

Pour $k = m$, l'ensemble S_m coïncide avec tout un voisinage du point a. Donc, *dans un voisinage du point a, les composantes des vecteurs* $\xi^1, \ldots, \xi^m$, *à partir de la* $(m + 1)$*-ème, s'annulent.* Or, la représentation $\xi^k = \{\xi_1^k, \ldots, \xi_m^k, 0, \ldots, 0\}$ veut dire que *le vecteur* ξ^k *est tangent à chaque surface* $P = \{t_1, \ldots, t_m, h_{m+1}, \ldots, h_n\}$ *où les paramètres* $t_1, \ldots, t_m$ *varient, les quantités* $h_{m+1}, \ldots, h_n$ *étant fixes*: notamment, il est tangent au point $a_0 = (t_1^0, \ldots, t_m^0, h_{m+1}^0, \ldots, h_n^0) \in P$ à la courbe $L = \{t_1^0 + \tau \xi_1^k(a_0), \ldots, t_m^0 + \tau \xi_m^k(a_0), h_{m+1}^0, \ldots, h_n^0\}$, $|\tau| \leqslant \delta$ se trouvant sur la surface P. Nous voyons que ces surfaces sont m-dimensionnelles et passent par chaque point au voisinage du point a. Par conséquent, l'existence des surfaces enveloppes est démontrée.

Il reste à établir l'unicité de la surface enveloppe passant par le point donné a_0. Supposons qu'il y ait une autre surface enveloppe $\widetilde{P}$ donnée par une fonction $x = \widetilde{x}(u)$ continûment dérivable qui ne coïncide avec la surface P dans aucun voisinage du point a_0. Choisissons sur la surface $\widetilde{P}$ une courbe $L = \{x = x(\tau)\}$ qui passe par le point a et ne se trouve pas entièrement sur P. L'équation de cette courbe en nouvelles coordonnées est de la forme:

$$x = x(\tau) = \{t_1(\tau), \ldots, t_m(\tau), h_{m+1}(\tau), \ldots, h_n(\tau)\},$$

avec pour vecteur directeur

$$x'(\tau) = \{t_1'(\tau), \ldots, t_m'(\tau), h_{m+1}'(\tau), \ldots, h_n'(\tau)\}.$$

Or, par hypothèse, la courbe L est tangente en tout point x à la variété linéaire $T(x)$ engendrée par les vecteurs $\xi^1(x), \ldots, \xi^m(x)$. Nous avons vu que les $n - m$ dernières composantes de ces vecteurs, en nouvelles coordonnées, sont nulles; il en est donc de même du vecteur $x'(\tau)$. Ainsi, $h_{m+1}'(\tau) = \ldots = h_n'(\tau) = 0$, c'est-à-dire que les fonctions $h_{m+1}(\tau), \ldots, h_n(\tau)$ sont constantes sur la courbe L; alors la courbe L se trouve sur la surface $P(a)$ en contradiction avec l'hypothèse. Le théorème est démontré.

2.47. Existe-t-il des systèmes de coordonnées avec m vecteurs directeurs donnés? Soient toujours $\xi^1(x), \ldots, \xi^m(x)$ des champs vectoriels donnés dans un domaine $G \subset X = R_n$ et linéairement indépendants à tout point $x \in G$. On pose la question suivante: existe-t-il dans le domaine G, ou au moins dans un voisinage d'un point donné $a \in G$, un tel système de coordonnées (curvilignes) que les vecteurs $\xi^1(x), \ldots, \ldots, \xi^m(x)$ coïncident avec les vecteurs directeurs de m lignes de coordonnées?

Signalons d'abord une condition nécessaire de la réponse positive. Soit $(t_1, \ldots, t_m, h_{m+1}, \ldots, h_n)$ un système de coordonnées cherché, le vecteur $\xi^j(x)$ étant vecteur directeur de la ligne où seule la coordonnée t_j varie et les autres sont fixes. Dans ce système le vecteur $\xi^j(x)$ a la forme

$$\xi^j(x) = (0, \ldots, \underset{j}{1}, \ldots, 0).$$

En calculant dans ce système $[\xi^i, \xi^j]$ et appliquant la formule *2.44* on voit que $[\xi^i, \xi^j] \equiv 0$. Ainsi, la condition $[\xi^i, \xi^j] \equiv 0$ s'avère *nécessaire* pour qu'un système de coordonnées cherché existe. Montrons que cette condition est également *suffisante*. Montrons que le système de coordonnées $t_1, \ldots, t_m, h_{m+1}, \ldots, h_n$ construit à partir des champs $\xi^1(x), \ldots, \xi^m(x)$ selon la règle *2.45* possède la propriété voulue. En effet, l'hypothèse du théorème *2.46a* étant remplie, les composantes des vecteurs $\xi^j(x)$ ont, dans les nouvelles coordonnées, la forme

$$\xi^j(x) = \{\lambda_1^j(x), \ldots, \lambda_m^j(x), 0, \ldots, 0\}.$$

Sur l'ensemble $S_{i+1} = \{t_1, \ldots, t_{i+1}, 0, \ldots, h_{m+1}, \ldots, h_n\}$, a lieu l'égalité *2.45* (2):

$$0 = [\xi^i, \xi^{i+1}] = \frac{\partial \xi^i}{\partial t_{i+1}}.$$

Il en découle que le vecteur ξ^i a les composantes constantes sur les lignes vectorielles du champ ξ^{i+1}. Or, pour $t_{i+1} = 0$, i.e. sur S_i, le champ ξ^i a les composantes $\{0, \ldots, 1, \ldots, 0\}$; on a vu qu'il a les mêmes composantes sur S_{i+1}. En appliquant d'une façon analogue l'égalité $[\xi^i, \xi^{i+2}] = 0$ nous voyons que le champ ξ^i a les mêmes composantes sur S_{i+2}; en continuant on établit que ce champ a les mêmes composantes sur S_m, donc dans un voisinage du point a. Ainsi le champ ξ^i se trouve directeur pour la i-ème ligne de coordonnées du nouveau système de coordonnées, ce qu'il fallait démontrer.

Si $m = 1$, c'est-à-dire s'il n'y a qu'un seul champ vectoriel $\xi^1 = \xi(x)$, l'hypothèse du théorème est automatiquement remplie; ainsi, pour un seul champ vectoriel $\xi(x)$ il existe toujours un système de coordonnées pour lequel ce champ $\xi(x)$ s'avère directeur d'une ligne de coordonnées.

Soulignons encore une fois que l'existence d'un système cherché de coordonnées n'est garantie que dans un voisinage du point donné $a \in G$ où les raisonnements *2.45* sont valables.

§ 2. 5. Théorème de Frobenius

2.51. Position du problème. Nous allons étudier une équation différentielle de la forme

$$y'(x) = \Phi(x, y(x)). \tag{1}$$

La fonction inconnue $y(x)$ doit être définie dans un voisinage au moins d'un point donné a d'un espace normé X et prendre ses valeurs dans un espace normé Y. Pour que ce problème ait un sens, il faut supposer que la fonction $\Phi(x, y)$ soit définie dans le produit de deux domaines quelconques $U \ni a$ et V des espaces X et Y respectivement et qu'elle prenne ses valeurs dans le même espace que $y'(x)$, i.e. dans $L(X, Y)$.

L'équation (1) est complétée par la condition initiale

$$y(a) = b \in V. \tag{2}$$

Si $X = R_1$, alors (1) représente une équation différentielle ordinaire. Dans ce cas, comme nous l'avons vu dans *1.61-1.62*, il existe un voisinage du point a où la solution cherchée $y(x)$ existe et est unique, pourvu que soient remplies certaines conditions pour la fonction $\Phi(x, y)$ (continuité, condition de Lipschitz par rapport à y, dérivabilité). Dans le cas général où $X \neq R_1$, les conditions de dérivabilité, aussi dures soient-elles, ne sont pas suffisantes pour la

résolubilité de l'équation (1) : la fonction $\Phi(x, y)$ doit satisfaire à certaines équations spéciales.

Envisageons le cas où $X = R_2$, $Y = R_1$. L'équation (1) est alors équivalente au système de deux équations à dérivées partielles

$$\left.\begin{aligned} \frac{\partial y(x_1, x_2)}{\partial x_1} &= \Phi_1(x_1, x_2, y), \\ \frac{\partial y(x_1, x_2)}{\partial x_2} &= \Phi_2(x_1, x_2, y). \end{aligned}\right\} \tag{3}$$

Supposons que les fonctions Φ_1 et Φ_2 soient dérivables dans le domaine $U \times V$; alors, une fois la solution existe, elle est déjà une fonction deux fois dérivable, donc vérifie la condition de symétrie de la dérivée seconde (*2.32c*) :

$$\frac{\partial^2 y}{\partial x_1 \partial x_2} = \frac{\partial^2 y}{\partial x_2 \partial x_1}.$$

Avec les équations (3) nous avons

$$\frac{\partial}{\partial x_2} \Phi_1(x_1, x_2, y(x_1, x_2)) = \frac{\partial}{\partial x_1} \Phi_2(x_1, x_2, y(x_1, x_2)),$$

ce qui, complété toujours de (3), amène à la relation

$$\frac{\partial \Phi_1(x_1, x_2, y)}{\partial x_2} + \frac{\partial \Phi_1(x_1, x_2, y)}{\partial y} \Phi_2(x_1, x_2, y) =$$

$$= \frac{\partial \Phi_2(x_1, x_2, y)}{\partial x_1} + \frac{\partial \Phi_2(x_1, x_2, y)}{\partial y} \Phi_1(x_1, x_2, y)$$

qui, par conséquent, est remplie dès qu'une solution du système (3) existe. Pour une solution donnée $y(x_1, x_2)$, elle est satisfaite identiquement par rapport à toutes les valeurs de x_1, x_2 et $y(x_1, x_2)$ et si, *quels que soient trois nombres* a_1, a_2, b où $\{a_1, a_2\} \in U$, $b \in V$, il existe une solution vérifiant la condition $y(a_1, a_2) = b$, alors les relations (3) sont satisfaites identiquement par rapport à toutes les valeurs $x_1 = a_1$, $x_2 = a_2$, $y = b$.

Cet exemple nous suggère les conditions nécessaires de la résolubilité de l'équation (1) dans le cas général. S'il existe une solution $y(x)$ de l'équation (1) et que la fonction $\Phi(x, y)$ est dérivable dans le domaine $U \times V$, alors la fonction $y(x)$ est deux fois dérivable, donc la condition de symétrie de la dérivée seconde *2.32a* est remplie : pour n'importe quels h et k de X, on a

$$y''(x)\, hk = y''(x)\, kh.$$

D'après l'équation (1), en développant la dérivée totale et utilisant encore une fois (1), nous avons la relation

$$\left(\frac{\partial \Phi}{\partial x} + \frac{\partial \Phi}{\partial y} \Phi\right) hk = \left(\frac{\partial \Phi}{\partial x} + \frac{\partial \Phi}{\partial y} \Phi\right) kh \tag{4}$$

qui *a lieu pour tous les h et k dès qu'il existe une solution de l'équation* (1); qui plus est, la relation (4) a lieu *pour tous les* $x = a$, $y = b$, du moment que l'équation (1) possède une solution $y(x)$ satisfaisant à n'importe quelle condition donnée $y(x) = b$.

2.52. Supposons d'abord qu'une solution de l'équation (1) avec la condition (2) existe dans un voisinage $W = \{x: |x - a| < r\}$ du point a et en établissons certaines propriétés.

Soit $h \in X$ un vecteur quelconque, $|h| < r$; considérons la solution $y(x)$ sur le segment $x(t) = a + th$, $0 \leqslant t \leqslant 1$. Là, on peut traiter $y(x)$ comme fonction du paramètre numérique $t \in [0, 1]$. Désignons $\varphi(t) = y[x(t)]$. Alors $\varphi'(t) = y'[x(t)]\, x'(t) = y'[x(t)]\, h$, et l'équation (1) prend la forme

$$\varphi'(t) = \Phi(a + th, \varphi(t))\, h. \tag{5}$$

C'est une *équation différentielle ordinaire par rapport à la fonction* $\varphi(t)$ *de la variable* $t \in [0, 1]$. Pour $t = 0$ nous avons $\varphi(0) = y(a) = b$, de sorte que l'équation (5) est complétée par la condition initiale

$$\varphi(0) = b. \tag{6}$$

Si la fonction $\Phi(x, y)$ est dérivable dans le domaine $U \times V$, alors la fonction $\Phi(a + th, \varphi(t))\, h$ qui figure dans l'équation (5) est elle aussi dérivable (en t), donc l'équation (5) satisfait aux conditions d'existence et d'unicité de solution d'une équation différentielle donnée avec une condition initiale donnée. Il en découle la conclusion que *toute solution de l'équation* (5) *avec la condition* (6) *ne peut qu'être unique*. Par conséquent, *les valeurs d'une solution de l'équation* (1) *avec la condition* (2) *sont définies de façon unique sur tout segment* $x = a + th$, donc *elles sont définies de façon unique dans le voisinage du point a*. Même plus, le résultat que nous disposons permet de trouver cette solution, dès qu'on sache qu'elle existe: il ne faut pour cela que d'intégrer l'équation (5) avec la condition (6) sur tout segment $x = a + th$ et de poser $y[x(t)] = \varphi(t)$.

2.53. Supposons à présent que l'on ne sache rien sur la résolubilité de l'équation (1) et allons construire sa solution. Nous en connaissons la méthode: il faut intégrer l'équation (5) avec la condition (6) sur chaque segment $x = a + th$, $0 \leqslant t \leqslant 1$. Cette équation contient le paramètre h. Le théorème sur la résolution d'une équation ordinaire avec un paramètre (*1.61*) nous permet d'affirmer qu'il existe deux nombres $\delta > 0$, $\rho > 0$ et une fonction $\varphi(t, h)$ dérivable en t qui, pour tous les t, h, $0 \leqslant t \leqslant \delta$, $|h| \leqslant \rho$, représente une solution du problème (5)-(6).

En nous bornant pour le moment à un seul segment $x = a + th$, t_0 étant donné, posons $x_0 = a + t_0 h$ et déterminons la quantité $y(x_0) = \varphi(t_0, h)$. Nous avons à montrer que la valeur $y(x_0)$ dépend non pas de t_0 et h considérés séparément, mais de x_0 tout entier. Soit

k un vecteur colinéaire avec h et $t_0 h = \tau_0 k$. Les fonctions $\varphi(t, h)$ et $\varphi(\tau, k)$ sont des solutions des équations (5) et

$$\varphi'(\tau, k) = \Phi(a + \tau k, \varphi(\tau, k))\, k \tag{7}$$

respectivement, pour la même condition initiale (6). Faisons dans (7) le changement de la variable $\tau = t\frac{\tau_0}{t_0}$ et notons $\varphi\left(t\frac{\tau_0}{t_0}, k\right) = \psi(\ )$. Alors $\psi'(t) = \varphi'\left(t\frac{\tau_0}{t_0}, k\right)\frac{\tau_0}{t_0}$, de sorte que la fonction $\psi(t)$ satisfait à l'équation

$$\psi'(t) = \Phi\left(a + t\frac{\tau_0}{t_0}k, \psi(t)\right) k\frac{\tau_0}{t_0} = \Phi(a + th, \psi(t))\, h$$

qui coïncide avec (5). En vertu du théorème d'unicité, $\psi(t) = \varphi(t, h)$ d'où $\varphi(\tau_0, k) = \psi(t_0) = \psi(t_0, h) = y(x_0)$, ce qu'il nous fallait.

Finalement, la fonction $y(x)$ est définie de façon univoque sur tout segment $x = a + th$ pour $0 \leqslant t \leqslant \delta$ et $|h| \leqslant \rho$. Par conséquent, elle l'est dans un voisinage du point a (notamment, dans la boule $|x - a| \leqslant \delta\rho$). On a démontré qu'au point a même la fonction $y(x)$ possède la dérivée suivant toute demi-droite issue de a et que la valeur de cette dérivée sur tout vecteur $h \in X$ coïncide avec $\Phi(a, b)\, h$.

2.54. Montrons à présent que, si $\Phi(x, y)$ satisfait à la condition (4) et possède les dérivées secondes, alors la fonction $y(x)$, dans un voisinage du point a, est aussi bien dérivable suivant les directions autres que celle du vecteur allant du point a au point x.

Soit $\gamma = \{x \in X;\ x = a + th + sk\}$ un plan à deux dimensions engendré par deux vecteurs non colinéaires h et k. Construisons d'abord sur ce plan une fonction $\varphi(t, s)$ vérifiant l'équation (1) et la condition initiale (2) et montrons ensuite qu'elle coïncide (sur γ) avec la fonction $y(x)$ définie ci-dessus.

A cette fin, remarquons que, dès que la solution cherchée $y(x)$ de l'équation (1) avec la condition (2) existe, la fonction $\varphi(t, s) = y(a + th + sk)$ satisfait au système de deux équations à dérivées partielles:

$$\frac{\partial\varphi}{\partial t} = y'(x)\, x'(t) = \Phi(x, y)\, h = \Phi(a + th + sk, \varphi)\, h,$$

$$\frac{\partial\varphi}{\partial s} = y'(x)\, x'(s) = \Phi(x, y)\, k = \Phi(a + th + sk, \varphi)\, k$$

avec la condition initiale

$$\varphi(0, 0) = y(a) = b.$$

La notation abrégée du système est :

$$\left.\begin{aligned} \frac{\partial \varphi}{\partial t} &= \Phi_1(t, s, \varphi), \quad \Phi_1 \equiv \Phi(a+th+sk, \varphi)\, h, \\ \frac{\partial \varphi}{\partial s} &= \Phi_2(t, s, \varphi), \quad \Phi_2 \equiv \Phi(a+th+sk, \varphi)\, k. \end{aligned}\right\} \tag{8}$$

(Les vecteurs h et k y figurent en tant que paramètres.)

Rejetons la supposition d'existence de la solution $y(x)$ et considérons le système (8) avec la condition

$$\varphi(0, 0) = b. \tag{9}$$

Il est facile de prouver que les fonctions $\Phi_1(t, s, \varphi)$ et $\Phi_2(t, s, \varphi)$ satisfont à une relation découlant de la condition de symétrie (4). En effet, il est évident que

$$\frac{\partial \Phi_1}{\partial s} = \frac{\partial \Phi}{\partial x} kh, \quad \frac{\partial \Phi_2}{\partial t} = \frac{\partial \Phi}{\partial x} hk.$$

Ensuite, calculons $\frac{\partial \Phi_1}{\partial \varphi}\Phi_2$. C'est la partie linéaire principale de l'accroissement de la fonction $\Phi_1(t, s, \varphi)$, lorsque φ subit un accroissement Φ_2, i.e. la partie linéaire principale (par rapport à Φ_2) de la différence

$$\Phi_1(t, s, \varphi + \Phi_2) - \Phi_1(t, s, \varphi) = \\ = [\Phi(a + th + sk, \varphi + \Phi_2) - \Phi(a + th + sk, \varphi)]\, h.$$

Comme la fonction $\Phi(x, y)$ est dérivable, cette partie linéaire principale coïncide avec $\frac{\partial \Phi}{\partial y}\Phi_2 h$, donc avec $\frac{\partial \Phi}{\partial y}\Phi kh$. D'une manière analogue, $\frac{\partial \Phi_2}{\partial \varphi}\Phi_1 = \frac{\partial \Phi}{\partial y}\Phi hk$.

Ainsi, l'égalité (4) prend la forme

$$\frac{\partial \Phi_2}{\partial t} + \frac{\partial \Phi_2}{\partial \varphi}\Phi_1 = \frac{\partial \Phi_1}{\partial s} + \frac{\partial \Phi_1}{\partial \varphi}\Phi_2; \tag{10}$$

c'est bien ce qu'on a dû attendre, vu l'exemple de *2.51*.

Passons à la résolution du système (8). D'après la deuxième équation (8), la fonction $\psi(s) = \varphi(0, s)$ satisfait à l'équation différentielle ordinaire

$$\frac{\partial \psi}{\partial s} = \Phi_2(0, s, \psi) \tag{11}$$

et à la condition initiale $\psi(0) = \varphi(0, 0) = b$. (C'est la même équation qu'on a considérée dans *2.52*, bien que ce fait ne soit pas important pour le moment.) En vertu du théorème *1.61*, il existe $\delta_1 > 0$, $\delta_1 \leqslant \delta$ et $\rho_1 > 0$, $\rho_1 \leqslant \rho$ tels que, pour toutes les valeurs $|h| < \rho_1$, $|k| < \rho_1$ sur le segment $0 \leqslant s \leqslant \delta_1$, la solution $\psi(s)$ existe. Maintenant considérons, sur une droite $s = s_0 \in [0, \delta_1]$ dans

le plan $\{t, s\}$, l'équation différentielle ordinaire

$$\frac{\partial \varphi(t, s_0)}{\partial t} = \Phi_1(t, s_0, \varphi(t, s_0)) \tag{12}$$

avec la condition initiale

$$\varphi(0, s_0) = \psi(s_0). \tag{13}$$

Toujours en vertu du théorème *1.61*, il existe $\rho_2 > 0$, $\rho_2 \leqslant \rho_1$ et $\delta_2 > 0$, $\delta_2 \leqslant \delta_1$, tels que, pour $|h| \leqslant \rho_2$, $|k| < \rho_2$, $s_0 \in [0, \delta_2]$ le problème (12)-(13) est résoluble sur le segment $t \in [0, \delta_2]$. D'après *2.37b*, la solution $\varphi(t, s)$ est deux fois dérivable par rapport au paramètre s, d'après *2.37c*, $\frac{\partial \varphi(t, s)}{\partial t}$ est aussi dérivable en s; à son tour, $\frac{\partial \varphi(t, s)}{\partial s}$ est dérivable en t (*1.66*), et toutes les dérivées mentionnées sont continues. Montrons que $\varphi(t, s)$ satisfait à l'équation

$$\frac{\partial \varphi(t, s)}{\partial s} = \Phi_2(t, s, \varphi). \tag{14}$$

Pour le faire, composons la fonction

$$\Psi(t, s) = \frac{\partial \varphi(t, s)}{\partial s} - \Phi_2(t, s, \varphi(t, s)).$$

On a par construction

$$\Psi(0, s) = \frac{\partial \varphi(0, s)}{\partial s} - \Phi_2(0, s, \varphi(0, s)) = \frac{\partial \psi}{\partial s} - \Phi_2(0, s, \psi) = 0. \tag{15}$$

Ensuite,

$$\frac{\partial \psi}{\partial t} = \frac{\partial^2 \varphi(t, s)}{\partial t\, \partial s} - \frac{\partial}{\partial t} \Phi_2(t, s, \varphi(t, s)).$$

En utilisant la symétrie de la dérivée mixte (*2.32c*) et développant la dérivée totale dans le deuxième terme nous obtenons

$$\frac{\partial \Psi}{\partial t} = \frac{\partial^2 \varphi(t, s)}{\partial s\, \partial t} - \frac{\partial \Phi_2}{\partial t} - \frac{\partial \Phi_2}{\partial \varphi} \frac{\partial \varphi}{\partial t} = \frac{\partial}{\partial s} \Phi_1(t, s, \varphi(t, s)) - \frac{\partial \Phi_2}{\partial t} - \frac{\partial \Phi_2}{\partial \varphi} \Phi_1.$$

Il en résulte d'après (10)

$$\begin{aligned} \frac{\partial \Psi}{\partial t} &= \frac{\partial \Phi_1}{\partial s} + \frac{\partial \Phi_1}{\partial \varphi} \frac{\partial \varphi}{\partial s} - \frac{\partial \Phi_2}{\partial t} - \frac{\partial \Phi_2}{\partial \varphi} \Phi_1 = \\ &= \frac{\partial \Phi_1}{\partial s} + \frac{\partial \Phi_1}{\partial \varphi} \frac{\partial \varphi}{\partial s} - \frac{\partial \Phi_1}{\partial s} - \frac{\partial \Phi_1}{\partial \varphi} \Phi_2 = \\ &= \frac{\partial \Phi_1}{\partial \varphi} \left(\frac{\partial \varphi}{\partial s} - \Phi_2 \right) = \frac{\partial \Phi_1}{\partial \varphi} \Psi. \end{aligned} \tag{16}$$

Nous voyons que la fonction $\Psi(t, s)$ satisfait à l'équation différentielle ordinaire (16) avec la condition initiale (15); donc, vu le théorème d'unicité, $\Psi(t, s) \equiv 0$, et l'égalité (14) est démontrée.

Ainsi, la fonction $\varphi(t, s)$ vérifie les deux équations du système (8) dans le carré $t \in [0, \delta_2]$, $s \in [0, \delta_2]$. Par conséquent, la fonction $z(x) = \varphi(t, s)$ $(x = a + th + sk)$ représente une solution de l'équation (1) avec la condition (2) dans le plan γ, donc, y coïncide avec la fonction $y(x)$ construite dans *2.53*. Il en découle que la fonction $y(x)$, de même que $z(x)$, est dérivable pour $x = a + th$ suivant la direction du vecteur k et que

$$y'(x)\,k = \frac{\partial \varphi(t, 0)}{\partial s} = \Phi_2(t, 0, \varphi) = \Phi(x, y)\,k. \tag{17}$$

2.55. Nous sommes maintenant en mesure de donner l'énoncé du théorème fondamental du présent paragraphe.

Théorème (Frobenius). *Soit $\Phi(x, y)$ une fonction définie dans le produit de deux domaines $U \subset X$ et $V \subset Y$ et à valeurs dans l'espace $L(U \times Y)$; supposons ensuite que la fonction $\Phi(x, y)$ soit deux fois dérivable dans le domaine $U \times V$ et qu'elle y satisfasse à la relation*

$$\left(\frac{\partial \Phi}{\partial x} + \frac{\partial \Phi}{\partial y}\,\Phi\right) hk = \left(\frac{\partial \Phi}{\partial x} + \frac{\partial \Phi}{\partial y}\,\Phi\right) kh$$

pour n'importe quels vecteurs k et h des espaces X et Y respectivement. Alors il existe un nombre $\delta > 0$ et une fonction $y = y(x): X \to Y$ tels que cette dernière satisfait, dans la boule $|x - a| < \delta$, à l'équation différentielle

$$y'(x) = \Phi(x, y(x))$$

avec la condition initiale $(a \in U, b \in V)$

$$y(a) = b.$$

Dans la boule indiquée, la solution $y(x)$ est unique.

Démonstration. Nous avons vu dans *2.53* que, dans un voisinage du point a, il existe une fonction $y(x)$, $y(a) = b$ qui, d'après *2.54*, possède à tout point x une dérivée suivant chaque direction, cette dérivée vérifiant la condition (17). Par hypothèse, la fonction $\Phi(x, y)$ représente, pour tout x et $y = y(x)$, un opérateur linéaire de l'espace X dans l'espace Y et continu en x. Ceci étant, la fonction $y(x)$ est dérivable d'après *1.48d* et sa dérivée coïncide avec l'opérateur $\Phi(x, y)$. Le théorème est démontré.

2.56. Existence d'une primitive. Supposons que, dans l'hypothèse du théorème *2.55*, la fonction $\Phi(x, y)$ ne dépende pas de y, de sorte que nous avons affaire à l'équation

$$y'(x) = \Phi(x), \quad y(a) = b. \tag{1}$$

C'est le problème de la recherche d'une primitive de la fonction $\Phi(x)$ $(X \to L(X, Y))$, c'est-à-dire d'une fonction $y(x)$ dont la dérivée coïncide avec $\Phi(x)$. Un tel problème n'est pas résoluble pour une fonction $\Phi(x)$ quelconque. Une condition nécessaire de

l'existence d'une primitive est la symétrie de la dérivée seconde de $y(x)$, ce qui amène à la relation

$$\Phi'(x)\,hk = \Phi'(x)\,kh, \qquad h \in X, \qquad k \in X. \tag{2}$$

Il résulte du théorème *2.55* que *l'égalité* (2) *est en même temps une condition suffisante de l'existence d'une primitive dans un voisinage du point a.*

§ 2.6. Systèmes d'équations à dérivées partielles et applications géométriques

2.61.a. Considérons un système d'équations différentielles de la forme:

$$\left.\begin{aligned} \frac{\partial y(x_1, \ldots, x_n)}{\partial x_1} &= \Phi_1(x_1, \ldots, x_n, y), \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ \frac{\partial y(x_1, \ldots, x_n)}{\partial x_n} &= \Phi_n(x_1, \ldots, x_n, y) \end{aligned}\right\} \tag{1}$$

avec une fonction inconnue $y(x) \equiv y(x_1, \ldots, x_n)$ qu'on cherche au voisinage d'un point donné $a = (a_1, \ldots, a_n) \in R_n$, cette fonction prenant ses valeurs dans un espace Y et satisfaisant à la condition initiale

$$y(a) = b \in Y. \tag{2}$$

En ce qui concerne les fonctions $\Phi_i(x_1, \ldots, x_n, y)$, nous supposons qu'elles soient définies dans le produit de domaines $U_1 \times \ldots \times U_n \times V$, $U_i \ni a_i$, $V \ni b$ $(i = 1, \ldots, n)$ et représentent, pour n'importe quels $x_i \in U_i$ et $y \in V$, des opérateurs linéaires $X_i \to L(X_i, Y)$. Dans ce cas on peut mettre le système (1) sous forme d'une seule équation

$$y'(x) = \Phi(x, y): X \to L(X, Y), \tag{3}$$

où $U = U_1 \times \ldots \times U_n \in R_n$, $\Phi(x, y)\,h = \sum_{i=1}^{n} \Phi_i(x, y)\,h_i$, $h = \{h_1, \ldots, h_n\} \in R_n$, après quoi toute la situation s'avère un cas particulier de la situation décrite par le théorème de Frobenius *2.55*. La condition nécessaire de la résolubilité *2.51* (4) peut être mise, comme on l'a déjà fait pour les systèmes *2.54* (8) et *2.51* (3), sous la forme:

$$\frac{\partial \Phi_i(x, y)}{\partial x_j} + \frac{\partial \Phi_i(x, y)}{\partial y}\Phi_j(x, y) = \frac{\partial \Phi_j(x, y)}{\partial x_i} + \frac{\partial \Phi_j(x, y)}{\partial y}\Phi_i(x, y) \quad (i, j = 1, \ldots, n). \tag{4}$$

Ceci fait, le théorème de Frobenius appliqué à notre cas affirme: *si les fonctions* $\Phi_i(x_1, \ldots, x_n, y)$ *sont deux fois dérivables et satisfont dans le domaine* $U \times V$ *aux conditions* (4), *alors, pour tout point* $a \in U$ *et pour tout* $b \in V$, *il existe* $\delta > 0$ *tel qu'une solution* $y = y(x)$ *du système* (1) *avec la condition initiale* (2) *existe et est unique dans la boule* $|x - a| < \delta$.

b. Pour un système plus simple:

$$\left.\begin{aligned} \frac{\partial y(x_1, \ldots, x_n)}{\partial x_1} &= \Phi_1(x_1, \ldots, x_n), \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \frac{\partial y(x_1, \ldots, x_n)}{\partial x_n} &= \Phi_n(x_1, \ldots, x_n) \end{aligned}\right\} \tag{5}$$

le résultat obtenu ci-dessus reste, bien entendu, valable; notons seulement que la condition (4) prend ici une forme très simple:

$$\frac{\partial \Phi_i(x, y)}{\partial x_j} = \frac{\partial \Phi_j(x, y)}{\partial x_i} \quad (i, j = 1, \ldots, n). \tag{6}$$

c. Le schéma général enferme également les systèmes à plusieurs fonctions inconnues $z_k(x): G \subset R_n \to Y_k$, $k = 1, \ldots, m$:

$$\frac{\partial z_k}{\partial x_i} = f_{ki}(x_1, \ldots, x_n, z_1, \ldots, z_m) \tag{7}$$

$$i = 1, \ldots, n; \quad k = 1, \ldots, m$$

et avec les conditions initiales arbitraires

$$z_k(x_1^0, \ldots, x_n^0) = z_k^0 \in G \quad (k = 1, \ldots, m). \tag{8}$$

Il résulte de a (où l'on doit poser Y égal à la somme droite $Y_1 + \ldots + Y_m$) que la condition nécessaire et suffisante de la résolubilité du problème (7)-(8) a la forme

$$\frac{\partial f_{ki}}{\partial x_p} + \sum_{j=1}^{m} \frac{\partial f_{ki}}{\partial z_j} f_{jp} \equiv \frac{\partial f_{kp}}{\partial x_i} + \sum_{j=1}^{m} \frac{\partial f_{kp}}{\partial z_j} f_{ji}$$

$$(k = 1, \ldots, m; \quad i, p = 1, \ldots, n)$$

quelles que soient les données initiales dans le domaine de définition des fonctions $f_{ki}(x, z)$.

2.62. Interprétation géométrique. Pour une équation différentielle ordinaire

$$y' = \Phi(x, y) \quad (x \in U \subset R_1, \; y \in V \subset R_1), \tag{1}$$

on a une interprétation géométrique connue: à tout point $\{x_0, y_0\}$ du domaine $U \times V \subset R_2$, on fait correspondre une droite

$$y - y_0 = \Phi(x_0, y_0)(x - x_0), \tag{2}$$

et l'on cherche une courbe passant par le point donné $\{a, b\} \in U \times V$ et ayant pour tangente à tout son point $\{x_0, y_0\}$ la droite (2). Le théorème d'existence et d'unicité de la solution garantit la résolubilité de ce problème et l'unicité de sa solution dès que la fonction $\Phi(x, y)$ soit assez lisse.

Une interprétation géométrique analogue a lieu pour l'équation générale

$$y' = \Phi(x, y) \quad (x \in U \subset X,\ y \in V \subset Y).$$

Ici, à tout point $\{x_0, y_0\} \in U \times V$ on fait correspondre une variété linéaire dans l'espace $X \times Y$:

$$y - y_0 = \Phi(x_0, y_0)(x - x_0), \tag{3}$$

et l'on cherche une fonction $y = y(x)$, $y(a) = b$ dont le graphique dans l'espace $X \times Y$ (« variété intégrale ») ait, à tout son point $\{x_0, y_0\}$, pour plan tangent (*1.26c*) la variété linéaire correspondante (3). Dans ce cas, on ne peut plus dire que toute fonction $\Phi(x, y)$ (aussi lisse qu'elle soit) amène à un problème résoluble; nous savons que c'est la condition *2.51* (4) (nécessaire et suffisante) qui est à estimer.

Le plus simple exemple non trivial peut déjà être construit pour $X = R_2$, $Y = R_1$. Dans un domaine $U \times V$ ($U \subset R_2$, $V \subset R_1$), on fait passer par tout point $\{x_1^0, x_2^0, y_0\}$ un plan

$$y - y_0 = \Phi_1(x_1^0, x_2^0, y_0)(x_1 - x_1^0) + \Phi_2(x_1^0, x_2^0, y_0)(x_2 - x_2^0), \tag{4}$$

et l'on cherche une surface $y = y(x_1, x_2)$ qui passe par le point donné $\{a_1, a_2, b\} \in U \times V$ et qui ait pour plan tangent à tout son point $\{x_1^0, x_2^0, y_0\}$ le plan correspondant (4). Ce problème possède une solution non pas pour tout couple de fonctions $\Phi_1(x_1, x_2, y)$ et $\Phi_2(x_1, x_2, y)$, mais seulement pour les couples qui vérifient la condition qui nous est déjà familière

$$\frac{\partial \Phi_1}{\partial x_2} + \frac{\partial \Phi_1}{\partial y}\Phi_2 = \frac{\partial \Phi_2}{\partial x_1} + \frac{\partial \Phi_2}{\partial y}\Phi_1$$

(à condition d'existence des dérivées secondes).

Dans *2.63-2.64* on indique encore quelques problèmes géométriques liés au théorème de Frobenius.

2.63. a. Supposons que pour tout point x d'un domaine $V \subset R_n$ soit donné un plan $(n-1)$-dimensionnel $\gamma(x) \ni x$ qui dépend continûment de x; il faut construire une fonction $w = f(x): V \to R_1$ dont chaque surface de niveau $f(x) = C$ soit tangente, à tout son point

$x \in V$, au plan $\gamma(x)$. Puisque le problème, comme d'habitude, se pose localement, nous supposons qu'aucun des plans $\gamma(x)$, dans le domaine V, ne contienne de droites parallèles à l'axe des x_n (sinon on permute les axes des coordonnées). C'est la raison pour laquelle on modifie les notations: désignons par y la coordonnée x_n, par x l'ensemble des autres coordonnées $\{x_1, \ldots, x_{n-1}\}$. Chaque point du domaine V sera donc désigné par $\{x, y\}$. Au lieu de $\gamma(x)$ nous écrirons naturellement $\gamma(x, y)$. L'équation du plan $\gamma(x^0, y^0)$ aura la forme

$$y - y_0 = \sum_{i=1}^{n-1} \Phi_i(x^0, y^0)(x_i - x_i^0), \tag{1}$$

les fonctions $\Phi_i(x, y)$ étant connues par hypothèse. Nous les supposons deux fois dérivables. La solution cherchée est notée $w = f(x, y)$. Supposons qu'une solution $w = f(x, y)$ existe. Fixons un point $a \in R_{n-1}$ appartenant à la projection sur R_{n-1} du domaine $V \subset R_n$ et considérons $f(a, y)$. Il est évident que $\frac{\partial f(a, y)}{\partial y} \not\equiv 0$, sinon la droite $x = a$ appartiendrait à une surface de niveau de la fonction $f(x, y)$, ce qui n'est pas par hypothèse. Remarquons ensuite que, avec $f(x, y)$, toute fonction $F(f(x, y))$ représente une solution de notre problème; nous pouvons nous en servir afin d'obtenir une solution $f(x, y)$ qui (sur une partie Δ au moins de la droite $x = a$ à laquelle nous nous bornerons dans ce qui suit) satisfasse à la condition $f(a, y) \equiv y$. Par conséquent, l'équation de la surface de niveau de la fonction $f(x, y)$ qui passe par le point $\{a, b\} \leqslant \Delta$ a la forme $f(x, y) = b$. D'après le théorème sur la fonction implicite, au voisinage du point $\{a, b\}$, cette équation est équivalente à une équation $y = y(x, b)$. Le plan tangent à cette surface au point $\{x^0, y^0\}$ est

$$y - y_0 = \sum_{i=1}^{n-1} \frac{\partial y(x^0, b)}{\partial x_i}(x_i - x_i^0). \tag{2}$$

Comme les plans (2) et (1) coïncident, la fonction $y(x, b)$ vérifie le système d'équations

$$\frac{\partial y(x, b)}{\partial x_i} = \Phi_i(x, y), \quad i = 1, \ldots, n-1 \tag{3}$$

avec les conditions initiales

$$y(a, b) = b. \tag{4}$$

D'après le théorème de Frobenius, pour qu'une solution du problème (3)-(4) existe, il faut et il suffit que soient remplies les conditions d'intégrabilité $(i, j = 1, \ldots, n-1)$

$$\frac{\partial \Phi_i(x, y)}{\partial x_j} + \frac{\partial \Phi_i(x, y)}{\partial y}\Phi_j(x, y) = \frac{\partial \Phi_j(x, y)}{\partial x_i} + \frac{\partial \Phi_j(x, y)}{\partial y}\Phi_i(x, y). \tag{5}$$

La condition (5) est donc *nécessaire* pour la résolubilité du problème posé. Montrons qu'elle est aussi *suffisante*, au moins dans un voisinage du point $\{a, b\}$. La condition (5) étant remplie, le théorème de Frobenius implique l'existence, pour tout b et pour un voisinage du point a, d'une solution $y(x, b)$ du système (3) avec la condition (4). Il est clair que $\frac{\partial y(a, b)}{\partial b} = 1$; par conséquent, on peut résoudre l'équation $y = y(x, b)$ par rapport à b, d'après le théorème sur la fonction implicite, de sorte qu'elle se trouve équivalente à une équation de la forme $b = f(x, y)$. Nous affirmons que la fonction $f(x, y)$ est la fonction cherchée. En effet, le plan tangent à la surface $f(x, y) = b$ est le même que pour la surface $y = y(x, b)$, c'est donc le plan (2) ; or, en raison des équations (3), le plan (2) coïncide avec le plan (1), ce qu'il nous fallait.

b. Si $n - 1 = 1$, i.e. $n = 2$, alors le système (3) ne comprend qu'une seule équation et est toujours compatible. Ainsi, si pour tout point $\{x, y\}$ d'un domaine V du plan on donne une droite $\gamma(x, y)$, il est toujours possible de choisir une fonction $f(x, y)$ dont toute ligne de niveau $f(x, y) = C$ (au moins dans un voisinage du point donné) soit tangente à la droite correspondante $\gamma(x, y)$. (Naturellement, on doit exiger que le coefficient angulaire de la droite $\gamma(x, y)$ soit assez lisse.) Si $n > 2$, alors, comme nous l'avons vu, le problème analogue n'est pas toujours résoluble.

2.64. Supposons que, pour tout point x d'un domaine $V \subset R_n$, il y ait n vecteurs linéairement indépendants $g_1(x), \ldots, g_n(x)$ (dont la dépendance de x soit suffisamment lisse). On se demande, s'il est possible d'introduire, dans le domaine V, un nouveau système des coordonnées $w_1, \ldots, w_n$ tel que chaque ligne de coordonnées, i.e. une ligne sur laquelle toutes les coordonnées à l'exception d'une seule, mettons w_j, restent constantes, soit tangente à tout son point au vecteur correspondant $g_j(x)$.

Si un tel système des fonctions $w_k(x_1, \ldots, x_n)$ $(k = 1, \ldots, n)$ existe, alors chaque surface de niveau $w_k = C$ est tangente à tout son point aux vecteurs correspondants $g_1, \ldots, g_{k-1}, g_{k+1}, \ldots, g_n$ et possède donc un plan tangent donné, notamment celui qui est engendré par les vecteurs $g_1, \ldots, g_{k-1}, g_{k+1}, \ldots, g_n$.

Inversement, si, pour $k = 1, \ldots, n$, on réussit à trouver une fonction suffisamment lisse $w_k(x)$ dont chaque surface de niveau soit tangente, à tout son point, au plan engendré par les vecteurs $g_1, \ldots$ $\ldots, g_{k-1}, g_{k+1}, \ldots, g_n$, alors le problème sera résolu, car, dans ce cas, chaque ligne de coordonnées $w_1 = C_1, \ldots, w_{k-1} = C_{k-1}, w_{k+1} = C_{k+1}, \ldots, w_n = C_n$ sera tangente à tout son point au vecteur g_k. Nous voyons qu'en principe notre problème est réduit au précédent. Pour $n = 2$, une solution existe toujours (à condition que les vecteurs donnés soient convenablement lisses) ; pour $n > 2$, une solution existe dès que soient remplies certaines conditions d'intégrabilité

que nous allons obtenir. L'équation du plan engendré par les vecteurs $g_1(x^0), \ldots, g_{k-1}(x^0), g_{k+1}(x^0), \ldots, g_n(x^0)$ a la forme:

$$\begin{vmatrix} x_1 - x_1^0 & g_{11}(x^0) \ldots g_{k-1,1}(x^0) & g_{k+1,1}(x^0) \ldots g_{n1}(x^0) \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n - x_n^0 & g_{1n}(x^0) \ldots g_{k-1,n}(x^0) & g_{k+1,n}(x^0) \ldots g_{nn}(x^0) \end{vmatrix} = 0, \qquad (1)$$

où $g_{ij}(x^0)$ sont les composantes du vecteur $g_i(x^0)$.

Soit $A_{ik}(x^0)$ le complément algébrique d'un élément $g_{ik}(x^0)$ de la $n \times n$-matrice $\| g_{ij}(x^0) \|$. Alors, en développant le déterminant dans (1) suivant la première colonne, nous mettons l'équation (1) sous la forme

$$\sum_{i=1}^{n} (x_i - x_i^0) A_{ik}(x^0) = 0,$$

soit, pour $A_{nk}(x^0) \neq 0$ (ce qu'on peut supposer sans restreindre la généralité, les vecteurs g_i étant linéairement indépendants par hypothèse),

$$x_n - x_n^0 = -\sum_{i=1}^{n-1} \frac{A_{ik}(x^0)}{A_{nk}(x^0)} (x_i - x_i^0).$$

Ainsi les fonctions $\Phi_i(x^0, y^0)$ figurant dans l'équation *2.63* (1) sont définies. Dans le présent cas, elles dépendent de l'indice k ($=1, \ldots, n$). Ceci étant, nous avons à vérifier les conditions *2.63* (5) et, dès qu'elles soient remplies pour tout $k = 1, \ldots, n$ (et dans ce cas seulement), notre problème est résoluble, au moins localement.

2.65. Système qui ne vérifie pas l'hypothèse du théorème de Frobenius. Considérons le système suivant:

$$\frac{\partial z_k}{\partial x_i} = f_{ki}(x_1, \ldots, x_n, z_1, \ldots, z_m) \qquad (i = 1, \ldots, n;\ k = 1, \ldots, m) \qquad (1)$$

dans un domaine $U \times V$, $U \subset R_n$, $V \subset R_m$, sans supposer que l'hypothèse du théorème de Frobenius

$$\frac{\partial f_{ki}}{\partial x_p} + \sum_{j=1}^{m} \frac{\partial f_{ki}}{\partial z_j} f_{jp} = \frac{\partial f_{kp}}{\partial x_i} + \sum_{j=1}^{m} \frac{\partial f_{kp}}{\partial z_j} f_{ji} \qquad (2)$$

$$(k = 1, \ldots, m; \quad i, p = 1, \ldots, n)$$

soit remplie identiquement dans tout le domaine $U \times V$, de sorte que l'ensemble S des points (x, z) où les conditions (2) aient lieu ne soit qu'une variété dans $U \times V$ de dimension inférieure à celle du domaine. Le système (1) ne possède plus de solution vérifiant les données initiales arbitraires

$$x^0 \in U, \quad z_k(x_1^0, \ldots, x_n^0) = z_k^0 \in V \quad (k = 1, \ldots, m), \qquad (3)$$

puisque toute solution doit appartenir totalement à la variété S. Or, en général, ce n'est pas par chaque point de la variété S qu'une solution passe; il se

trouve que, en plus des conditions (2), il y a encore toute une série de conditions à estimer. Pour étudier les possibilités qui se présentent et ayant en vue maintes dérivations successives à effectuer, nous supposons maintenant que les fonctions f_{ki} dans (1) soient indéfiniment dérivables. Notons les égalités (2) d'une façon abrégée

$$E_\alpha^{(1)}(x, z) = 0. \tag{4}$$

Soit $z = z(x)$ une solution du système (1); nous avons vu qu'elle satisfait également au système (4). Dérivons chaque équation du système (4), en y posant préalablement $z = z(x)$, par rapport à chaque variable x_i et remplaçons les dérivées $\frac{\partial z_k}{\partial x_i}$ qui apparaissent par leurs expressions du système (1); nous déduisons de (4) un nouveau système de relations finies auquel la solution $z(x)$ doit satisfaire; désignons ce système comme suit:

$$E_\alpha^{(2)}(x, z) = 0. \tag{5}$$

En appliquant le même raisonnement au système (5) nous en déduisons encore un système d'équations:

$$E_\alpha^{(3)}(x, z) = 0; \ldots \tag{6}$$

ce procédé peut être continué indéfiniment et a pour résultat une suite des systèmes d'équations:

$$E_\alpha^{(s)}(x, z) = 0, \quad s = 1, 2, \ldots; \tag{7}$$

la solution $z(x)$ satisfait nécessairement à chacun de ces systèmes.

Considérons à présent la variété W composée par tous les points $\{x, z\}$ vérifiant tous les systèmes (7). Elle peut s'avérer vide ou bien dégénérée à ce sens qu'il ne contienne aucune « surface » du type $z = z(x)$ définie au moins sur un sous-domaine du domaine U. Dans de tels cas, il n'y a à fortiori aucune solution du système (1). Nous n'avons donc à considérer que le cas où la variété W n'est pas dégénérée, i.e. contient certaines « surfaces ». C'est parmi ces dernières que se trouvent les solutions éventuelles du système (1). Pour discuter l'existence de telles « surfaces », on peut considérer la matrice jacobienne

$$J = \left\| \frac{\partial E_\alpha^{(s)}(x, z)}{\partial z_k} \right\| \quad (s = 1, 2, \ldots; \ \alpha = 1, 2, \ldots; \ k = 1, \ldots, m) \tag{8}$$

à m colonnes et à un nombre infini de lignes. En tout point $(x, z) \in W$ cette matrice a un certain rang qui ne dépasse pas m; trouvons un point $\{a, b\} \in W$ où le rang de la matrice J atteint sa valeur maximale, par exemple r, $r \leqslant m$. En raison de continuité, un mineur d'ordre r qui n'est pas nul au point $\{a, b\}$ reste non nul dans un voisinage de ce point. D'après le théorème du rang *1.74b*, pour tout ensemble fini d'équations du système (7), on peut choisir un voisinage Q du point $\{a, b\}$ dans lequel le lieu géométrique des points satisfaisant à ces équations est représenté par un système d'équations de la forme

$$z_j = \Phi_j(x, z_{r+1}, \ldots, z_m), \tag{9}$$

où $\Phi_j(a, b_{r+1}, \ldots, b_m) = b_j$, $j = 1, \ldots, r$, les fonctions Φ_j dans les seconds membres étant dérivables. Les fonctions Φ_j sont définies dans un voisinage $Q^{(n+m-r)}$ du point $(a, b_{r+1}, \ldots, b_m) \in R_{n+m-r}$ et prennent leurs valeurs dans un voisinage $Q^{(r)}$ du point $(b_1, \ldots, b_r) \in R_r$, de sorte que $Q = Q^{(r)} \times Q^{(n+m-r)}$. Lorsque le nombre d'équations choisies du système (7) croît, le voisinage Q peut diminuer, et, en général, il n'y a pas de garantie qu'il existe un voisinage Q où le système infini (7) soit équivalent au système (9).

Théorème (Veblen et Thomas). *Si le point* $\{a, b\} \in W$ *admet un voisinage* Q *de la forme* $Q^{(r)} \times Q^{(n+m-r)}$ *dans lequel le système infini* (7) *soit équivalent au système fini* (9), *le système* (1) *possède une solution* $z = z(x)$ *avec* $z(a) = b$.

Démonstration. Par hypothèse, si l'on dérive en x_i les équations du système (9) et qu'on y remplace les dérivées $\frac{\partial z_k}{\partial x_i}$ par leurs expressions du système (1), alors on arrive aux relations

$$f_{ji}(x, z) = \frac{\partial \Phi_j(x, z)}{\partial x_i} + \sum_{k=r+1}^{m} \frac{\partial \Phi_j(x, z)}{\partial z_k} f_{ki}(x, z) \tag{10}$$

$$(i = 1, \ldots, n; \quad j = 1, \ldots, r)$$

valables sur toute la variété W (dans les limites du voisinage Q).

Considérons le système suivant d'équations par rapport aux fonctions inconnues $z_{r+1}(x), \ldots, z_m(x)$:

$$\frac{\partial z_k}{\partial x_i} = F_{ki}(x_1, \ldots, x_n, z_{r+1}, \ldots, z_m) \tag{11}$$

$$(i = 1, \ldots, n; \quad k = r+1, \ldots, m),$$

où les fonctions F_{ki} sont obtenues des fonctions f_{ki} figurant dans le système (1) en substituant aux variables $z_1, \ldots, z_r$ leurs expressions (9) par les variables $x, z_{r+1}, \ldots, z_m$. Montrons que le système (11) vérifie l'hypothèse du théorème de Frobenius dans le voisinage $Q^{(n+m-r)}$. En effet, nous avons sur $Q^{(n+m-r)}$:

$$\frac{\partial F_{ki}}{\partial x_s} + \sum_{l=r+1}^{m} \frac{\partial F_{ki}}{\partial z_l} F_{ls} = \frac{\partial f_{ki}}{\partial x_s} + \sum_{p=1}^{r} \frac{\partial f_{ki}}{\partial z_p} \frac{\partial \Phi_p}{\partial x_s} +$$

$$+ \sum_{l=r+1}^{m} \left[\sum_{p=1}^{r} \frac{\partial f_{ki}}{\partial z_p} \frac{\partial \Phi_p}{\partial z_l} + \frac{\partial f_{ki}}{\partial z_l} \right] F_{ls} =$$

$$= \frac{\partial f_{ki}}{\partial x_s} + \sum_{p=1}^{r} \frac{\partial f_{ki}}{\partial z_p} \left[\frac{\partial \Phi_p}{\partial x_s} + \sum_{l=r+1}^{m} \frac{\partial \Phi_p}{\partial z_l} F_{ls} \right] + \sum_{l=r+1}^{m} \frac{\partial f_{ki}}{\partial z_l} F_{ls} =$$

$$= \frac{\partial f_{ki}}{\partial x_s} + \sum_{p=1}^{r} \frac{\partial f_{ki}}{\partial z_p} F_{ps} + \sum_{l=r+1}^{m} \frac{\partial f_{ki}}{\partial z_l} F_{ls} = \frac{\partial f_{ki}}{\partial x_s} + \sum_{p=1}^{m} \frac{\partial f_{ki}}{\partial z_p} F_{ps}, \tag{12}$$

où l'avant-dernière égalité est obtenue avec (10). D'une façon analogue, toujours sur $Q^{(n+m-r)}$, on a

$$\frac{\partial F_{ks}}{\partial x_i} + \sum_{l=r+1}^{m} \frac{\partial F_{ks}}{\partial z_l} F_{li} = \frac{\partial f_{ks}}{\partial x_i} + \sum_{p=1}^{m} \frac{\partial f_{ks}}{\partial z_p} F_{pi}. \tag{13}$$

Les égalités

$$\frac{\partial f_{ki}}{\partial x_s} + \sum_{p=1}^{m} \frac{\partial f_{ki}}{\partial z_p} f_{ps} = \frac{\partial f_{ks}}{\partial x_i} + \sum_{p=1}^{m} \frac{\partial f_{ks}}{\partial z_p} f_{pi}$$

($k = 1, \ldots, m$; $i, s = 1, \ldots, n$) sont valables sur l'ensemble W par construction ; en y portant les expressions (9), nous voyons que les seconds membres de (12) et (13) coïncident sur $Q^{(n+m-r)}$, ce que nous cherchons.

En vertu de *2.61c*, le système (11) admet une solution $\{ z^0_{r+1}(x), \ldots, z^0_m(x) \}$ définie éventuellement dans un voisinage plus petit $Q^{(n)}_1 \subset Q^{(n)}$ et satisfaisant aux conditions

$$z^0_{r+1}(a) = b_{r+1}, \ldots, z^0_n(a) = b_n.$$

Nous affirmons que le système des fonctions $z(x) = \{ z_1(x), \ldots, z_m(x) \}$ de la forme

$$\begin{aligned} z_1(x) &= \Phi_1(x, z^0_{r+1}(x), \ldots, z^0_m(x)), \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ z_r(x) &= \Phi_r(x, z^0_{r+1}(x), \ldots, z^0_m(x)), \\ z_{r+1}(x) &= z^0_{r+1}(x) \\ &\ldots\ldots\ldots \\ z_m(x) &= z^0_m(x) \end{aligned}$$

est une solution du système (1) avec la condition $z(a) = b$. En effet, les fonctions ci-dessus satisfont par construction aux équations du système (1) avec $k \geqslant r+1$. Pour $k \leqslant r$ nous avons d'après (10):

$$\frac{\partial z_k(x)}{\partial x_i} = \frac{\partial \Phi_k}{\partial x_i} + \sum_{p=r+1}^{m} \frac{\partial \Phi_k}{\partial z_p} \frac{\partial z^0_p}{\partial x_i} =$$

$$= \frac{\partial \Phi_k}{\partial x_i} + \sum_{p=r+1}^{m} \frac{\partial \Phi_k}{\partial z_p} F_{pi} = F_{ki} = f_{ki}(x, z^0_1(x), \ldots, z^0_m(x)),$$

et ces équations sont satisfaites elles aussi. Le théorème est démontré.

§ 2. 7. Théorème de Taylor et son inversion

2.71. a. Théorème de Taylor. *Soit* $y = f(x)$ $(V \subset X \to Y)$ *une fonction définie dans une boule* $V = \{x: |x - a| < r\}$ *d'un espace* X, *à valeurs dans un espace* Y, *ayant les dérivées successives y compris la p-ème. Alors, pour* $|h| < r$, *a lieu la formule de Taylor:*

$$f(a+h) = f(a) + f'(a)h + 1/2 f''(a)hh + \ldots$$

$$\ldots + \frac{1}{p!} f^{(p)}(a) \underbrace{h \ldots h}_{p \text{ fois}} + R_p, \qquad (1)$$

où

$$|R_p| \leqslant \frac{|h|^p}{p!} \sup_{x \in V} \| f^{(p)}(x) - f^{(p)}(a) \|. \qquad (2)$$

Démonstration. En fixant h, $|h| < r$, considérons la fonction suivante de la variable réelle t, $0 \leqslant t \leqslant 1$:

$$\varphi(t) = f(a + th).$$

D'après le théorème sur la dérivée d'une fonction composée *2.34*, la fonction $\varphi(t)$ possède, avec $f(x)$, les p premières dérivées ; on a :

$$\varphi'(t) = f'(a+th)\,h, \qquad \varphi'(0) = f'(a)\,h,$$
$$\varphi''(t) = f''(a+th)\,hh, \qquad \varphi''(0) = f''(a)\,hh,$$
$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$
$$\varphi^{(p)}(t) = f^{p}(a+th)\underbrace{h\dots h}_{p \text{ fois}}, \quad \varphi^{(p)}(0) = f^{(p)}(a)\underbrace{h\dots h}_{p \text{ fois}}.$$

Pour la fonction $\varphi(t)$ a lieu la formule de Taylor *U12.64c* :

$$\varphi(1) = \varphi(0) + \varphi'(0) + \frac{1}{2}\varphi''(0) + \dots + \frac{1}{(p-1)!}\varphi^{(p-1)}(0) + Q_p, \quad (3)$$

où Q_p est terme résiduel que l'on peut mettre sous sa forme intégrale

$$Q_p = \int_0^1 \frac{(1-\xi)^{p-1}}{(p-1)!}\varphi^{(p)}(\xi)\,d\xi,$$

soit, après avoir séparé sous le signe somme le terme $\varphi^{(p)}(0)$,

$$Q_p = \int_0^1 \frac{(1-\xi)^{p-1}}{(p-1)!}[\varphi^{(p)}(\xi) - \varphi^{(p)}(0)]\,d\xi + \varphi^{(p)}(0)\int_0^1 \frac{(1-\xi)^{p-1}}{(p-1)!}\,d\xi =$$
$$= \int_0^1 \frac{(1-\xi)^{p-1}}{(p-1)!}[\varphi^{(p)}(\xi) - \varphi^{(p)}(0)]\,d\xi + \frac{1}{p!}\varphi^{(p)}(0).$$

La quantité

$$R_p = \int_0^1 \frac{(1-\xi)^{p-1}}{(p-1)!}[\varphi^{(p)}(\xi) - \varphi^{(p)}(0)]\,d\xi$$

admet, à son tour, l'estimation

$$|R_p| \leqslant \sup_{0\leqslant\xi\leqslant 1} \|\varphi^{(p)}(\xi) - \varphi^{(p)}(0)\| \int_0^1 \frac{(1-\xi)^{p-1}}{(p-1)!}\,d\xi =$$
$$= \sup_{0\leqslant t\leqslant 1} \|[f^{(p)}(a+th) - f^{(p)}(a)]\underbrace{h\dots h}_{p \text{ fois}}\| \cdot \frac{1}{p!} \leqslant$$
$$\leqslant \sup_{x\in V} \|f^{(p)}(x) - f^{(p)}(a)\| \cdot \frac{|h|^p}{p!}.$$

Comme $\varphi^{(p)}(0) = f^{(p)}(a)\,h\dots h$, la relation (1) est démontrée.

b. Conséquence. *Si la fonction $f^{(p)}(x)$ est continue pour $x = a$, alors, quel que soit $\varepsilon > 0$, il existe $\delta > 0$ tel que la relation*

$$|R_p| \leqslant \varepsilon |h|^p$$

a lieu dès que $h \in X$, $h < \delta$, *c'est-à-dire que la quantité* $| R_p |$ *a un ordre de petitesse supérieur à celui de* $| h |^p$.

En effet, la fonction $f^{(p)}(x)$ étant continue au point $x = a$ et $\varepsilon > 0$ étant donné, il suffit de déterminer $\delta > 0$ de façon que $| h | < \delta$ implique $\| f^{(p)}(x) - f^{(p)}(a) \| < \varepsilon p!$.

2.72. Inversion du théorème de Taylor. La formule de Taylor montre que l'accroissement d'une fonction $f(x)$ p fois dérivable dans un domaine V peut être localement approché par une forme de degré p par rapport au vecteur déplacement h. La réciproque, a-t-elle lieu, si l'on a une fonction $f(x) : V \to Y$ dont l'accroissement $f(x + h) - f(x)$ est approximable par une forme de degré p par rapport au vecteur déplacement h au sens de représentation

$$f(x + h) = f(x) + a_1(x)\, h + a_2(x)\, hh + \ldots$$
$$\ldots + a_p(x)\, \underbrace{h \ldots h}_{p \text{ fois}} + R_p,$$

où

$$R_p = o(| h |^p),$$

alors peut-on dire que cette fonction soit p fois (au moins) dérivable? (Naturellement, les $a_1(x), \ldots, a_p(x)$ sont ici des opérateurs linéaires; $a_1(x) : V \to Y_1, \ldots, a_p(x) : V \to Y_p$, comme dans *2.24.*)

En général, c'est faux (exercice 6). Mais si l'on suppose que les fonctions $a_1(x), \ldots, a_p(x)$ soient continues et que le rapport $R_p / | h |^p$ tende vers 0 uniformément par rapport à x, alors *la proposition formulée s'avère valable.* Nous donnerons la démonstration plus bas, dans *2.74.*

2.73. Schémas aux différences.

a. Soit $a_1, \ldots, a_p$ une suite de p ($\geqslant 2$) éléments qui sont des nombres ou des vecteurs d'un espace vectoriel. Composons-en un « *schéma aux différences d'ordre* p », i.e. un tableau triangulaire

$$\begin{array}{llll} a_{11} & a_{12} & \ldots & a_{1,\,p-1}\ a_{1p}, \\ a_{21} & a_{22} & \ldots & a_{2,\,p-1}, \\ \ldots & \ldots & \ldots & \ldots \\ a_{p-1,\,1} & a_{p-1,\,2}, & & \\ a_{p1} & & & \end{array}$$

d'après la règle suivante:

$$a_{11} = a_1,\ a_{21} = a_2,\ \ldots,\ a_{p1} = a_p,$$
$$a_{12} = a_{21} - a_{11},\ a_{22} = a_{31} - a_{21},\ \ldots,\ a_{p-1,\,2} = a_{p1} - a_{p-1,\,1},$$
$$\ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots$$
$$a_{1p} = a_{2,\,p-1} - a_{1,\,p-1},$$

c'est-à-dire que la première colonne est formée des éléments $a_1, \ldots, a_p$ mêmes, et, à partir de la seconde colonne, chaque élément du tableau est la différence de deux éléments de la colonne précédente (l'élément de la ligne donnée est soustrait de celui de la ligne suivante). L'élément a_{1p} s'appelle *résultat du schéma aux différences.* Désignons-le par $Z = Z(a_1, \ldots, a_p)$. Nous allons mettre en évidence certaines propriétés de la fonction $Z(a_1, \ldots, a_p)$.

b. *La fonction* $Z(a_1, \ldots, a_p)$ *est une fonction linéaire de la suite* $a_1, \ldots, a_p$; *autrement dit, quels que soient* $a_1, \ldots, a_p, b_1, \ldots, b_p$ *et les nombres* α *et* β, *on a*

$$Z(\alpha a_1 + \beta b_1, \ldots, \alpha a_p + \beta b_p) = \\ = \alpha Z(a_1, \ldots, a_p) + \beta Z(b_1, \ldots, b_p).$$

En effet, toute opération de soustraction possède la propriété en question, il en est donc de même du résultat.

c. *Supposons que* k *désigne* 0 *ou un nombre naturel*; *alors, pour* $k < p - 1$, *on a*

$$Z(1^k, 2^k, \ldots, p^k) = 0.$$

Effectuons la démonstration par récurrence en p. Pour $p = 2$, la seule valeur possible de k est 0, et le calcul directe montre que

$$Z(1^0, 2^0) = 1 - 1 = 0.$$

Supposons que la proposition soit valable pour $p = 2, 3, \ldots, q - 1$; montrons qu'elle le reste pour $p = q$. Il s'agit de la quantité $Z(1^k, 2^k, \ldots, q^k)$, $k < q - 1$. Enlevons la première colonne du schéma aux différences correspondant; nous obtenons, évidemment, un autre schéma aux différences d'ordre $q - 1$ qui est déterminé par $q - 1$ éléments $2^k - 1^k, \ldots, q^k - (q-1)^k$ et qui a le même résultat. Remarquons maintenant que, pour tout $m = 1, \ldots, q - 1$,

$$(m+1)^k - m^k = km^{k-1} + \frac{k(k-1)}{1 \cdot 2} m^{k-2} + \ldots$$

Tous les exposants dans le second membre sont strictement inférieurs à $q - 2 = (q - 1) - 1$, puisque $k < q - 1$. Il en résulte d'après b et par hypothèse de récurrence que

$$Z(1^k, 2^k, \ldots, q^k) = Z(2^k - 1^k, \ldots, q^k - (q-1)^k) = \\ = kZ(1^{k-1}, 2^{k-1}, \ldots, (q-1)^{k-1}) = 0;$$

la proposition est démontrée.

d. *A lieu l'égalité*

$$Z(1^{p-1}, 2^{p-1}, \ldots, p^{p-1}) = (p-1)!.$$

En effet, composons les différences premières et appliquons, comme plus haut, la relation

$$(m+1)^{p-1} - m^{p-1} = (p-1)\, m^{p-2} + \frac{(p-1)(p-2)}{1\cdot 2}\, m^{p-3} + \ldots,$$
$$m = 1, \ldots, p-1.$$

Le résultat du schéma aux différences d'ordre $p-1$ formé par les termes du second membre à partir du deuxième est nul d'après c. On peut donc, sans nuire au résultat, remplacer la deuxième colonne du schéma initial par les termes $(p-1)\, m^{p-2}$. De même, dans la troisième colonne, on peut se borner aux termes $(p-1)(p-2)\, m^{p-3}$; en arrivant à la p-ème colonne, on obtient $(p-1)!$, ce qu'il fallait.

2.74. Abordons l'inversion du théorème de Taylor.

Théorème. *Supposons que, pour tous les $x \in V$ et pour tous les h admissibles, on ait la relation*

$$f(x+h) - f(x) = a_1(x)\, h + a_2(x)\, hh + \ldots + a_p(x) \underbrace{h \ldots h}_{p \text{ fois}} + R_p(x, h), \quad (1)$$

où les fonctions $a_1(x), \ldots, a_p(x)$ sont bornées et continues dans V, le reste $R_p(x, h)$ possédant la propriété suivante: quel que soit $\varepsilon > 0$, il existe $\delta > 0$ tel que, pour $x \in V$ et $|h| < \delta$ on a

$$|R_p(x, h)| < \varepsilon\, |h|^p. \quad (2)$$

Alors la fonction $f(x)$ est p fois dérivable dans le domaine V et $a_k(x) = \frac{1}{k!} f^{(k)}(x)$ $(k = 1, \ldots, p)$.

Démonstration. Comme l'expression (1) admet une partie linéaire principale $a_1(x)\, h$, la fonction $f(x)$ est dérivable *une fois* au moins et $f'(x) = a_1(x)$. Supposons qu'elle soit m fois dérivable $(m < p)$ et que les relations $a_1(x) = f'(x), \ldots, a_m(x) = \frac{1}{m!} f^{(m)}(x)$ aient lieu; montrons qu'elle est $(m+1)$-dérivable et que $a_{m+1} = \frac{f^{(m+1)}(x)}{(m+1)!}$. On peut mettre (1) sous la forme

$$f(x+h) - f(x) = f'(x)\, h + \frac{1}{2} f''(x)\, hh + \ldots + \frac{1}{m!} f^{(m)}(x) \underbrace{h \ldots h}_{m \text{ fois}} + a_{m+1}(x) \underbrace{h \ldots h}_{m+1 \text{ fois}} + R_{m+1}(x, h) \quad (3)$$

tout en sachant que, pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que

$$|R_{m+1}(x, h)| < \varepsilon\, |h|^{m+1} \quad (4)$$

du moment que $x \in V$ et $|h| < \delta$; ceci a lieu d'après (2) et parce que les fonctions $a_{m+2}(x), \ldots, a_p(x)$ sont bornées. A côté de (3) nous avons les relations :

$$f(x+h+k) - f(x+h) = f'(x+h)\,k + \frac{1}{2} f''(x+h)\,kk + \ldots$$

$$\ldots + \frac{1}{m!} f^{(m)}(x+h) \underbrace{k \ldots k}_{m \text{ fois}} +$$

$$+ a_{m+1}(x+h) \underbrace{k \ldots k}_{m+1 \text{ fois}} + R_{m+1}(x+h,\ k), \qquad (5)$$

où, pour $|k| < \delta$,

$$|R_{m+1}(x+h,\ k)| < \varepsilon\,|k|^{m+1}, \qquad (6)$$

et

$$f(x+h+k) - f(x) = f'(x)\,(h+k) + \frac{1}{2} f''(x)\,(h+k)\,(h+k) + \ldots$$

$$\ldots + \frac{1}{m!} f^{(m)}(x) \underbrace{(h+k) \ldots (h+k)}_{m \text{ fois}} +$$

$$+ a_{m+1}(x) \underbrace{(h+k) \ldots (h+k)}_{m+1 \text{ fois}} + R_{m+1}(x,\ h+k), \qquad (7)$$

où, pour $|h+k| < \delta$,

$$|R_{m+1}(x,\ h+k)| < \varepsilon\,|h+k|^{m+1}. \qquad (8)$$

En additionnant les relations (3) et (5), puis soustrayant (7) nous trouvons

$$0 = f'(x+h)\,k - f'(x)\,k + \frac{1}{2} f''(x+h)\,kk -$$

$$- \frac{1}{2} [f''(x)\,(h+k)\,(h+k) - f''(x)\,hh] + \ldots$$

$$\ldots + \frac{1}{m!} f^{(m)}(x+h) \underbrace{k \ldots k}_{m \text{ fois}} -$$

$$- \frac{1}{m!} [f^{(m)}(x) \underbrace{(h+k) \ldots (h+k)}_{m \text{ fois}} - f^{(m)}(x) \underbrace{h \ldots h}_{m \text{ fois}}] +$$

$$+ a_{m+1}(x+h) \underbrace{k \ldots k}_{m+1 \text{ fois}} - [a_{m+1}(x) \underbrace{(h+k) \ldots (h+k)}_{m+1 \text{ fois}} -$$

$$- a_{m+1}(x) \underbrace{h \ldots h}_{m+1 \text{ fois}}] + R'_{m+1}(x,\ h,\ k), \qquad (9)$$

où, pour $|h| < \delta,\ |k| < \delta,\ |h+k| < \delta$,

$$|R'_{m+1}(x,\ h,\ k)| \leqslant \varepsilon\,(|k|^{m+1} + |h|^{m+1} + |k+h|^{m+1}). \qquad (10)$$

La fonction $a_{m+1}(x)$ étant continue par hypothèse, on peut choisir un nombre $\delta_1 \leqslant \delta$ tel que, pour $|h| \leqslant \delta_1$, on aura

$$|a_{m+1}(x+h) - a_{m+1}(x)| \leqslant \varepsilon;$$

par conséquent, à la place de (9) on peut écrire:

$$\begin{aligned}
0 = {} & f'(x+h)\,k - f'(x)\,k + \frac{1}{2} f''(x+h)\,kk - \\
& - \frac{1}{2}[f''(x)(h+k)(h+k) - f''(x)\,hh] + \ldots \\
& \ldots + \frac{1}{m!} f^{(m)}(x+h) \underbrace{k \ldots k}_{m \text{ fois}} - \\
& - \frac{1}{m!}[f^{(m)}(x) \underbrace{(h+k) \ldots (h+k)}_{m \text{ fois}} - f^{(m)}(x) \underbrace{h \ldots h}_{m \text{ fois}}] + \\
& + a_{m+1}(x) \underbrace{k \ldots k}_{m+1 \text{ fois}} - a_{m+1}(x) \underbrace{(h+k) \ldots (h+k)}_{m+1 \text{ fois}} + \\
& + a_{m+1}(x) \underbrace{h \ldots h}_{m+1 \text{ fois}} + R''(x, h, k), \qquad (11)
\end{aligned}$$

où, dans les mêmes conditions imposées à x, h, k,

$$|R''(x, h, k)| \leqslant |R'_{m+1}(x, h, k)| + \varepsilon |k|^{m+1}. \qquad (12)$$

En remplaçant ici k par $2k$, $3k$, ..., $(m+1)\,k$ on obtient, à côté de (11), un système de $m+1$ relations:

$$\begin{aligned}
& f'(x+h)\,2k - f'(x)\,2k + \frac{1}{2} f''(x+h)\,2k\,2k - \\
& - \frac{1}{2}[f''(x)(h+2k)(h+2k) - f''(x)\,hh] + \ldots \\
& \ldots + \frac{1}{m!} f^{(m)}(x+h)\,2k \ldots 2k - \\
& - \frac{1}{m!}[f^{(m)}(x)(h+2k) \ldots (h+2k) - f^{(m)}(x) \underbrace{h \ldots h}_{m \text{ fois}}] + \\
& + a_{m+1}(x)\,2k \ldots 2k - a_{m+1}(x)(h+2k) \ldots (h+2k) + \\
& + a_{m+1}(x) \underbrace{h \ldots h}_{m+1 \text{ fois}} + R''(x, h, 2k) = 0,
\end{aligned}$$

. . . . . .

$$\begin{aligned}
& f'(x+h)(m+1)\,k - f'(x)(m+1)\,k + \frac{1}{2} f''(x+h)(m+1)\,k\,(m+1)\,k - \\
& - \frac{1}{2}[f''(x)(h+(m+1)\,k)(h+(m+1)\,k) - f''(x)\,hh] + \ldots \\
& \ldots + \frac{1}{m!} f^{(m)}(x+h) \underbrace{(m+1)\,k \ldots (m+1)\,k}_{m \text{ fois}} -
\end{aligned}$$

$$-\frac{1}{m!}[f^{(m)}(x)\underbrace{(h+(m+1)k)\ldots(h+(m+1)k)}_{m\text{ fois}}-$$

$$-f^{(m)}(x)\underbrace{h\ldots h}_{m\text{ fois}}]+$$

$$+a_{m+1}(x)\underbrace{(m+1)k\ldots(m+1)k}_{m+1\text{ fois}}-$$

$$-a_{m+1}(x)\underbrace{(h+(m+1)k)\ldots(h+(m+1)k)}_{m+1\text{ fois}}+$$

$$+a_{m+1}(x)\underbrace{h\ldots h}_{m+1\text{ fois}}+R''(x,\ h,\ (m+1)k)=0.$$

Transformons ce système. Soustrayons de chaque égalité, à partir de la deuxième, la précédente ; nous aboutissons à un autre système comprenant déjà m égalités. Le troisième système construit de façon analogue contient $m-1$ égalités ; enfin le dernier, le $(m+1)$-ème système en contiendra une seule que nous tenons à expliciter.

Il est évident que, pour chaque terme de ces égalités, on peut construire le résultat d'une manière indépendante, comme résultat du schéma aux différences correspondant de $(m+1)$-ème ordre. Les premiers termes forment la colonne (notée en ligne) :

$$\{1\cdot f'(x+h)k,\ 2f'(x+h)k,\ \ldots,\ (m+1)f'(x+h)k\},$$

le résultat du schéma aux différences correspondant étant nul d'après *2.73b* et *c*.

La situation est analogue pour tous les termes sauf les derniers. Notamment, il nous faut envisager les colonnes (notées ici en lignes) :

$$\frac{1}{m!}f^{(m)}(x+h)\underbrace{k\ldots k}_{m\text{ fois}}\{1^m,\ 2^m,\ldots,(m+1)^m\},\qquad(13)$$

$$\frac{1}{m!}f^{(m)}(x)\{\underbrace{(h+k)\ldots(h+k)}_{m\text{ fois}},\ \underbrace{(h+2k)\ldots(h+2k)}_{m\text{ fois}},\ \ldots$$

$$\ldots,\ \underbrace{(h+(m+1)k)\ldots(h+(m+1)k)}_{m\text{ fois}}\},\qquad(14)$$

$$a_{m+1}(x)\{\underbrace{k\ldots k}_{m+1\text{ fois}}-\underbrace{(h+k)\ldots(h+k)}_{m+1\text{ fois}},$$

$$2^m\underbrace{k\ldots k}_{m+1\text{ fois}}-\underbrace{(h+2k)\ldots(h+2k)}_{m+1\text{ fois}},\ \ldots,\ (m+1)^m\underbrace{k\ldots k}_{m+1\text{ fois}}-$$

$$-\underbrace{(h+(m+1)k)\ldots(h+(m+1)k)}_{m+1\text{ fois}}\}.\qquad(15)$$

En vertu de *2.73d*, où l'on pose $p = m + 1$, le résultat du schéma aux différences pour la colonne (13) vaut

$$f^{(m)}(x+h)\underbrace{k \ldots k}_{m \text{ fois}}.$$

En supprimant les parenthèses dans la colonne (14) et utilisant la multilinéarité de la forme $f^{(m)}(x)\, h \ldots h$, nous sommes en droit de ne garder que les termes $\frac{1}{m!} f^{(m)}(x)\, k \ldots k \{1^m, 2^m, \ldots$ $\ldots, (m+1)^m\}$ (les autres se réduisent en vertu de *2.73c*) ; d'après *2.73d*, nous aurons pour résultat

$$f^{(m)}(x)\underbrace{k \ldots k}_{m \text{ fois}}.$$

Dans les différences qui forment la colonne (15), les termes dominants, de degré $m + 1$ par rapport à k, se réduisent. On peut ne pas tenir compte des termes de degrés $\leqslant m - 1$, de même que dans la colonne (14). Ainsi, sans nuire au résultat du schéma aux différences correspondant, on peut remplacer la colonne (15) par celle-ci :

$$-a_{m+1}(x)\{(m+1)\underbrace{hk \ldots k}_{m \text{ fois}},\ 2^m(m+1)\underbrace{hk \ldots k}_{m \text{ fois}},\ \ldots$$

$$\ldots,\ (m+1)(m+1)^m\underbrace{hk \ldots k}_{m \text{ fois}}\},$$

dont le résultat vaut

$$-a_{m+1}(x)(m+1)!\ \underbrace{hk \ldots k}_{m \text{ fois}}.$$

Définitivement, nous arrivons à la relation

$$[f^{(m)}(x+h) - f^{(m)}(x)]\underbrace{k \ldots k}_{m \text{ fois}} = (m+1)!\ a_{m+1}(x)\underbrace{hk \ldots k}_{m \text{ fois}} + R \qquad (16)$$

dans laquelle

$$|R| \leqslant C_m \cdot 2\varepsilon(|k|^{m+1} + |h|^{m+1} + |k+h|^{m+1})$$

pour $|h| \leqslant \frac{\delta}{2(m+1)}$, $|k| \leqslant \frac{\delta}{2(m+1)}$. Fixons-y h. Alors, pour tous les k dont la norme est égale à celle de h, on a

$$|R| \leqslant C'_m \varepsilon |h|^{m+1}; \qquad (17)$$

or, ceci veut dire que la m-forme par rapport à k

$$\{[f^{(m)}(x+h) - f^{(m)}(x)] - (m+1)!\ a_{m+1}(x)\, h\}\underbrace{k \ldots k}_{m \text{ fois}}$$

ne dépasse pas en norme le nombre $C'_m \varepsilon |h|^{m+1}$ sur la sphère de rayon $|h|$. Par conséquent, la même forme ne dépasse pas en norme

le nombre $C'_m\varepsilon\,|h|$, sur la sphère de rayon 1. En raison de *2.24e*, nous obtenons l'inégalité

$$\|f^{(m)}(x+h)-f^{(m)}(x)-(m+1)!\,a_{m+1}(x)\,h\|\leqslant C''_m\varepsilon\,|h|$$

pour l'opérateur correspondant. Comme la constante C''_m ne dépend pas de h, *il en résulte la dérivabilité de la fonction* $f^{(m)}(x)$ et la relation

$$f^{(m+1)}(x)=(m+1)!\,a_{m+1}(x),$$

ce qu'il nous faut. En posant $m=1,\ \ldots,\ p-1$ nous achevons la démonstration du théorème.

EXERCICES

1. Montrer que si l'un au moins des déterminants de Sylvester d'ordre pair (*2.14f*) est négatif, alors la forme quadratique correspondante n'est pas définie (change de signe), indépendamment des valeurs et des signes des autres déterminants.

2. Ecrire la dérivée seconde de la fonction x^{-1} (*2.35*).

3. Ecrire la dérivée seconde de la fonction $y=f(x)$ inverse par rapport à une fonction inversible et deux fois dérivable $x=\varphi(y)$ (*2.36b*).

4. Supposons qu'un maximum lié d'une fonction $f(x)\colon G\to R_1$ sous une condition $\varphi(x)=C\in Z$ soit réalisé à un point $a\in G\subset X$ (*1.72a*). Supposons encore qu'une fonction $\psi(x)\colon G\to Z$ satisfasse aux relations $\psi(a)=\varphi(a)$, $\psi'(a)=\varphi'(a)$ et que l'opérateur $f'(a)\colon X\to X$ soit inversible.

Montrer que le point a est un point stationnaire (au sens d'extrémum lié) de la fonction $f(x)$ aussi bien sous la condition $\psi(x)=C$. Mais son caractère peut changer (maximum, minimum, absence d'extrémum).

5. Montrer que, pour qu'une famille biparamétrique de lignes hélicoïdales dans R_3 (φ est la variable, r et α les paramètres)

$$x=r\cos(\varphi-\alpha),\quad y=r\sin(\varphi-\alpha),\quad z=A(r)\,\varphi$$

admette les surfaces orthogonales, il faut et il suffit que $A(r)\equiv Cr^2$.

6. Etant donnée la fonction $f(x)=x^3\sin\dfrac{1}{x}$, on a le développement

$$f(x+h)=f(x)+a_1(x)\,h+a_2(x)\,\frac{h^2}{2}+o(h^2) \qquad (1)$$

pour $x\neq 0$, $|h|<|x|$. Pour $x=0$ on a également

$$h^3\sin\frac{1}{h}=0+0\cdot h+0\cdot\frac{h^2}{2}+o(h^2),$$

de sorte que l'égalité (1) est aussi bien valable pour $x=0$ si l'on pose $f(0)=0$, $a_1(0)=a_2(0)=0$. Or, $f(x)$ ne possède pas la dérivée seconde en $x=0$. Expliquer la contradiction apparente avec le théorème *2.74*.

7. Enoncer et démontrer la condition de résolubilité pour l'équation

$$\frac{\partial y}{\partial x_1}=\Phi(x,\,y),$$

où $X=X_1+X_2$ est une décomposition de l'espace X en somme droite.

8. On donne une application $y=y(x)$ d'un espace hilbertien H dans lui-même qui admet les dérivées première et seconde. On sait que $y'(x)=cT(x)$ où c est une constante, $T(x)$ un opérateur orthogonal. Montrer que $T(x)$ ne dépend pas de x et que l'application $y(x)$ est une combinaison de translation, dilatation et rotation.

9. On donne une application $y(x): H \to H$ admettant les dérivées première et seconde. On sait que $y'(x) = \dfrac{c}{|x - x_0|^2} T(x)$ où x_0 est un vecteur fixe, $T(x)$ un opérateur orthogonal. Montrer que $y(x)$ est une combinaison d'inversion

$$z(x) = \frac{x - x_0}{|x - x_0|^2},$$

translation, dilatation et rotation.

10. Une application $y(x): H \to H$ est dite conforme si $y'(x) = c(x)\, T(x)$ où $c(x)$ est une fonction numérique, $T(x)$ un opérateur orthogonal. Montrer que, pour une application conforme et deux fois dérivable $y(x)$ et pour n'importe quels trois vecteurs deux à deux orthogonaux h, k, l, a lieu la relation $(y''hk,\ y'l) = 0$.

11. (Suite.) Dans l'hypothèse de l'exercice précédent montrer que $y''hk = \mu y'h + \nu y'k$ et trouver les coefficients μ et ν.

12. (Suite.) Pour une application conforme trois fois dérivable $y(x)$ et pour $\rho(x) = \dfrac{1}{c(x)}$, montrer que $\rho''(x)\, hk = 0$, quels que soient deux vecteurs orthogonaux h et k.

13. (Suite.) Montrer que $\rho''hk = \sigma(h, k)$, où $\sigma = \text{const}$.

14. (Suite.) Montrer que $\rho(x) = \alpha |x - x_0|^2 + \beta$, où α et β sont des constantes.

15. (Suite.) Montrer que les points x et $y(x)$ ($y(x)$ est une application conforme) satisfont à la relation

$$(\alpha |x - x_0|^2 + \beta)(\gamma |y - y_0|^2 + \delta) = 1,$$

où x_0, y_0 sont des points fixes, γ, δ, des constantes.

16. (Suite.) Considérer l'équation

$$|dy| = c(x)\, |dx|$$

sur une demi-droite issue du point x_0 et, en l'intégrant, montrer que dans l'expression de la fonction $\rho(x)$ (exercice 14) ou bien $\alpha = 0$, ou bien $\beta = 0$.

Remarque. Ce résultat complété par ceux des exercices 8 et 9 montre que toute application conforme d'un espace hilbertien dans lui-même se réduit à une combinaison de translation, dilatation, rotation et inversion (Nevanlinna).

17. Le système

$$\frac{\partial z(x, y)}{\partial x} = z, \qquad \frac{\partial z(x, y)}{\partial y} = z^2$$

est compatible, il admet la solution immédiate $z(x, y) \equiv 0$. La condition nécessaire de compatibilité *2.61* (4) n'est pourtant pas satisfaite. Comment expliquer la contradiction apparente avec le théorème *2.61a*?

18. Montrer que, pour qu'un système d'équations

$$\frac{\partial z}{\partial x_1} = f_1(z), \ldots, \frac{\partial z}{\partial x_n} = f_n(z)$$

à une fonction inconnue $z = z(x_1, \ldots, x_n)$ ait une solution $z(x)$, quelles que soient les conditions initiales $z_0 = z(x_0)$, il faut et il suffit que toutes les fonctions $f_j(z)$ ($j = 1, \ldots, n$) ne diffèrent l'une de l'autre que par les facteurs numériques.

Historique

Les créateurs de l'analyse infinitésimale Newton et Leibniz appliquent les différentielles d'ordre supérieur (second) à déduire et à résoudre des équations différentielles ordinaires. L'étude générale des différentielles d'ordre supérieur est entreprise par Euler (1730) et achevée par Cauchy un siècle plus tard à l'aide de la théorie des limites. L'extension de la théorie sur les fonctions dans les espaces normés devient possible dès que Fréchet donne sa définition de la différentielle (1911). L'inversion de la formule de Taylor peut être trouvée, par exemple, dans le livre de Л. А. Люстерник и В. И. Соболев, Элементы функционального анализа (Москва, 1965). Le théorème classique de Frobenius (pour les fonctions $y\,(x)$: $R_n \to R_m$), qui a réalisé à son temps (1876) une étape importante du développement de la théorie générale des systèmes d'équations linéaires à dérivées partielles du premier ordre, est étendu sur les fonctions dans les espaces normés par M. Kerner (1933), pour l'équation $y' = \Phi\,(x, y)$ par J. Dieudonné (1960, cas général). Le théorème de Veblen et Thomas est publié en 1926.

Intégration dans les espaces à plusieurs dimensions

L'intégration des fonctions d'une variable multidimensionnelle est un des plus puissants moyens des mathématiques. Les théories abstraites modernes de l'intégration ont affaire à des fonctions sur un ensemble quelconque, pourvu qu'une mesure σ-additive y soit donnée. En visant les applications purement analytiques nous nous bornons à des ensembles relativement simples (« espaces chargés »), ce qui permet de construire une théorie d'intégration par analogie avec l'intégrale de Riemann unidimensionnelle, sans toucher à la σ-additivité de la mesure. L'application d'une telle théorie à l'espace euclidien n-dimensionnel conduit à la théorie classique des intégrales multiples (§ 3.5), des intégrales de surfaces (§ 3.6), après quoi nous considérons les intégrales impropres (§ 3.7).

Pour ne pas surcharger l'exposé, nous étudions les fonctions à valeurs réelles. Néanmoins, toutes les conclusions restent valables pour les fonctions à valeurs dans un espace vectoriel normé, les rares exceptions étant du type (*3.14* (7))

$$mX \inf_X f(x) \leqslant \int_X f(x)\,dx \leqslant mX \sup_X f(x).$$

Or, cette formule admet elle aussi un homologue dans le cas de fonctions à valeurs dans un espace vectoriel normé (*3.14* (8)).

§ 3. 1. Intégrale de Riemann sur un espace chargé

3.11. Avant d'aborder la construction de l'intégrale pour les fonctions d'une variable multidimensionnelle, rappelons la définition de l'intégrale d'une fonction $f(x)$ d'une variable x parcourant un intervalle $a \leqslant x \leqslant b$ (*U9.13*).

Désignons par Π une partition de l'intervalle $a \leqslant x \leqslant b$:

$$\Pi = \{a = x_0 \leqslant x_1 \leqslant \ldots \leqslant x_i \leqslant x_{i+1} \leqslant \ldots \leqslant x_n = b\}.$$

Posons $\Delta x_i = x_{i+1} - x_i$ et $d(\Pi) = \max_i \Delta x_i$. Choisissons dans l'intervalle $x_i \leqslant x \leqslant x_{i+1}$ un point quelconque ξ_i et formons la somme intégrale :

$$S_\Pi(f) = \sum_{i=0}^{n-1} f(\xi_i)\,\Delta x_i.$$

Un nombre If s'appelle *intégrale de Riemann de la fonction* $f(x)$ *sur l'intervalle* $[a, b]$ si, pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que l'inégalité

$$| If - S_{\Pi}(f) | < \varepsilon$$

a lieu pour toute partition Π avec $d(\Pi) < \delta$.

Une définition équivalente peut être exprimée en termes de suites. Considérons une suite de partitions $\Pi_1, \Pi_2, \ldots$ de l'intervalle $[a, b]$, telle que $d(\Pi_k) \to 0$; nous l'appelons *morcellement illimité d'une partition.* Lorsque, pour toute suite $\{\Pi_k\}$ de ce type, les nombres $S_{\Pi_k}(f)$ tendent à une limite fixe qui ne dépend pas du choix de la suite Π_k ni des points marqués ξ_i, cette limite est appelée *intégrale de la fonction* $f(x)$ *sur l'intervalle* $[a, b]$:

$$\int_a^b f(x)\,dx = \lim_{k\to\infty} S_{\Pi_k}(f).$$

Signalons, enfin, encore une définition équivalente que l'on exprime en termes de limite suivant une direction (*U4.12*). Soit E l'ensemble des partitions Π à points marqués; pour un $\delta > 0$ donné, désignons par E_δ le sous-ensemble formé des partitions $\Pi \in E$ avec $d(\Pi) < \delta$. Les sous-ensembles E_δ, pour δ différents, déterminent une direction sur E que nous désignons par $d(\Pi) \to 0$. L'intégrale de la fonction $f(x)$ est *la limite* Jf *de ses sommes intégrales suivant cette direction*; elle existe en même temps que les intégrales If et $\int_a^b f(x)\,dx$ et a la même valeur.

Il est, par exemple, démontré que l'intégrale d'une fonction continue $f(x)$ existe.

3.12. Passons à la définition générale de l'intégrale de Riemann. A la place de l'intervalle $a \leqslant x \leqslant b$, nous considérons un ensemble quelconque X. A la place d'un ensemble de parties de l'intervalle $[a, b]$, nous introduisons un système $\mathfrak{A} = \mathfrak{A}(X)$ de sous-ensembles de X qui satisfont aux conditions suivantes:

a. L'ensemble X lui-même et l'ensemble vide appartiennent au système $\mathfrak{A}$.

b. Si $A_1 \in \mathfrak{A}$, $A_2 \in \mathfrak{A}$, alors l'intersection $A_1 A_2$ appartient elle aussi à $\mathfrak{A}$.

c. Si $A_1 \in \mathfrak{A}$, $A \in \mathfrak{A}$, $A_1 \subset A$, alors il existe dans $\mathfrak{A}$ des ensembles $A_2, \ldots, A_p$ tels que $A = A_1 \cup A_2 \cup \ldots \cup A_p$ et que $A_1, \ldots, A_p$ sont deux à deux disjoints.

Un système $\mathfrak{A}$ de sous-ensembles de X qui vérifie les conditions a, b, c est appelé *demi-anneau.* Ainsi, le système de tous les intervalles faisant partie d'un intervalle $[a, b]$ représente un demi-anneau d'ensembles.

Dans ce qui suit, nous supposons que X soit un espace métrique et que le demi-anneau $\mathfrak{A}(X)$ satisfasse à la condition:

d. Pour tout $\delta > 0$, il existe une partition de l'ensemble X en une réunion finie d'ensembles $A_1, \ldots, A_p$ de $\mathfrak{A}(X)$, ces ensembles étant deux à deux disjoints et leurs diamètres ne dépassant pas δ.

Une fois la condition *d* remplie, l'espace métrique X est précompact (*U3.93*), car il admet, pour tout $\delta > 0$, un δ-réseau fini.

Formulons enfin la dernière condition imposée au demi-anneau $\mathfrak{A}$:

e. Pour tout ensemble $A \in \mathfrak{A}$, il est défini un nombre non négatif mA de façon que, étant donnée une partition

$$A = A_1 \cup A_2 \cup \ldots \cup A_p,$$

où $A_1, \ldots, A_p$ sont deux à deux disjoints et appartiennent à $\mathfrak{A}$, on a

$$mA = mA_1 + \ldots + mA_p$$

(*condition d'additivité*).

Le nombre mA s'appelle *mesure* de l'ensemble A. Les ensembles $A \in \mathfrak{A}$ eux-mêmes, si les conditions *a-e* sont remplies, s'appellent *cellules*, l'espace métrique X *espace chargé* et le demi-anneau $\mathfrak{A}$ avec la mesure des cellules mA, *charge* de l'espace X.

3.13. Définition de l'intégrale. Soit $f(x)$ une fonction définie sur un espace chargé X. Désignons par Π une partition de l'ensemble X en p cellules $A_1, \ldots, A_p$ deux à deux disjointes. Pour chacune de ces cellules A_i, désignons par $d(A_i)$ son diamètre, i.e. la borne supérieure exacte des distances entre les points de la cellule; par $d(\Pi)$ la valeur maximale des diamètres des cellules A_i.

Choisissons dans toute cellule A_i un point quelconque ξ_i et composons la *somme intégrale*

$$S_\Pi(f) = \sum_{i=1}^{p} f(\xi_i)\, mA_i. \tag{1}$$

Le nombre

$$I_{\mathfrak{A}} f = \int_{X,\,\mathfrak{A}} f(x)\, dx \tag{2}$$

s'appelle *intégrale de la fonction $f(x)$ sur l'espace X avec la charge $\mathfrak{A}$* si, pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que l'inégalité

$$| I_{\mathfrak{A}} f - S_\Pi(f) | < \varepsilon$$

a lieu pour toute partition Π avec $d(\Pi) < \delta$.

Nous voyons que cette définition est parfaitement analogue à la première définition de l'intégrale d'une fonction $f(x)$ sur un intervalle fermé.

En continuant le raisonnement par analogie, formulons la deuxième définition en termes de suites. Notamment, considérons une suite arbitraire $\Pi_1, \ldots, \Pi_k \ldots$ de partitions de l'espace X vérifiant la condition $d(\Pi_k) \to 0$; appelons-la *morcellement illimité de la partition*. Si, pour toute suite $\{\Pi_k\}$ de ce type, les nombres $S_{\Pi_k}(f)$ tendent vers une limite fixe qui ne dépend pas du choix de la suite $\{\Pi_k\}$ et des points ξ_i, cette limite est appelée *intégrale de la fonction f sur l'espace chargé X*.

Passons enfin à la définition en termes de limite suivant une direction. Soit E l'ensemble de toutes les partitions (à points marqués) de l'espace X ; pour $\delta > 0$ donné, désignons par E_δ l'ensemble des partitions pour lesquelles $d(\Pi) < \delta$. Deux δ étant différents, les ensembles correspondants E_δ sont inclus l'un dans l'autre ; l'intersection de tous les E_δ est vide ; par conséquent, les ensembles E_δ déterminent une direction sur E ; nous la désignons, naturellement, par $d(\Pi) \to 0$. Par définition, *l'intégrale de la fonction* $f(x)$ *est la limite des sommes intégrales suivant la direction* $d(\Pi) \to 0$.

L'équivalence de toutes les trois définitions formulées résulte des propriétés fondamentales de la limite suivant une direction (cf. *U4.71*).

Toute fonction $f(x)$ possédant l'intégrale sur un espace X avec une charge $\mathfrak{A}$ est dite *intégrable sur X avec la charge* $\mathfrak{A}$ ou *intégrable* tout court, si l'espace X et la charge $\mathfrak{A}$ sont fixes ; dans le dernier cas nous omettons le symbole $\mathfrak{A}$ dans la désignation de l'intégrale.

3.14. Signalons quelques propriétés fondamentales de l'intégrale dans la seule hypothèse qu'elle existe, sans nous baser sur les propriétés des fonctions intégrées. Nous ne donnons pas les démonstrations car elles suivent la même voie que dans *U9.15* où l'on considérait les intégrales sur un intervalle fermé : le passage à la limite des sommes intégrales.

a. *Une fonction* $f(x) \equiv C$ (*constante*) *est intégrable sur un espace chargé X, et*

$$\int_X f(x)\,dx = CmX. \tag{1}$$

b. *Si une fonction* $f(x)$ *est intégrable sur un espace chargé X, alors, quelle que soit une constante C, la fonction* $Cf(x)$ *est elle aussi intégrable sur X, et*

$$\int_X Cf(x)\,dx = C\int_X f(x)\,dx. \tag{2}$$

c. *Si deux fonctions* $f(x)$ *et* $g(x)$ *sont intégrables sur un espace chargé X, leur somme* $f(x) + g(x)$ *est elle aussi intégrable sur X, et*

$$\int_X [f(x) + g(x)]\,dx = \int_X f(x)\,dx + \int_X g(x)\,dx. \tag{3}$$

d. *Toute fonction intégrable sur un espace chargé X est bornée sur X.*

e. *Si deux fonctions $f(x)$ et $g(x)$ intégrables sur l'espace X satisfont à l'inégalité $f(x) \leqslant g(x)$, alors*

$$\int_X f(x)\,dx \leqslant \int_X g(x)\,dx. \tag{4}$$

En particulier, si les fonctions $f(x)$ et $|f(x)|$ sont intégrables sur X, alors

$$\left|\int_X f(x)\,dx\right| \leqslant \int_X |f(x)|\,dx. \tag{5}$$

Si les valeurs d'une fonction $f(x)$ intégrable sur l'espace X sont comprises dans les limites $c \leqslant f(x) \leqslant C$, alors

$$cmX \leqslant \int_X f(x)\,dx \leqslant CmX. \tag{6}$$

Par exemple, on a toujours

$$\inf_X f(x) \cdot mX \leqslant \int_X f(x)\,dx \leqslant \sup_X f(x) \cdot mX. \tag{7}$$

Pour les fonctions à valeurs dans un espace normé, les formules (6) et (7) cèdent leur place à la formule (*U.12.62g*) :

$$\frac{1}{mX}\int_X f(x)\,dx \in \overline{V(E)}, \tag{8}$$

où E est l'ensemble des valeurs de la fonction $f(x)$ sur X et $\overline{V(E)}$ l'enveloppe convexe fermée de l'ensemble E.

f. Signalons enfin encore un théorème dont la démonstration répète mot à mot celle du théorème *U9.72* en remplaçant l'intervalle $[a, b]$ par un espace chargé X.

Théorème. *Si une suite $f_1(x)$, $f_2(x)$, . . . de fonctions intégrables converge uniformément sur l'espace chargé X vers une fonction $f(x)$, alors $f(x)$ est elle aussi intégrable et*

$$\int_X f(x)\,dx = \lim_{n\to\infty} \int_X f_n(x)\,dx.$$

Un théorème analogue a lieu pour la série $\varphi_1(x) + \varphi_2(x) + \ldots$ de fonctions intégrables.

3.15. Quelques réalisations concrètes de l'intégrale.

a. Si l'on choisit pour espace X un intervalle $[a, b]$, pour cellules A n'importe quels intervalles (contenant ou non leurs extrémités)

et pour mesure d'une cellule A la longueur de l'intervalle correspondant, on retrouvera la définition ordinaire de l'intégrale d'une fonction d'une variable.

b. Soit X un pavé dans un espace n-dimensionnel:

$$X = \{x \in R_n : a_1 \leqslant x_1 \leqslant b_1, \ldots, a_n \leqslant x_n \leqslant b_n\}.$$

En tant que cellules, choisissons tous les sous-pavés $A \subset X$:

$$A = \{x \in X : \alpha_1 \leqslant x_1 \leqslant \beta_1, \ldots, \alpha_n \leqslant x_n \leqslant \beta_n\}$$

et tous les ensembles qui se déduisent des ensembles définis ci-dessus en remplaçant certains des symboles $\leqslant$ par $<$. Appelons mesure d'une cellule A son volume euclidien

$$mA = \prod_{i=1}^{n} (\beta_i - \alpha_i).$$

On peut prouver sans trop de peine que les conditions *3.12 a-e* sont satisfaites. Remarquons tout de même que la preuve formelle de la condition *3.12e* est assez laborieuse, mais cela ne nous empêche pas de l'omettre pour le moment, puisque, dans *3.17*, nous démontrerons une proposition plus générale. L'intégrale que l'on obtient avec cette construction s'appelle *intégrale de Riemann n-uple* et se désigne par

$$\int_X f(x)\,dx = \int_{a_1}^{b_1} \ldots \int_{a_n}^{b_n} f(x_1, \ldots, x_n)\,dx_1 \ldots dx_n.$$

3.16. Lemmes sur les demi-anneaux. Le dernier exemple admet une généralisation essentielle: ayant donnés des espaces chargés $X_1, \ldots, X_n$, on peut construire l'espace chargé $X = \prod_{i=1}^{n} X_i$ de la même façon que l'on construit un pavé n-dimensionnel avec les intervalles unidimensionnels. Afin de réaliser cette idée, ce qu'on fait dans *3.17*, il faut une préparation.

a. Lemme. *Si les ensembles $A_1, \ldots, A_k$ d'un demi-anneau $\mathfrak{A}$ sont deux à deux disjoints et contenus dans un ensemble $A \in \mathfrak{A}$, il existe dans $\mathfrak{A}$ des ensembles $B_{k+1}, \ldots, B_r$ tels que*

$$A_1 \cup \ldots \cup A_k \cup B_{k+1} \cup \ldots \cup B_r = A, \qquad (1)$$

les ensembles dans le premier membre étant eux aussi disjoints deux à deux.

Démonstration. Pour $k = 1$, la proposition est contenue dans la définition d'un demi-anneau (*3.12c*). Supposons qu'elle soit valable pour un nombre k et démontrons-la pour le nombre $k + 1$. Nous avons les ensembles disjoints $A_1, \ldots, A_k, A_{k+1}$ appartenant au demi-anneau $\mathfrak{A}$ et contenus dans l'ensemble $A \in \mathfrak{A}$. Par supposition de récurrence, il existe dans $\mathfrak{A}$ des sous-ensembles

$B_{k+1}, \ldots, B_r$ réalisant la partition (1). L'ensemble A_{k+1} est disjoint avec chacun des ensembles $A_1, \ldots, A_k$, donc contenu dans la réunion $B_{k+1} \cup \ldots \cup B_r$, ce qui implique

$$A_{k+1} = A_{k+1}B_{k+1} \cup \ldots \cup A_{k+1}B_r. \tag{2}$$

Par définition d'un demi-anneau, on a les partitions

$$\left.\begin{aligned} B_{k+1} &= A_{k+1}B_{k+1} \cup B_{k+1}^{(1)} \cup \ldots \cup B_{k+1}^{(p_{k+1})}, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ B_r &= A_{k+1}B_r \cup B_r^{(1)} \cup \ldots \cup B_r^{(p_r)}, \end{aligned}\right\} \tag{3}$$

où les ensembles dans les seconds membres appartiennent à $\mathfrak{A}$ et sont disjoints deux à deux. En portant (3) dans (1) et utilisant (2), nous arrivons à la relation

$$A = A_1 \cup \ldots \cup A_k \cup A_{k+1}B_{k+1} \cup \ldots$$
$$\ldots \cup A_{k+1}B_r \cup B_{k+1}^{(1)} \cup \ldots \cup B_r^{(p_r)} =$$
$$= A_1 \cup \ldots \cup A_k \cup A_{k+1} \cup B_{k+1}^{(1)} \cup \ldots \cup B_r^{(p_r)},$$

ce qu'il nous fallait.

b. L e m m e. *Soit un ensemble fini $A_1, \ldots, A_k$ d'ensembles d'un demi-anneau $\mathfrak{A}$. Il existe dans $\mathfrak{A}$ des ensembles disjoints $B_1, \ldots, B_p$ tels que $\bigcup_{i=1}^{k} A_i = \bigcup_{j=1}^{p} B_j$ et que chacun des ensembles A_i $(i = 1, \ldots, k)$ représente une réunion de certains ensembles B_j.*

D é m o n s t r a t i o n. Pour $k = 1$ la proposition est immédiate : on peut poser $B_1 = A_1$. En la supposant valable pour un k naturel démontrons qu'elle reste valable pour $k+1$. Par supposition de récurrence, il existe dans $\mathfrak{A}$ des ensembles disjoints $B_1^{(k)}, \ldots, B_p^{(k)}$ tels que $\bigcup_{i=1}^{k} A_i = \bigcup_{j=1}^{p} B_j^{(k)}$ et que chaque ensemble A_i, $i = 1, \ldots, k$ est une réunion de certains ensembles $B_j^{(k)}$. Par définition d'un demi-anneau, on peut écrire, pour tout $B_j^{(k)}$, sa partition en ensembles disjoints dont le premier est $B_j^{(k)} \cap A_{k+1}$:

$$B_j^{(k)} = (B_j^{(k)} \cap A_{k+1}) \cup B_{j1} \cup \ldots \cup B_{jp_j} \tag{4}$$
$$(j = 1, \ldots, p).$$

Ensuite, d'après le lemme *a*, l'ensemble A_{k+1} peut être mis sous la forme de réunion d'ensembles disjoints de $\mathfrak{A}$ dont p premiers sont les intersections de A_{k+1} avec les ensembles $B_j^{(k)}$, $j = 1, \ldots, p$:

$$A_{k+1} = (A_{k+1} \cap B_1^{(k)}) \cup \ldots \cup (A_{k+1} \cap B_p^{(k)}) \cup B_1 \cup \ldots \cup B_r. \tag{5}$$

On voit que tous les ensembles que l'on réunit dans les seconds membres de (4) et (5) appartiennent au demi-anneau $\mathfrak{A}$ et sont deux à deux disjoints. Il est clair qu'en réunissant certains de ces ensembles on peut obtenir chacun des ensembles $A_1, \ldots, A_{k+1}$. Le lemme est démontré.

3.17. Produit d'espaces chargés. Soient $X_1, \ldots, X_n$ des espaces chargés; considérons le produit Z des ensembles $X_1, \ldots, X_n$, i.e. l'ensemble de tous les complexes $x = (x_1, \ldots, x_n)$. Introduisons-y une structure d'espace chargé. Pour ne pas compliquer l'exposé, bornons-nous au cas $n = 2$ et notons $X_1 = X$, $X_2 = Y$.

Appelons cellules de l'espace $Z = X \times Y$ les produits $C = A \times B$ des cellules des espaces X et Y respectivement. L'ensemble de ces nouvelles cellules désignons par $\mathfrak{A}(Z)$. Montrons que $\mathfrak{A}(Z)$ est encore un demi-anneau. Tout l'espace $Z = X \times Y$ appartient au système $\mathfrak{A}(Z)$ car $X \in \mathfrak{A}(X)$, $Y \in \mathfrak{A}(Y)$. Si $C_1 = A_1 \times B_1$, $C_2 = A_2 \times B_2$ sont deux cellules sur Z, alors $C_1 C_2 = A_1 A_2 \times B_1 B_2$ en est encore une. Soient $C_1 = A_1 \times B_1 \in \mathfrak{A}(Z)$ et $C = A \times B \in \mathfrak{A}(Z)$, de sorte que $C_1 \subset C$. Cela veut dire que $A_1 \subset A$, $B_1 \subset B$. Dans ce cas, nous trouvons des cellules disjointes $A_2, \ldots, A_p$ de $\mathfrak{A}(X)$ et $B_2, \ldots, B_q$ de $\mathfrak{A}(Y)$ telles que

$$A = A_1 \cup A_2 \cup \ldots \cup A_p, \quad B = B_1 \cup B_2 \cup \ldots \cup B_q.$$

L'expression

$$A \times B = (A_1 \times B_1) \cup (A_1 \times B_2) \cup \ldots \cup (A_1 \times B_q) \cup$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\cup (A_p \times B_1) \cup (A_p \times B_2) \cup \ldots \cup (A_p \times B_q)$$

représente la cellule $A \times B$ comme réunion de la cellule $A_1 \times B_1$ avec certaines autres cellules disjointes. Il en découle que $\mathfrak{A}(Z)$ est un demi-anneau, les axiomes *3.12 a-c* étant remplis.

L'espace Z est métrisé de la façon ordinaire, comme produit d'espaces métriques, par exemple d'après la formule (*U3.16*):

$$\rho(x_1 \times y_1, x_2 \times y_2) = \sqrt{\rho^2(x_1, x_2) + \rho^2(y_1, y_2)}.$$

Si $X = A_1 \cup \ldots \cup A_p$ et $Y = B_1 \cup \ldots \cup B_q$ sont les décompositions des espaces X et Y respectivement en cellules disjointes de diamètres $\leqslant \delta$, alors $Z = (A_1 \times B_1) \cup (A_1 \times B_2) \cup \ldots \cup (A_p \times B_q)$ représente une décomposition de l'espace Z en cellules disjointes de diamètres $\leqslant \delta\sqrt{2}$, de sorte que l'axiome *3.12d* se trouve lui aussi rempli.

Enfin, pour une cellule $C = A \times B$, posons $mC = mA \cdot mB$. Montrons que la mesure des cellules dans Z ainsi définie satisfait à l'axiome de l'additivité *3.12e*. Soit

$$C = A \times B = (A_1 \times B_1) \cup \ldots \cup (A_k \times B_k) \qquad (1)$$

une partition de la cellule C en cellules disjointes. Mettons A sous la forme de réunion $A = A_1 \cup \ldots \cup A_k$ d'ensembles qui ne sont pas obligatoirement disjoints.

D'après *3.16b*, il existe dans $\mathfrak{A}(X)$ des cellules disjointes $\widetilde{A}_1, \ldots, \widetilde{A}_p$ telles que $A = \bigcup_{j=1}^{p} \widetilde{A}_j$ et que chaque cellule A_i est la réunion de certaines cellules $\widetilde{A}_j$. De même, il existe dans $\mathfrak{A}(Y)$ des cellules disjointes $\widetilde{B}_1, \ldots, \widetilde{B}_q$ telles que $B = \bigcup_{1}^{q} \widetilde{B}_j$ et que chaque cellule B_i est la réunion de certaines cellules $\widetilde{B}_j$. Nous notons les décompositions correspondantes comme suit:

$$A_i = \bigcup_{r} \widetilde{A}_r^{(i)}, \quad B_j = \bigcup_{s} \widetilde{B}_s^{(j)}.$$

L'expression (1) se met alors sous la forme

$$C = A \times B = (\widetilde{A}_1^{(1)} \times \widetilde{B}_1^{(1)}) \cup (\widetilde{A}_2^{(1)} \times \widetilde{B}_1^{(1)}) \cup \ldots \cup (\widetilde{A}_r^{(k)} \times \widetilde{B}_s^{(k)})$$

et l'on a

$$mC = mA \cdot mB = \sum_{i=1}^{p} m\widetilde{A}_i \cdot \sum_{j=1}^{q} m\widetilde{B}_j.$$

D'autre part,

$$\sum_{i=1}^{k} m(A_i \times B_i) = \sum_{i=1}^{k} mA_i \times mB_i =$$

$$= \sum_{i=1}^{k} \left(\sum_{r} m\widetilde{A}_r^{(i)}\right)\left(\sum_{s} m\widetilde{B}_s^{(i)}\right) = \sum_{i=1}^{p} \sum_{j=1}^{q} m\widetilde{A}_i m\widetilde{B}_j$$

car, les cellules $A_i \times B_i$ étant disjointes, chaque terme à droite est égal à un terme à gauche, et vice versa. D'où l'égalité cherchée:

$$mC = \sum_{i=1}^{k} m(A_i \times B_i).$$

Ainsi donc, un produit d'espaces chargés admet lui aussi une structure d'espace chargé. La structure décrite ici s'appelle *produit des charges des espaces X et Y*. La construction peut être étendue par récurrence à n'importe quel nombre de facteurs.

3.18. Ensembles élémentaires.

a. Nous appelons *ensembles élémentaires* les réunions finies de cellules d'un espace chargé.

En vertu de *3.16b*, tout ensemble élémentaire P peut être représenté comme réunion finie de cellules disjointes.

b. Lemme. *S'il y a des ensembles élémentaires P et Q, et que $P = A_1 \cup \ldots \cup A_p$, $Q = B_1 \cup \ldots \cup B_q$ sont leurs représentations en forme de réunions de cellules deux à deux disjointes, alors l'inclusion $P \subset Q$ implique*

$$\sum_{i=1}^{p} mA_i \leqslant \sum_{j=1}^{q} mB_j. \tag{1}$$

Démonstration. Comme $P \subset Q$, on a $A_i \subset Q$ d'où

$$A_i = A_iQ = \bigcup_{j=1}^{q} A_iB_j,$$

les termes à droite étant disjoints. Il en résulte d'après *3.12e*:

$$mA_i = \sum_{j=1}^{q} m(A_iB_j).$$

Pour j fixe, les cellules A_iB_j sont disjointes; d'après *3.12c*, il existe des cellules disjointes $B_j^{(1)}, \ldots, B_j^{(r_j)}$ telles que

$$B_j = (A_iB_j) \cup \ldots \cup (A_pB_j) \cup B_j^{(1)} \cup \ldots \cup B_j^{(r_j)} \quad (j = 1, \ldots q).$$

Vu *3.12e*, on en tire

$$mB_j = \sum_{i=1}^{p} m(A_iB_j) + mB_j^{(1)} + \ldots + mB_j^{(r_j)} \geqslant \sum_{i=1}^{p} m(A_iB_j).$$

Par conséquent,

$$\sum_{j=1}^{q} mB_j \geqslant \sum_{j=1}^{q} \sum_{i=1}^{p} m(A_iB_j) = \sum_{i=1}^{p} \Big(\sum_{j=1}^{q} m(A_iB_j) \Big) = \sum_{i=1}^{p} mA_i,$$

ce qu'il fallait démontrer.

c. En particulier, si l'on pose $P = Q$, on a:

$$\sum_{i=1}^{p} mA_i = \sum_{j=1}^{q} mB_j. \tag{2}$$

Nous pouvons donc définir une mesure mP de tout ensemble élémentaire P en tant que somme des mesures des cellules disjointes qui le composent; la relation (2) assure que cette définition soit correcte.

d. L'inégalité (1) a également lieu dans le cas où les cellules B_j ne sont pas disjointes (en conservant les autres conditions). En effet, d'après *3.16b* on peut trouver une collection de cellules disjointes $\widetilde{B}_r$ $(r = 1, \ldots, s)$ de façon que $\bigcup_j B_j = \bigcup_r \widetilde{B}_r$ et que chaque B_j est réunion de certains $\widetilde{B}_r$. La mesure de chaque B_j vaut la somme des mesures de toutes les $\widetilde{B}_r$ qui y sont inclues, et la somme des mesures

de toutes les B_j est au moins égale à la somme des mesures de toutes les $\widetilde{B}_r$, puisque chaque cellule $\widetilde{B}_r$ est inclue dans l'une des B_j. Nous avons maintenant:

$$\bigcup_{i=1}^{p} A_i \subset \bigcup_{j=1}^{q} B_j = \bigcup_{r=1}^{s} \widetilde{B}_r$$

et, en vertu de a,

$$\sum_{i=1}^{p} mA_i \leqslant \sum_{r=1}^{s} m\widetilde{B}_r \leqslant \sum_{j=1}^{q} mB_j,$$

ce qu'il fallait.

e. *Si un ensemble élémentaire P est la réunion de certaines cellules, disjointes ou non, $B_1, \ldots, B_q$, alors on a l'inégalité*

$$mP \leqslant \sum_{j=1}^{q} mB_j.$$

En effet, il existe d'après d des ensembles disjoints $A_1, \ldots, A_p$ tels que $\bigcup_i A_i = P = \bigcup_j B_j$; la définition c et le résultat d impliquent alors

$$mP = \sum_{i=1}^{p} mA_i \leqslant \sum_{j=1}^{q} mB_j,$$

ce qu'il fallait.

§ 3. 2. Théorèmes d'existence

Dans le présent paragraphe, on démontrera l'intégrabilité des fonctions continues ainsi que des fonctions dont l'ensemble des points de discontinuité est réduit (au sens qui sera précisé).

3.21. Nous montrerons ici qu'en démontrant l'existence de l'intégrale on peut se borner aux partitions qui suivent les partitions suffisamment morcelées.

Comme d'ordinaire, une partition Π' d'un ensemble X est dite *suivante* par rapport à une partition Π si les cellules de la partition Π sont des réunions (sans point commun) des cellules de la partition Π'. Pour toute partition Π et tout $\delta > 0$, il existe une partition suivante Π' avec $d(\Pi') < \delta$; afin d'en construire une, considérons n'importe quelle partition Π'' avec $d(\Pi'') < \delta$ qui existe d'après *3.12d* et composons la partition Π' de toutes les intersections des cellules de Π avec celles de Π''.

a. L e m m e. *Supposons qu'on ait une partition Π et que, pour un $\varepsilon > 0$ et pour toute partition suivante Π', soit valable l'inégalité*

$$| S_{\Pi}(f) - S_{\Pi'}(f) | \leqslant \varepsilon.$$

Si Π_1 *est une autre partition possédant la même propriété, i.e. telle que, pour toute partition suivante* Π'_1, *on a*

$$| S_{\Pi_1}(f) - S_{\Pi'_1}(f) | \leqslant \varepsilon,$$

alo s

$$| S_{\Pi}(f) - S_{\Pi_1}(f) | \leqslant 2\varepsilon.$$

Démonstration. Considérons une nouvelle partition Π_2 formée des intersections des cellules de Π et Π_1, les points marqués étant choisis arbitrairement. La partition Π_2 est suivante par rapport à Π et Π_1. On a par hypothèse:

$$| S_{\Pi}(f) - S_{\Pi_2}(f) | \leqslant \varepsilon,$$
$$| S_{\Pi_1}(f) - S_{\Pi_2}(f) | \leqslant \varepsilon,$$

d'où la relation cherchée.

b. Conséquence. *Si, pour tout* $\varepsilon > 0$, *il existe* $\delta > 0$ *tel que, quelles que soient une partition* Π *d'un espace chargé* X *avec* $d(\Pi) < \delta$ *et une partition suivante* Π', *a lieu l'inégalité*

$$| S_{\Pi}(f) - S_{\Pi'}(f) | < \varepsilon,$$

alors la fonction $f(x)$ *est intégrable sur l'espace* X.

En effet, dans l'hypothèse formulée l'inégalité

$$| S_{\Pi}(f) - S_{\Pi_1}(f) | < 2\varepsilon$$

a lieu d'après *a*, quelles que soient les partitions Π et Π_1 avec $d(\Pi) < \delta$ et $d(\Pi_1) < \delta$. Ainsi, on peut appliquer le critère de Cauchy aux sommes intégrales $S_{\Pi}(f)$ et à la direction $d(\Pi) \to 0$, ce qui prouve l'existence de l'intégrale.

3.22. Théorème d'existence de l'intégrale d'une fonction continue.

a. Soit P une partie d'un espace métrique X. Désignons par

$$\omega_f(P, \delta) = \sup_{\substack{\rho(x', x'') < \delta \\ x' \in P,\ x'' \in P}} | f(x') - f(x'') |$$

l'oscillation d'une fonction $f(x)$ sur l'ensemble $P \subset X$.

b. Soient $P \subset X$ un ensemble élémentaire de mesure mP (*3.18c*) et $\Pi = \{P = \bigcup_{i=1}^{p} A_i\}$ sa partition en cellules disjointes. Nous désignons la somme intégrale sur l'ensemble P, i.e. la somme de la forme $\sum_{i=1}^{p} f(\xi_i)\, mA_i$, par $S_{\Pi}(f, P)$ en réservant le symbole $S_{\Pi}(f)$ au cas $P = X$. Soit $\delta = \max \operatorname{diam} A_i$.

Lemme. *Pour toute partition suivante* Π' *de l'ensemble mentionné* P, *on a*:

$$| S_{\Pi}(f, P) - S_{\Pi'}(f, P) | \leqslant \omega_f(P, \delta)\, mP. \qquad (1)$$

Démonstration. Soit

$$\Pi' = \{P = A_{11} \cup \ldots \cup A_{1r_1} \cup \ldots \cup A_{p1} \cup \ldots \cup A_{pr_p}\}$$

où $A_i = \bigcup_{j=1}^{r_i} A_{ij}$. La partie de la somme intégrale $S_{\Pi'}(f, P)$ qui correspond à la cellule A_i, a la forme:

$$\sum_{j=1}^{r_i} f(\xi_{ij})\, mA_{ij},\ \xi_{ij} \in A_{ij}.$$

Le terme correspondant $f(\xi_i)\, mA_i$ de la somme intégrale $S_\Pi(f, P)$ $(\xi_i \in A_i)$ vérifie l'inégalité

$$\left| f(\xi_i)\, mA_i - \sum_{j=1}^{r_i} f(\xi_{ij})\, mA_{ij} \right| =$$

$$= \left| \sum_{j=1}^{r_i} [f(\xi_i) - f(\xi_{ij})]\, mA_{ij} \right| \leqslant \omega_f(P, \delta)\, mA_i,$$

d'où (1).

Le théorème suivant est fondamental. Les définitions et les résultats *a*, *b* y sont utilisés en supposant que $P = X$. Le cas $P \subset X$, $P \neq X$ sera utile dans la suite (*3.24*).

c. Théorème. *Toute fonction $f(x)$ uniformément continue sur un espace chargé X est intégrable.*

Démonstration. Pour $\varepsilon > 0$ donné, on peut trouver $\delta > 0$ tel que $\omega_f(X, \delta) < \varepsilon/(2mX)$. Par conséquent, quelles que soient une partition Π avec $d(\Pi) < \delta$ et une partition suivante Π', l'inégalité (1) a pour conséquence

$$|S_{\Pi'}(f) - S_\Pi(f)| \leqslant \omega_f(X, \delta)\, mX \leqslant \varepsilon,$$

et il ne reste qu'à appliquer *3.21b*.

d. Conséquence. *Toute fonction continue sur un compact chargé X est intégrable.*

En effet, toute fonction continue $f(x)$ sur un compact X est uniformément continue (*U5.17*), et le résultat cherché découle de *c*.

3.23. Nous aurons besoin des intégrales de certaines fonctions discontinues dont l'ensemble des points de discontinuité est relativement petit. Pour décrire de tels ensembles, formulons les définitions suivantes:

a. Soient A un sous-ensemble d'un espace métrique X et $x \in X$ un point quelconque. Le nombre

$$\rho(x, A) = \inf_{y \in A} \rho(x, y)$$

s'appelle *distance du point x à l'ensemble A*.

b. L'ensemble

$$U_\delta(A) = \{x \in X : \rho(x, A) < \delta\}$$

s'appelle ρ-*voisinage d'un ensemble* $A \subset X$.

c. Un ensemble $A \subset X$ est dit contenu *strictement à l'intérieur d'un ensemble* $B \subset X$, si l'on a l'inclusion $U_\delta(A) \subset B$ avec un certain $\delta > 0$.

d. Soient A et B deux sous-ensembles d'un espace métrique X. Posons

$$d(A, B) = \inf_{\substack{x \in A \\ y \in B}} \rho(x, y).$$

Il est clair que $d(A, B) = 0$ si, et seulement si, les ensembles A et B possèdent au moins un point adhérent commun ; sinon $d(A, B) > 0$.

Il est évident que, si A est strictement à l'intérieur de B, alors $d(A, X - B) > 0$; réciproquement, si $d(A, X - B) > 0$, alors A est situé strictement à l'intérieur de B.

e. Un ensemble Z dans un espace chargé X est dit *négligeable* si, pour tout $\varepsilon > 0$, on peut l'inclure strictement dans l'intérieur d'une réunion finie de cellules $A_1, \ldots, A_p$ avec $\sum_{j=1}^{p} mA_j < \varepsilon$.

f. *Toute réunion finie* Z *d'ensembles négligeables* $Z_1, \ldots, Z_n$ *est encore un ensemble négligeable* : en effet, si Z_k est contenu strictement à l'intérieur de la réunion des cellules $A_1^{(k)}, \ldots, A_{s_k}^{(k)}$, $k = 1, \ldots, n$, alors Z est à l'intérieur de la réunion des cellules $A_1^{(1)}, \ldots, A_{s_1}^{(1)}, \ldots, A_1^{(n)}, \ldots, A_{s_n}^{(n)}$; si, de plus, les cellules $A_j^{(k)}$ sont choisies de façon que $\sum_{j=1}^{s_k} mA_j^{(k)} < \frac{\varepsilon}{n}$, alors $\sum_{k=1}^{n} \sum_{j=1}^{s_k} mA_j^{(k)} < \varepsilon$, ce qu'il fallait.

3.24.a. Dans la théorie de l'intégration des fonctions d'une variable x, $a \leqslant x \leqslant b$, on a établi le théorème sur l'intégrabilité des fonctions continues par morceaux (*U9.16*).

Voici l'analogue de ce théorème pour le cas de fonctions définies sur un espace chargé :

Théorème. *Soient* X *un espace chargé et* $Z \subset X$ *un ensemble négligeable. Toute fonction bornée* $f(x)$ *qui est uniformément continue en dehors de chaque voisinage de l'ensemble* Z *est intégrable sur* X.

Démonstration. Montrons que, pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que

$$| S_\Pi(f) - S_{\Pi'}(f) | \leqslant \varepsilon,$$

quelles que soient une partition Π avec $d(\Pi) < \delta$ et une partition Π' suivante ; alors la conclusion du théorème découlera de *3.21b*.

Soit $M = \sup |f(x)|$. Par hypothèse, pour tout $\varepsilon > 0$ il existe des cellules $A_1, \ldots, A_p$ avec $\sum_{j=1}^{p} mA_j < \varepsilon/(4M)$ dont la réunion contient l'ensemble Z strictement à son intérieur. Posons $B = X - \bigcup_{j=1}^{p} A_j$. D'après *3.23d*, $d(Z, B) \equiv 2\rho > 0$. La fonction $f(x)$ est uniformément continue en dehors du ρ-voisinage de l'ensemble Z; il existe donc un $\tau > 0$ tel que les relations $\rho(x', x'') < 2\tau$, $x' \in B$, $x'' \in B$ impliquent $|f(x') - f(x'')| < \varepsilon/(2mX)$. Considérons n'importe quelle partition Π de l'espace X avec $d(\Pi) < \delta = \min(\tau, \rho)$ et n'importe quelle partition suivante Π'.

Partageons en deux classes l'ensemble de toutes les cellules $C_1, \ldots, C_n$ de la partition Π. La première sera formée des cellules qui se trouvent entièrement dans la réunion des cellules $A_1, \ldots, A_p$, la seconde des cellules qui restent et qui ont donc les points communs avec l'ensemble B.

Les cellules de la seconde classe sont entièrement en dehors du ρ-voisinage de l'ensemble Z puisque leurs diamètres sont inférieurs à $\delta \leqslant \rho$ et elles contiennent des points dont la distance de Z dépasse 2ρ. Soient P la réunion des cellules de la première classe et Q, celle de la seconde. Partageons en deux classes respectives les cellules D_i de la partition suivante Π'; la première classe sera formée des cellules contenues dans l'ensemble P, la seconde des cellules contenues dans Q. Evaluons la différence des sommes intégrales $S_\Pi(f)$ et $S_{\Pi'}(f)$. Nous avons:

$$|S_\Pi(f) - S_{\Pi'}(f)| \leqslant |S_\Pi(f, P)| + |S_{\Pi'}(f, P)| + |S_\Pi(f, Q) - S_{\Pi'}(f, Q)|.$$

Estimons deux premiers termes en appliquant le lemme *3.18b*:

$$|S_\Pi(f, P)| \leqslant M \sum_{C_i \subset P} mC_i \leqslant M \sum_{j=1}^{p} mA_j \leqslant M \cdot \frac{\varepsilon}{4M} = \frac{\varepsilon}{4}, \qquad (1)$$

$$|S_{\Pi'}(f, P)| \leqslant M \sum_{D_i \subset P} mD_i \leqslant M \sum_{j=1}^{p} mA_i \leqslant M \cdot \frac{\varepsilon}{4M} = \frac{\varepsilon}{4}, \qquad (2)$$

vu le lemme *3.22b*, nous avons

$$|S_\Pi(f, Q) - S_{\Pi'}(f, Q)| \leqslant \omega_f(Q, \delta)\, mX \leqslant \frac{\varepsilon}{2mX} mX = \frac{\varepsilon}{2}. \qquad (3)$$

Il résulte de (1), (2) et (3) que

$$|S_\Pi(f) - S_{\Pi'}(f)| \leqslant \varepsilon,$$

ce qu'il fallait.

b. Conséquence. *Une fonction bornée $f(x)$, qui est nulle partout sur un espace chargé X à l'exception d'un ensemble négligeable Z est intégrable sur X, et son intégrale est nulle.*

En effet, la fonction $f(x)$ est nulle, donc uniformément continue en dehors de chaque voisinage de l'ensemble Z; son intégrabilité résulte du théorème *a*. Ensuite, avec les désignations de ce théorème, nous avons

$$S_\Pi(f) = S_\Pi(f, P) + S_\Pi(f, Q) = S_\Pi(f, P),$$

car la fonction $f(x)$ est nulle sur l'ensemble $Q \subset X - U_\rho(Z)$. Vu (1), on en tire

$$|S_\Pi(f)| = |S_\Pi(f, P)| \leqslant \frac{\varepsilon}{4},$$

donc,

$$\int_X f(x)\,dx = \lim_{d(\Pi)\to 0} S_\Pi(f) = 0.$$

c. Conséquence. *Soient Π_1, Π_2, . . . une suite de partitions d'un espace chargé X avec $d(\Pi_n) \to 0$, et $m_n(Z)$ la somme des mesures des cellules de la partition Π_n qui contiennent des points d'un ensemble négligeable fixe Z. Alors $m_n(Z) \to 0$.*

En effet, $m_n(Z)$ est la somme intégrale de la fonction $f(x)$ qui vaut 1 sur l'ensemble Z et 0 ailleurs, si seulement on choisit pour point marqué ξ_i un point de $C_i \cap Z$. Par conséquent, la proposition *c* résulte de *b*.

3.25. Dans le cas où l'espace chargé X est *compact*, l'hypothèse du théorème *3.24* peut être simplifiée, puisque l'on peut ne plus s'intéresser à la continuité uniforme de la fonction en dehors des voisinages de l'ensemble négligeable donné. Etablissons d'abord un lemme:

a. Lemme. *Une fonction $f(x)$ bornée sur un compact chargé X et dont les points de discontinuité forment un ensemble négligeable Z est uniformément continue en dehors de tout voisinage de l'ensemble Z.*

Démonstration. Admettons le contraire: pour certains $\delta > 0$ et $\varepsilon > 0$, il existe une suite de points x'_n, x''_n $(n = 1, 2, \ldots)$ telle que $\rho(x'_n, Z) \geqslant \delta$, $\rho(x''_n, Z) \geqslant \delta$, $\rho(x'_n, x''_n) < \frac{1}{n}$, $|f(x'_n) - f(x''_n)| \geqslant \varepsilon$. Comme X est compact, on peut dire, en passant éventuellement aux sous-suites, que les suites x'_n et x''_n ont une limite commune z. Le point z est donc un point de discontinuité de la fonction $f(x)$ (cf. *U5.17b*) et, comme tel, appartient à Z. Or, il résulte des inégalités $\rho(x'_n, Z) \geqslant \delta$ que $\rho(z, Z) \geqslant \delta$; ceci contredit le fait que Z est l'ensemble de *tous* les points de discontinuité de la fonction $f(x)$. Le lemme est démontré.

b. Théorème. *Si l'ensemble Z des points de discontinuité d'une fonction $f(x)$ bornée sur un compact chargé X est négligeable, alors la fonction $f(x)$ est intégrable sur X.*

Démonstration. D'après le lemme *a*, la fonction $f(x)$ est uniformément continue en dehors de chaque voisinage de l'ensemble Z ; il reste à appliquer *3.24a*.

§ 3. 3. Ensembles jordaniens

3.31.a. Un sous-ensemble G d'un espace chargé avec $\mathfrak{A}$ pour charge s'appelle *ensemble jordanien* (plus précisément, jordanien par rapport à la charge $\mathfrak{A}$), si sa *frontière*, i.e. l'ensemble des points communs à la fermeture de G et à celle de $X - G$, est un ensemble négligeable. Un ensemble jordanien fermé A dont l'intérieur est partout dense dans A est également appelé *corps jordanien*.

La *fonction caractéristique* $\chi_G(x)$ d'un ensemble G, i.e . la fonction qui vaut 1 pour $x \in G$ et 0 pour $x \notin G$, a pour points de discontinuité ceux de la frontière de l'ensemble G. Par conséquent, du moment que G est un ensemble jordanien sur un compact chargé X, la fonction $\chi_G(x)$ est intégrable (*3.25b*). L'intégrale de la fonction $\chi_G(x)$ s'appelle *volume de l'ensemble* G et est désignée par $|G|$ (ou, si besoin est, par $|G|_{\mathfrak{A}}$).

b. Les cellules dans un compact chargé X ne sont pas forcément ensembles jordaniens (voir l'exercice 7). Mais, *dès qu'une cellule* A *est un ensemble jordanien, son volume* $|A|$ *est égal à la mesure initiale* mA *de la cellule.* En effet, considérons une somme intégrale pour la fonction $\chi_A(x)$ construite d'après une partition quelconque Π. Si les points marqués ξ_i des cellules de Π qui ont des points communs avec A sont choisis dans A, la valeur de cette somme s'avère égale à la somme des mesures des cellules en question, donc au moins égale à la mesure de la cellule A. Et si les points ξ_j des cellules de Π qui ont des points communs avec $X - A$ sont choisis dans $X - A$, la valeur de la somme intégrale s'avère égale à la somme des mesures des cellules de Π qui appartiennent entièrement à A, donc au plus égale à la mesure de A. La cellule A étant jordanienne par hypothèse, son volume, i.e. l'intégrale de sa fonction caractéristique, est égal à la mesure de A en tant que limite de nombres au moins égaux à la mesure de A et de nombres au plus égaux à cette mesure, ce qu'il fallait démontrer.

c. Un espace chargé dans lequel toutes les cellules sont jordaniennes sera appelé *espace normalement chargé*, et la charge correspondante *charge normale*. Nous avons montré dans *b* que le volume $|A|$ de chaque cellule A dans un espace normalement chargé est égale à la mesure mA. A côté du symbole $|G|$, on utilise le symbole mG pour désigner le volume d'un ensemble jordanien G dans un espace normalement chargé.

En appliquant le raisonnement *b* à n'importe quel ensemble jordanien G nous voyons que *le volume de tout ensemble jordanien* G *dans un espace normalement chargé est au moins égal à la somme*

$\mu_{\Pi_1}(G)$ *des volumes des cellules contenues dans* G *de n'importe quelle partition* Π_1 *et au plus égal à la somme* $\mu^{\Pi_2}(G)$ *des volumes des cellules ayant des points communs avec* G *de n'importe quelle partition* Π_2. Par conséquent, en passant aux bornes exactes nous avons

$$\sup_{\Pi} \mu_{\Pi}(G) \leqslant |G| \leqslant \inf_{\Pi} \mu^{\Pi}(G), \tag{1}$$

et, l'ensemble G étant jordanien, donc la fonction $\chi_G(x)$ étant intégrable, on peut remplacer (1) par les égalités :

$$\sup_{\Pi} \mu_{\Pi}(G) = |G| = \inf_{\Pi} \mu^{\Pi}(G). \tag{2}$$

En particulier, *pour tout ensemble jordanien* G *et pour tout* $\varepsilon > 0$ ($\varepsilon < |G|$), *on peut indiquer deux ensembles élémentaires* P *et* Q (réunions de cellules disjointes) *tels que* $P \subset G \subset Q$ *et que*

$$|Q| \leqslant |G| + \varepsilon, \quad |P| \geqslant |G| - \varepsilon.$$

d. Lemme. *Dans un espace normalement chargé, pour toute cellule* A *et pour tout* $\varepsilon > 0$, *on peut indiquer un ensemble élémentaire* P *qui contient la cellule* A *strictement dans son intérieur et tel que* $mP < mA + \varepsilon$.

Démonstration. Soit Γ la frontière de la cellule A ; alors Γ est un ensemble négligeable. Par définition, il se trouve strictement à l'intérieur d'une réunion finie de cellules $A_1, \ldots, A_p$ avec la somme des mesures inférieure à ε. L'ensemble élémentaire $P = A \cup A_1 \cup \ldots \cup A_p$ contient déjà la cellule A strictement à son intérieur. Sa mesure, d'après *3.18d*, ne dépasse pas $mA + mA_1 + \ldots + mA_p \leqslant mA + \varepsilon$, ce qu'il fallait.

e. On peut simplifier la définition *3.23e* d'un ensemble négligeable dans un espace normalement chargé ; *dans un tel espace* X, *un ensemble* $Z \subset X$ *est négligeable si, pour tout* $\varepsilon > 0$, *il existe une réunion finie de cellules* $A_1, \ldots, A_p$ *qui recouvre* Z (*sans exiger que* Z *soit strictement à son intérieur*) *et tel que* $\sum_{j=1}^{p} mA_j < \varepsilon$.

Pour le prouver, il suffit de remarquer que, une fois le recouvrement mentionné existe, il en existe un autre que l'on obtient d'après *d* : on remplace toute cellule A_j, $j = 1, \ldots, p$, par un ensemble élémentaire P_j, $mP_j \leqslant mA_j + \varepsilon/p$ qui contient A_j strictement à son intérieur ; la réunion des ensembles P_j est un ensemble élémentaire P qui contient strictement à son intérieur toutes les cellules A_j et, en particulier, l'ensemble Z et dont la mesure ne dépasse pas ε.

Qui plus est, *un ensemble* $Z \subset X$ *est négligeable si, pour tout* $\varepsilon > 0$, *il existe un ensemble jordanien* $G_\varepsilon \supset Z$ *avec* $|G_\varepsilon| < \varepsilon$. En effet, une fois un ensemble $G_\varepsilon \supset Z$ avec $|G_\varepsilon| < \varepsilon$ trouvé d'après $\varepsilon > 0$ donné, on peut, en vertu de *c*, trouver un ensemble élé-

mentaire $Q_\varepsilon \supset G$ avec $|Q_\varepsilon| < 2\varepsilon$; $\varepsilon > 0$ étant arbitraire, l'ensemble Z est négligeable d'après ce qui précède.

3.32. Théorème. **a.** *L'intersection $G_1 \cap G_2$ de deux ensembles jordaniens G_1 et G_2 en est encore un.*

b. *La réunion $G_1 \cap G_2$ de deux ensembles jordaniens G_1 et G_2 en est encore un; si G_1 et G_2 sont disjoints, alors*

$$|G_1 \cap G_2| = |G_1| + |G_2|. \tag{1}$$

c. *Le complémentaire $E - G$ d'un ensemble jordanien G par rapport à un ensemble jordanien $E \supset G$ est encore un ensemble jordanien, et on a*

$$|E - G| = |E| - |G|. \tag{2}$$

Démonstration. **a.** La frontière de $G_1 \cap G_2$ ne contient aucun point qui soit intérieur à la fois ou bien à G_1 et G_2 ou bien à leurs complémentaires. La frontière de $G_1 \cap G_2$ est donc incluse à la réunion des frontières des ensembles G_1 et G_2 et représente un ensemble négligeable en vertu de *3.23f*. Par conséquent, l'ensemble $G_1 \cap G_2$ est jordanien.

b. Pour la même raison, la frontière de $G_1 \cup G_2$ est un ensemble négligeable, étant incluse dans la réunion des frontières des ensembles G_1 et G_2; par conséquent, l'ensemble $G_1 \cup G_2$ est jordanien. Si les ensembles G_1 et G_2 sont disjoints, alors $\chi_{G_1}(x) + \chi_{G_2}(x) = \chi_{G_1 \cup G_2}(x)$ et, d'après *3.14c*, on a

$$|G_1 \cup G_2| = \int_X \chi_{G_1 \cup G_2}(x)\,dx = \int_X \chi_{G_1}(x)\,dx + \int_X \chi_{G_2}(x)\,dx = |G_1| + |G_2|.$$

c. De même que plus haut, la frontière de $E - G$ est incluse dans la réunion des frontières de E et G, de sorte que l'ensemble $E - G$ est jordanien. Vu *b*, on a $|E - G| + |G| = |E|$, d'où (2).

3.33. Intégrale sur un ensemble jordanien. **a.** Soient $f(x)$ une fonction bornée sur un compact chargé X, $|f(x)| \leqslant M$, et $G \subset X$ un ensemble jordanien avec Γ pour frontière. Définissons l'intégrale de la fonction $f(x)$ sur l'ensemble G par la formule:

$$\int_G f(x)\,dx = \int_X f(x)\,\chi_G(x)\,dx. \tag{1}$$

Si la fonction $f(x)$ est continue sur G, à l'exception d'un ensemble négligeable Z, alors la fonction $f(x)\,\chi_G(x)$ est continue partout sur l'espace X, à l'exception de l'ensemble négligeable $Z \cup \Gamma$. L'intégrale (1) existe donc d'après le théorème *3.25b*.

En particulier (ce qu'on peut d'ailleurs voir directement), toutes les règles *3.14a-f* sont valables pour l'intégrale sur un ensemble jordanien; il faut seulement remplacer le nombre mX qui figure dans les estimations par $|G|$. On peut y ajouter la proposition suivante.

b. *Si une fonction $f(x)$ est intégrable sur chacun de deux ensembles jordaniens disjoints G_1 et G_2 d'un compact chargé X, alors elle est intégrable sur leur réunion, et l'on a*

$$\int\limits_{G_1 \cup G_2} f(x)\,dx = \int\limits_{G_1} f(x)\,dx + \int\limits_{G_2} f(x)\,dx.$$

La démonstration résulte immédiatement de la décomposition

$$\chi_{G_1 \cup G_2}(x) = \chi_{G_1}(x) + \chi_{G_2}(x).$$

Les ensembles G_1 et G_2 peuvent même avoir une partie commune à condition qu'elle soit un ensemble négligeable, en particulier, *ils peuvent avoir des points communs sur leurs frontières* (mais non à l'intérieur). Il en résulte que le *volume de l'ensemble jordanien $G_1 \cup G_2$ est égal à la somme des volumes de G_1 et G_2, si G_1 et G_2 n'ont aucun point intérieur commun.*

Il va de soi que les propositions formulées restent valables pour n'importe quel nombre (fini) d'ensembles $G_1, \ldots, G_k$, pourvu que ces derniers soient deux à deux sans points intérieurs communs.

c. Dans un espace normalement chargé (*3.31c*), l'intégrale d'une fonction $f(x)$ sur un ensemble G peut être définie indépendamment de l'intégrale sur tout l'espace. Appelons *partition jordanienne de l'ensemble G* une collection $E_1, \ldots, E_s$ d'ensembles jordaniens deux à deux sans points intérieurs communs qui vérifie la condition $\bigcup_{j=1}^{s} E_j = G$. Pour toute partition jordanienne Π, composons la somme intégrale

$$S_\Pi(f, G) = \sum f(\xi_j)\,|E_j|, \tag{2}$$

où ξ_j est un point de E_j. Soit $d(\Pi) = \max \operatorname{diam} E_j$.

Théorème. *Dans l'hypothèse formulée relativement à la fonction, l'intégrale* (1) *est la limite des sommes intégrales* (2) *suivant n'importe quelle suite $\Pi_1, \Pi_2, \ldots$ de partitions avec $d(\Pi_n) \to 0$.*

Démonstration. Chaque somme intégrale (2) est l'intégrale sur X de la fonction $f_\Pi(x)$ qui vaut $f(\xi_j)$ pour $x \in E_j$ et 0 en dehors de l'ensemble G. Cette fonction est intégrable, car elle est continue en dehors de Z et des frontières de tous les ensembles jordaniens E_j, la réunion de tous ces ensembles étant un ensemble négligeable Z_1 (*3.25b*). Soit $A_1, \ldots, A_p$ un système de cellules deux à deux disjointes qui contient à son intérieur l'ensemble Z_1, la somme des mesures des cellules étant inférieure à $\varepsilon/(4M)$, et soit $P = \bigcup_{i=1}^{p} A_i$. Comme X est un espace normalement chargé, P est un ensemble jordanien (*3.32b*), et $|P| = mP < \varepsilon/(4M)$. La fonction $f(x)$ est uniformément continue en dehors de tout voisinage de l'ensemble

Z_1 (*3.25a*); pour $\varepsilon > 0$, il existe $\delta > 0$ tel que:

$$\max_{x \in X-P} |f(x) - f_\Pi(x)| \leqslant \frac{\varepsilon}{2mX}$$

pour $d(\Pi) < \delta$. Par conséquent, en vertu de *b*, on a

$$\left| \int_G f(x)\,dx - S_\Pi(f, G) \right| = \left| \int_X f(x)\,\chi_G(x)\,dx - \int_X f_\Pi(x)\,\chi_G(x)\,dx \right| \leqslant$$

$$\leqslant \left| \int_P [f(x) - f_\Pi(x)]\,dx \right| + \int_{X-P} |f(x) - f_\Pi(x)|\,\chi_G(x)\,dx \leqslant$$

$$\leqslant \int_P |f(x)|\,dx + \int_P |f_\Pi(x)|\,dx + \int_{X-P} |f(x) - f_\Pi(x)|\,dx \leqslant$$

$$\leqslant \frac{\varepsilon}{4M} M + \frac{\varepsilon}{4M} M + \frac{\varepsilon}{2mX} mX = \varepsilon,$$

d'où notre affirmation.

Pour une fonction $f(x)$ continue sur un ensemble jordanien G, nous avons immédiatement:

$$\left| \int_G f(x)\,dx - S_\Pi(f, G) \right| = \left| \int_G [f(x) - f_\Pi(x)]\,dx \right| \leqslant \omega_f(G, \delta)\,|G|. \quad (3)$$

d. La définition *c* peut être interprétée comme suit. Un ensemble jordanien G dans un espace normalement chargé X est lui-même un espace chargé dont les cellules sont ses parties jordaniennes (ou même, seules les intersections des cellules $A \subset X$ avec l'ensemble G), la mesure de chaque cellule coïncidant avec son volume. Tous les axiomes *3.12a-e* sont valables d'après *3.32*. L'intégrale d'une fonction $f(x)$ sur G en tant qu'espace chargé est bien l'intégrale de la fonction $f(x)\,\chi_G(x)$ sur l'ensemble jordanien G au sens de la définition *a*. Dans le cas où la fonction $f(x)$ n'est définie que sur G, on peut la considérer sur tout X en posant, par exemple, $f(x) = 0$ pour $x \notin G$.

e. Enfin, on peut trouver l'intégrale d'une fonction $f(x)$ sur un ensemble jordanien G de la manière suivante. *Soit* $\Pi_1, \Pi_2, \ldots$ *une quelconque suite de partitions d'un compact* X, *avec* $d(\Pi_n) \to 0$; *désignons par* $C_1^{(n)}, \ldots, C_{k_n}^{(n)}$ *les cellules de la partition* Π_n *qui se trouvent inclues dans l'ensemble* G, *et soit* $\xi_j^{(n)} \in C_j^{(n)}$ *un point arbitrairement choisi. Alors les sommes*

$$\sum_{j=1}^{k_n} f(\xi_j^{(n)})\,mC_j^{(n)} \quad (4)$$

ont l'intégrale (1) *pour limite lorsque* $n \to \infty$.

En effet, ces sommes représentent les sommes intégrales de la fonction $f(x)\,\chi_G(x)$ pour la partition Π_n, du moment que l'on

marque dans les cellules $C_j^{(n)}$ les points $\xi_j^{(n)}$, et dans les autres cellules, des points n'appartenant pas à G. Comme la fonction $f(x)\,\chi_G(x)$ est intégrable sur X (a), les sommes (4) tendent vers son intégrale, c'est-à-dire, d'après a, vers l'intégrale de la fonction $f(x)$ sur l'ensemble G.

f. Charges équivalentes. Considérons, sur un même espace métrique X, deux charges, i.e. deux demi-anneaux $\mathfrak{A}$ et $\mathfrak{B}$ formés respectivement des cellules $A \in \mathfrak{A}$ et $B \in \mathfrak{B}$ avec les mesures respectives mA et μB; les axiomes *3.12a-e* sont donc remplis dans les deux cas. Les charges $\mathfrak{A}$ et $\mathfrak{B}$ sont dites *équivalentes* si toute fonction $f(x)$ intégrable sur l'espace X avec la charge $\mathfrak{A}$ est intégrable sur X avec la charge $\mathfrak{B}$ et vice versa, et si, de plus,

$$\int_{X,\,\mathfrak{A}} f(x)\,dx = \int_{X,\,\mathfrak{B}} f(x)\,dx. \tag{5}$$

Signalons un critère d'équivalence de deux charges normales.

Théorème. *Si chaque cellule $A \in \mathfrak{A}$ est un ensemble jordanien par rapport à la charge normale $\mathfrak{B}$, la relation $mA = |A|_{\mathfrak{B}}$ ayant lieu, et si chaque cellule $B \in \mathfrak{B}$ est un ensemble jordanien par rapport à la charge normale $\mathfrak{A}$, la relation $\mu B = |B|_{\mathfrak{A}}$ ayant lieu, alors les charges $\mathfrak{A}$ et $\mathfrak{B}$ sont équivalentes.*

Démonstration. Pour la raison de symétrie, il suffit d'envisager le cas où la fonction $f(x)$ est intégrable sur l'espace X avec $\mathfrak{A}$ pour charge et d'en déduire, dans l'hypothèse du théorème, son intégrabilité sur X avec la charge $\mathfrak{B}$ ainsi que l'égalité (5). Soient $f(x)$ une fonction vérifiant la condition imposée et $\sum f(\xi_j)\,\mu B_j$ ($\xi_j \in B_j$) sa somme intégrale par rapport à la charge $\mathfrak{B}$. On peut mettre cette somme sous la forme $\sum f(\xi_j)\,|B_j|_{\mathfrak{A}}$. En morcellant la partition la dernière somme tend, conformément à c, vers une limite qui est égale à l'intégrale de $f(x)$ pour la charge $\mathfrak{A}$; il s'avère donc que la fonction $f(x)$ est aussi bien intégrable par rapport à la charge $\mathfrak{B}$ et que les deux intégrales se confondent, ce qu'il fallait.

g. Exemple. Dans *3.15b*, on a muni un pavé n-dimensionnel X d'une charge à l'aide d'un système de sous-pavés

$$A = \{x \in X\colon \alpha_1 \leqslant x_1 \leqslant \beta_1,\ \ldots,\ \alpha_n \leqslant x_n \leqslant \beta_n\} \tag{6}$$

(où il est permis de remplacer n'importe quels des symboles $\leqslant$ par $<$), la mesure de chaque sous-pavé A étant égale à son volume vA. Nous choisissons ici pour pavé X le cube n-dimensionnel

$$X = \{x \in R_n\colon -1 \leqslant x_1 \leqslant 1,\ \ldots,\ -1 \leqslant x_n \leqslant 1\},$$

et pour cellules, les sous-pavés de la forme spéciale suivante :

$$B = \left\{ x \in X : \frac{p_1}{2^q} \leqslant x_1 \leqslant \frac{s_1}{2^q}, \ldots, \frac{p_n}{2^q} \leqslant x_n \leqslant \frac{s_n}{2^q} \right\},$$
$$q = 0, 1, 2, \ldots; \quad p_1, s_1, \ldots, p_n, s_n = 0, \pm 1, \pm 2, \ldots,$$
$$p_j \leqslant s_j \leqslant p_j + 1 \quad (j = 1, \ldots, n)$$

(où l'on peut remplacer $\leqslant$ par $<$). Le système de tous les pavés B sera désigné par $\mathfrak{B}$. Les pavés B représentent soit les cubes d'arête $1/2^q$, soit les parties de la frontière de tels cubes. Si l'on fait abstraction des frontières, alors deux cubes du système $\mathfrak{B}$ ou bien sont disjoints, ou bien l'un d'eux contient l'autre. Dans ce dernier cas, le grand cube peut être obtenu en complétant le petit par des cubes de mêmes dimensions (que le petit) du système $\mathfrak{B}$. Il en résulte que le système $\mathfrak{B}$ est un demi-anneau. Comme avant, nous posons la mesure d'une cellule B égale à son volume vB. Ainsi, le système $\mathfrak{B}$ détermine une charge ; nous tenons à montrer que la charge $\mathfrak{B}$ soit équivalente à la charge initiale que nous désignons par $\mathfrak{A}$.

Il suffit de prouver que tout sous-pavé A (6) est un ensemble jordanien dans la charge $\mathfrak{B}$ et que $| A |_{\mathfrak{B}} = vA$. Le premier résulte du fait que la frontière de chaque pavé A, pour tout $\varepsilon > 0$, peut évidemment être recouverte par un nombre fini de cubes du système $\mathfrak{B}$ de volume total $< \varepsilon$. Ensuite, étant donné un pavé A, on peut construire deux pavés P et Q avec les cubes du système $\mathfrak{B}$, de sorte que $P \subset A \subset Q$ et que la différence des volumes des pavés P, A, Q soit inférieure à ε. Donc, en vertu de *3.32c*, on a

$$| A |_{\mathfrak{B}} = \sup | P |_{\mathfrak{B}} = \sup vP = vA,$$

ce qu'il fallait.

Une construction analogue peut être réalisée pour n'importe quel cube $X = \{x \in R_n : | x_1 - a_1 | \leqslant d, \ldots, | x_n - a_n | \leqslant d\}$ avec les cellules de la forme

$$B = \left\{ x \in X : \frac{p_1}{2^q} \leqslant \frac{x_1}{d} \leqslant \frac{s_1}{2^q}, \ldots, \frac{p_n}{2^q} \leqslant \frac{x_n}{d} \leqslant \frac{s_n}{2^q} \right\},$$

où les valeurs admissibles de q, p_1, s_1, ..., p_n, s_n sont les mêmes.

3.34. Intégrale itérée.

a. Soient X et Y deux espaces chargés et $X \times Y = W$ leur produit avec la charge *3.17*.

Etant données deux cellules $A \subset X$, $B \subset Y$ de frontières $\Gamma(A) \subset X$, $\Gamma(B) \subset Y$ respectivement, la frontière de la cellule $A \times B \subset W$ est, évidemment, inclue dans l'ensemble $(\Gamma(A) \times B) \cup (A \times \Gamma(B))$. Si, de plus, $\Gamma(A)$ se trouve strictement à l'intérieur de la réunion des cellules $A_1, \ldots, A_p$, et $\Gamma(B)$, strictement à l'intérieur de la réunion des cellules $B_1, \ldots, B_q$, alors $(\Gamma(A) \times B) \cup (A \times \Gamma(B))$ se trouve strictement à l'intérieur de la

réunion des cellules

$$(A_1 \times B) \cup \ldots \cup (A_p \times B) \cup (A \times B_1) \cup \ldots \cup (A \times B_q)$$

dont la somme des mesures est au plus égale à

$$mA_1 mB + \ldots + mA_p mB + mA mB_1 + \ldots + mA mB_q =$$

$$= \Big(\sum_{i=1}^{p} mA_i\Big) mB + mA \Big(\sum_{j=1}^{q} mB_j\Big).$$

Il en découle aisément que la cellule $A \times B$ est jordanienne du moment que les cellules A et B le soient. Par conséquent, *si X et Y sont des espaces normalement chargés, alors $W = X \times Y$ est un espace normalement chargé.*

b. Soit $f(w) \equiv f(x, y)$ une fonction définie et uniformément continue sur un ensemble $U \subset X \times Y$.

Théorème. *Si l'ensemble U est jordanien et que, pour tout $y \in Y$ fixe, l'ensemble $X_y = \{x \in X : x \times y \in U\}$ (« section horizontale de l'ensemble U ») est aussi jordanien dans X, alors la fonction*

$$F(y) = \int_{X_y} f(x, y)\, dx \tag{1}$$

est intégrable en y, et l'on a

$$\int_{Y} F(y)\, dy = \int_{U} f(w)\, dw. \tag{2}$$

Démonstration. Soient, respectivement, Π_X, Π_Y deux partitions des espaces X et Y en cellules A_i, B_j de diamètres ne dépassant pas δ. Alors les produits $A_i \times B_j$ forment une partition de $X \times Y$ en cellules de diamètres $\leqslant \delta\sqrt{2}$. Les intersections $(A_i \times B_j) \cap U$ forment une partition jordanienne Π de l'ensemble U avec $d(\Pi) \leqslant \delta\sqrt{2}$. Les intersections $A_i \cap X_y$ forment une partition jordanienne Π_y de l'ensemble X_y avec $d(\Pi_y) \leqslant \delta$.

D'après *3.33* (3), on a

$$\Big|\int_{X_y} f(x, y)\, dx - \sum_i f(\xi_i, y)\, m(A_i \cap X_y)\Big| \leqslant$$

$$\leqslant \omega_f(\delta) \mid X_y \mid \leqslant \omega_f(\delta)\, mX.$$

Posons-y $y = \eta_j \in B_j$, multiplions par mB_j et faisons la somme sur j :

$$\Big|\sum_j \int_{X_{\eta_j}} f(x, \eta_j)\, dx \cdot mB_j - \sum_j \sum_i f(\xi_i, \eta_j)\, m(A_i \cap X_{\eta_j})\, mB_j\Big| \leqslant$$

$$\leqslant \omega_f(\delta)\, mX mY.$$

Envisageons, dans la somme double, les termes où $A_i \times B_j$ (cellules auxquelles appartiennent les points (ξ_i, η_j)) contiennent des points

de la frontière de l'ensemble U. Comme U est un ensemble jordanien, la somme des mesures $\nu(\delta)$ des cellules en question tend vers zéro avec δ, en vertu de *3.24c*. Si l'on remplace chacun des termes mentionnés par $f(\xi_i, \eta_j)\, m\,[(A_i \times B_j) \cap U]$, la variation de la somme sera $2M\nu(\delta)$, au plus. Par conséquent,

$$\left| \sum_j \int_{X_{\eta_j}} f(x, \eta_j)\, dx \cdot mB_j - \sum_{i,j} f(\xi_i, \eta_j)\, m\,[(A_i \times B_j) \cap U] \right| \leqslant$$

$$\leqslant \omega_f(\delta)\, mXmY + 2M\nu(\delta). \qquad (3)$$

Le second terme à gauche dans (3) représente une somme intégrale qui a pour limite la quantité

$$\int_U f(w)\, dw.$$

Ainsi, la première somme a une limite pour $\delta \to 0$, et comme elle est une somme intégrale de la fonction $F(y)$ sur l'espace chargé Y, on peut dire que la fonction $F(y)$ est intégrable et que l'égalité (2) est valable, ce qui achève la démonstration.

La relation (2) peut être écrite comme suit:

$$\int_U f(x, y)\, dx\, dy = \int_Y \left\{ \int_{X_y} f(x, y)\, dx \right\} dy. \qquad (4)$$

L'intégrale dans le second membre s'appelle *intégrale itérée*. L'égalité (4) réduit le calcul de l'intégrale sur un ensemble $U \subset W$ à celui d'une intégrale itérée qui contient les intégrations sur les ensembles X_y et Y considérés séparément.

c. En remplaçant X par Y dans l'énoncé du théorème nous obtenons le résultat suivant:

Théorème. *Si chaque ensemble* $Y_x = \{y \in Y : x \times y \in U\}$ (« *section verticale de l'ensemble* U ») *est jordanien* (*dans* Y), *alors la fonction*

$$\Phi(x) = \int_{Y_x} f(x, y)\, dy$$

est intégrable en x, *et l'on a*

$$\int_X \Phi(x)\, dx = \int_U f(w)\, dw.$$

3.35. Intégrales dépendant d'un paramètre. On considère ici les intégrales de la forme

$$\Phi(t) = \int_X f(x, t)\, dx, \qquad (1)$$

où X est un espace chargé, t un paramètre qui varie dans un espace métrique T.

On suppose que la fonction $f(x, t)$ soit intégrable en x pour tout $t \in T$. Nous avons à étudier les propriétés de la fonction $\Phi(t)$; il s'agit avant tout des conditions de sa continuité, puis, dans les suppositions spéciales sur l'espace T, de son intégrabilité et dérivabilité.

Le cas où $X = [a, b]$ a été considéré dans *U9.81-84.*

a. On peut démontrer le théorème sur la continuité de $\Phi(t)$ en suivant la même voie que dans *U9.81-84.*

Théorème. *Si la fonction $f(x, t)$ est uniformément continue sur l'espace métrique $X \times T$, i.e. si la quantité*

$$\omega_f(X \times T, \delta) \equiv \sup_{\substack{\rho(x', x'') \leqslant \delta \\ \rho(t', t'') \leqslant \delta}} |f(x', t') - f(x'', t'')|$$

tend vers zéro pour $\delta \to 0$, alors la fonction $\Phi(t)$ est uniformément continue sur l'espace T.

Démonstration. Pour $\varepsilon > 0$ donné, trouvons $\delta_0 > 0$ tel que $\delta < \delta_0$ impliqie la relation $\omega_f(X \times T, \delta) < \varepsilon/mX$. Soit $\rho(t', t'') \leqslant \delta < \delta_0$; alors

$$|\Phi(t') - \Phi(t'')| = \left| \int_X [f(x, t') - f(x, t'')]\, dx \right| \leqslant$$

$$\leqslant \int_X |f(x, t') - f(x, t'')|\, dx \leqslant \omega_f(X \times T, \delta)\, mX,$$

d'où

$$\omega_\Phi(T, \delta) \equiv \sup_{\rho(t', t'') < \delta} |\Phi(t') - \Phi(t'')| \leqslant \omega_f(X \times T, \delta)\, mX < \varepsilon,$$

ce qui démontre la continuité uniforme de la fonction $\Phi(t)$ sur l'espace T.

b. Supposons maintenant que non seulement X mais aussi T soit un espace chargé avec μ pour mesure. Alors la fonction $\Phi(t)$, en tant que fonction uniformément continue sur un espace chargé, est intégrable en t (*3.22c*).

Théorème. *Dans l'hypothèse formulée,*

$$\int_T \Phi(t)\, dt \equiv \int_T \left\{ \int_X f(x, t)\, dx \right\} dt = \int_X \left\{ \int_T f(x, t)\, dt \right\} dx. \qquad (2)$$

Démonstration. L'espace $X \times T$ est lui aussi chargé, la mesure de chaque sa cellule $C = A \times B$ valant $mA \cdot \mu B$ (*3.17*). La fonction $f(x, t)$ qui est uniformément continue sur $X \times T$ est intégrable sur cet espace (*3.22c*). L'égalité (2) n'exprime alors que la règle de réduction d'une intégrale à des intégrales itérées qu'on a établie dans *3.34.*

c. A présent, supposons que le paramètre t varie dans un espace vectoriel normé T et que la fonction $f(x, t)$ possède, pour tout $x \in X$ et pour $t = t_0$ une dérivée $\frac{\partial f(x, t_0)}{\partial t}$ au sens de *1.23.*

Théorème. *Si la fonction* $\frac{\partial f(x, t)}{\partial t}$ $(X \times T \to X \times L(T))$ *est continue en t pour* $t = t_0$, *uniformément par rapport à* $x \in X$, *alors la fonction* $\Phi(t)$ *est dérivable et l'on a*

$$\frac{d}{dt} \int_X f(x, t)\, dx \Big|_{t=t_0} = \int_X \frac{\partial f(x, t_0)}{\partial t}\, dx. \tag{3}$$

Démonstration. Appliquons le développement qui résulte de *1.42d*:

$$f(x, t) = f(x, t_0) + \frac{\partial f(x, t_0)}{\partial t}(t - t_0) + \varepsilon(x, t)(t - t_0),$$

où

$$\| \varepsilon(x, t) \| \leqslant \varepsilon \equiv \sup_{t_0 \leqslant \theta \leqslant t} \left| \frac{\partial f(x, \theta)}{\partial t} - \frac{\partial f(x, t_0)}{\partial t} \right|.$$

En intégrant terme à terme, nous obtenons:

$$\int_X f(x, t)\, dx = \int_X f(x, t_0)\, dx + (t - t_0) \cdot \int_X \frac{\partial f(x, t_0)}{\partial t}\, dx +$$

$$+ (t - t_0) \cdot \int_X \varepsilon(x, t)\, dx.$$

Comme

$$\int_X | \varepsilon(x, t) |\, dx \leqslant \varepsilon m X,$$

et que $\varepsilon \to 0$ pour $| t - t_0 | \to 0$ en vertu de la continuité uniforme de $\frac{\partial f(x, t)}{\partial t}$ pour $t = t_0$, la fonction $\Phi(t)$ admet une partie linéaire principale par rapport à $t - t_0$ qui est égale à l'intégrale dans le second membre de (3), cette intégrale agissant sur le vecteur $t - t_0$. Cela démontre la proposition.

d. Conséquence. *Dans l'hypothèse c, quelle que soit une direction* τ *dans l'espace T, l'existence et la continuité de la dérivée* $\frac{\partial f(x, t)}{\partial \tau}$ *impliquent la dérivabilité de la fonction* $\Phi(t)$ *suivant la direction* τ *et la relation*

$$\frac{\partial}{\partial \tau} \int_X f(x, t)\, dx = \int_X \frac{\partial f(x, t)}{\partial \tau}\, dx.$$

3.36. Suites en forme de delta. Un point y d'un espace chargé X sera appelé *point jordanien* s'il existe, pour tout $\delta > 0$, un ensemble jordanien U de diamètre $\leqslant \delta$ qui contient le point y strictement à son intérieur. Une suite $D_1(x)$, $D_2(x)$, ... de fonctions intégrables non négatives sera appelée *suite en forme de delta pour le point* y si, pour tout ensemble jordanien U qui contient le point y strictement à son intérieur, sont valables les relations limites

$$1) \lim_{n\to\infty} \int_U D_n(x)\,dx = 1,$$

$$2) \lim_{n\to\infty} \int_{X-U} D_n(x)\,dx = 0.$$

Alors, *étant donnée une fonction* $f(x)$ *continue pour* $x = y$ *et pour laquelle tous les produits* $f(x)\,D_n(x)$ *soient intégrables sur l'espace* X, *a lieu la relation*

$$\lim_{n\to\infty} \int_X D_n(x)\,f(x)\,dx = f(y).$$

La démonstration est analogue à celle du théorème *U12.55b* en y remplaçant la fonction $D_n(x, y)$ par $D_n(x)$, l'intervalle Q, par l'espace chargé X, et le voisinage $U_\delta(y)$ par un ensemble jordanien U qui contient le point y strictement à l'intérieur et dont le diamètre est si petit que la différence de chaque valeur de la fonction $f(x)$ sur U et de $f(y)$ ne dépasse pas un $\varepsilon > 0$ donné.

§ 3. 4. Applications des espaces chargés

3.41.a. Soit X un espace chargé avec un demi-anneau $\mathfrak{A}$ dont les cellules A ont la mesure mA. Nous savons que la mesure est une fonction non négative et additive des cellules. En plus de la mesure, il existe d'autres fonctions additives des cellules prenant éventuellement les valeurs de deux signes.

L'additivité d'une fonction des cellules $\Phi(A)$, de même que celle de la mesure mA, est exprimée par la relation

$$\Phi(A) = \Phi(A_1) + \ldots + \Phi(A_p) \qquad (1)$$

ayant lieu chaque fois qu'une cellule A représente la réunion de cellules disjointes $A_1, \ldots, A_p$. Une fonction $\Phi(A)$ est dite *fortement additive* si l'égalité (1) reste valable pour des cellules $A_1, \ldots, A_p$ dont les intersections sont négligeables.

Nous supposons dans ce qui suit que $\Phi(A) = 0$ pour toute cellule A avec $mA = 0$.

b. En tant qu'exemple considérons la fonction des cellules

$$\Phi(A) = \int_A f(x)\,dx, \tag{2}$$

où $f(x)$ est une fonction continue (uniformément) sur un espace normalement chargé X. Son additivité forte résulte de *3.33b*. Si la fonction $f(x)$ prend les valeurs de deux signes, il en est de même de la fonction $\Phi(A)$. Si $mA = 0$, on a évidemment $\Phi(A) = 0$.

c. Pour deux fonctions (fortement) additives $\Phi_1(A)$ et $\Phi_2(A)$, toute combinaison linéaire $\alpha_1\Phi_1(A) + \alpha_2\Phi_2(A)$ (α_1 et α_2 sont des coefficients réels) est encore une fonction des cellules, évidemment (fortement) additive elle aussi. Ainsi, *les fonctions (fortement) additives des cellules forment un espace vectoriel.*

d. Si $mA > 0$, le quotient $\Phi(A)/mA$ s'appelle *valeur moyenne de la fonction* $\Phi(A)$ sur la cellule A.

3.42.a. Nous dirons qu'une suite $A_1, \ldots, A_s, \ldots$ de cellules *se contracte vers un point* $y \in X$ pour $s \to \infty$ (et nous noterons $A_s \to y$) si le point y se trouve à l'intérieur ou sur la frontière de chaque cellule A_s et si toute boule centrée en y contient toutes ces cellules, à partir d'un numéro quelconque.

b. Soit $A_s \to y$; si la suite des nombres $\Phi(A_s)/mA_s$ ($mA_s > 0$) a une limite qui ne dépend pas du choix de la suite A_s avec $mA_s > 0$ se contractant vers le point y, cette limite s'appelle *densité de la fonction* $\Phi(A)$ *au point* y. La densité d'une fonction des cellules est désignée par la même lettre mais minuscule. Ainsi,

$$\varphi(y) = \lim_{A_s \to y} \frac{\Phi(A_s)}{mA_s}.$$

c. L'inégalité *3.14* (7) appliquée à une fonction (2) des cellules

$$\inf_A f(x) \cdot mA \leqslant \Phi(A) = \int_A f(x)\,dx \leqslant \sup_A f(x) \cdot mA$$

et la continuité de $f(x)$ sur X impliquent que *la fonction* (2) *des cellules possède une densité dont la valeur à tout point* $y \in X$ *est* $f(y)$. De plus, d'après la définition même de la fonction $\Phi(A)$, elle représente l'intégrale sur la cellule A de sa densité. Il se trouve que le dernier fait est d'un caractère général s'il s'agit d'un espace normalement chargé et complet X: nous montrerons dans *3.44* que chaque fonction des cellules $\Phi(A)$ qui soit (fortement) additive et qui possède une densité $\varphi(x)$ (uniformément) continue, peut être restituée d'après sa densité en intégrant sur les cellules.

3.43.a. L e m m e. *Si deux fonctions* $\Phi_1(A)$ *et* $\Phi_2(A)$ *des cellules possèdent, à un point* y, *les densités respectives* $\varphi_1(y)$ *et* $\varphi_2(y)$, *alors, quels que soient* α_1 *et* α_2 *réels, la fonction des cellules* $\Phi(A) = \alpha_1\Phi_1(A) + \alpha_2\Phi_2(A)$ *possède en* y *la densité* $\alpha_1\varphi_1(y) + \alpha_2\varphi_2(y)$.

Demonstration résulte de la relation

$$\frac{\Phi(A)}{mA} = \alpha_1 \frac{\Phi_1(A)}{mA} + \alpha_2 \frac{\Phi_2(A)}{mA}$$

si l'on passe à la limite pour $A \to y$.

b. Lemme. *Etant donnée une partition $A_1, \ldots, A_p$ d'une cellule A avec $mA > 0$ en cellules disjointes* (d'intersections négligeables), *si la valeur moyenne d'une fonction* (fortement) *additive Φ sur toute cellule A_j avec $mA_j > 0$ est en valeur absolue inférieure à une quantité γ, alors la valeur moyenne de la fonction Φ sur la cellule A est elle aussi inférieure à γ en valeur absolue.*

Démonstration. Il suit des relations

$$\frac{|\Phi(A_1)|}{mA_1} \leqslant \gamma, \ \ldots, \ \frac{|\Phi(A_p)|}{mA_p} \leqslant \gamma$$

que

$$|\Phi(A_1)| \leqslant \gamma mA_1, \ \ldots \ |\Phi(A_p)| \leqslant \gamma mA_p;$$

ainsi,

$$|\Phi(A)| = |\Phi(A_1) + \ldots + \Phi(A_p)| \leqslant$$
$$\leqslant \gamma(mA_1 + \ldots + mA_p) = \gamma mA,$$

d'où

$$\frac{|\Phi(A)|}{mA} \leqslant \gamma,$$

ce qu'il fallait démontrer.

c. Lemme. *Si la densité $\varphi(x)$ d'une fonction des cellules $\Phi(A)$, qui est additive sur un espace complet chargé X, est identiquement nulle, alors la fonction $\Phi(A)$ est nulle sur chaque cellule A.*

Le lemme est à fortiori valable pour une fonction fortement additive.

Démonstration. Soit $\Phi(A_1) \neq 0$ pour une cellule $A = A_1$, $mA_1 > 0$. Alors $\gamma = |\Phi(A_1)| / mA_1 > 0$.

Considérons une partition de la cellule A_1 en des cellules plus petites de diamètres inférieurs à 1. D'après le lemme *b*, il en existe une, désignons-la A_2, sur laquelle la valeur moyenne de la fonction Φ soit au moins égale à γ en module. D'une façon analogue, partageons la cellule A_2 en des cellules plus petites de diamètres inférieurs à 1/2 dont trouvons une, désignons-la A_3, sur laquelle la valeur moyenne de la fonction Φ soit toujours au moins égale à γ en module. Continuons ce procédé indéfiniment; nous aboutissons à une suite des cellules dont les diamètres tendent vers 0, sur chacune desquelles la valeur moyenne de la fonction Φ est au moins égale à γ en module. L'espace X étant complet par hypothèse, il existe d'après *U3.74a* un point $y \in X$ tel que chaque boule centrée en ce point contient toutes les cellules $A_1, A_2, \ldots, A_s, \ldots$ à partir d'une certaine, le point y appartenant à la fermeture de toutes ces cellules. Ainsi, la

suite $A_1, \ldots, A_s, \ldots$ se contracte vers le point y. Par hypothèse, nous avons $\lim_{s\to\infty} |\Phi(A_s)| / mA_s = 0$; or, par construction, $|\Phi(A_s)| / mA_s \geqslant \gamma > 0$. La contradiction obtenue montre qu'il n'existe pas de cellule A avec $mA > 0$ et $\Phi(A) \neq 0$. Le lemme est démontré.

3.44. Théorème. *Si une fonction additive* (fortement additive d'autant plus) *des cellules* $\Phi(A)$ *a une densité* $\varphi(x)$ *continue sur* X, *on a*

$$\Phi(A) = \int_A \varphi(x)\, dx \tag{1}$$

pour toute cellule A.

Démonstration. Considérons une autre fonction des cellules

$$\Psi(A) = \int_A \varphi(x)\, dx.$$

Nous avons vu dans *3.41b* et *3.42c* que cette fonction est additive et a pour densité la fonction $\varphi(x)$. La différence $\Phi(A) - \Psi(A)$ est encore une fonction additive des cellules (*3.41c*) avec 0 pour densité (*3.43a*). En vertu du lemme *3.43c*, cette différence est elle-même nulle sur toute cellule A ; ainsi

$$\Phi(A) = \Psi(A) = \int_A \varphi(x)\, dx, \tag{2}$$

ce qu'il fallait démontrer.

3.45. Soit $x = \theta(u)$ une application d'un espace normalement chargé et complet U avec μB pour mesure des cellules dans un espace normalement chargé et complet X avec mA pour mesure des cellules. L'application θ est supposée univoque, continue et *jordanienne*, c'est-à-dire transformant toute cellule B de mesure positive de l'espace U en un ensemble jordanien $\theta(B)$ de mesure positive de l'espace X et toutes deux cellules B_1 et B_2 sans point intérieur commun en deux ensembles jordaniens respectifs $\theta(B_1)$ et $\theta(B_2)$ toujours sans point intérieur commun. Supposons en outre que $\theta(U) = X$. Définissons sur les cellules B de l'espace U la fonction

$$\Phi(B) = m(\theta(B)).$$

(L'ensemble $\theta(B)$ étant jordanien sur X, le nombre $m(\theta(B)) = |\theta(B)|$ est déterminé d'une manière univoque.) La fonction $\Phi(B)$ est additive en vertu de *3.33d*. Supposons qu'elle possède une densité

$$\varphi(u) = \lim_{B\to u} \frac{\Phi(B)}{\mu B} = \lim_{B\to u} \frac{m(\theta(B))}{\mu B}$$

qui soit fonction continue de u. Cette fonction $\varphi(u)$ s'appelle *coefficient de distorsion de la mesure* μ pour l'application θ. A toute fonction $f(x)$ continue sur l'espace X on peut faire correspondre la fonction continue $g(u) = f(\theta(u))$ sur l'espace U. Etant donné le coefficient de distorsion de la mesure $\varphi(u)$ pour l'application θ, on peut lier l'intégrale de $f(x)$ sur l'espace X à l'intégrale d'une fonction sur l'espace U; notamment, *la formule suivante a lieu*:

$$\int_X f(x)\,dx = \int_U g(u)\,\varphi(u)\,du. \tag{1}$$

En abordant la preuve, considérons la fonction additive des cellules $B \subset U$:

$$\Psi(B) = \int_{\theta(B)} f(x)\,dx.$$

Trouvons sa densité. Pour $\mu B > 0$, nous avons

$$\frac{\int_{\theta(B)} f(x)\,dx}{\mu B} = \frac{\int_{\theta(B)} f(x)\,dx}{m(\theta(B))} \cdot \frac{m(\theta(B))}{\mu B}. \tag{2}$$

Supposons que la cellule B se contracte vers le point u. L'application θ étant continue, l'ensemble jordanien $\theta(B)$ se contracte vers le point $x = \theta(u)$. Comme la fonction $f(x)$ est continue, le premier rapport dans le second membre de (2) tend vers la limite $f(x)$. Le deuxième rapport, par hypothèse, a $\varphi(u)$ pour limite. Ainsi, la fonction $\Psi(B)$ possède une densité qui vaut $f(x)\,\varphi(u) = g(u)\,\varphi(u)$ et qui est continue sur U. D'après le théorème *3.44*, pour toute cellule de l'espace U, en particulier sur l'espace U même, on peut écrire la formule exprimant la fonction des cellules par sa densité:

$$\Psi(U) = \int_X f(x)\,dx = \int_U g(u)\,\varphi(u)\,du,$$

ce qu'il fallait.

§ 3. 5. Intégrale de Riemann dans un espace euclidien

3.51.a. Sur un pavé

$$X = \{x \in R_n : a_1 \leqslant x_1 \leqslant b_1, \ldots, a_n \leqslant x_n \leqslant b_n\}$$

nous avons établi une structure d'espace chargé (*3.15b*) en choisissant pour mesure de chaque cellule

$$A = \{x \in X : \alpha_1 \leqslant x_1 \leqslant \beta_1, \ldots, \alpha_n \leqslant x_n \leqslant \beta_n\} \tag{1}$$

son volume euclidien $|A| = \prod_{j=1}^{n} (\beta_j - \alpha_j)$. Cela concerne également les cellules qui s'obtiennent de celles-ci en remplaçant certains des symboles $\leqslant$ par $<$, c'est-à-dire en retranchant de A certaines parties de la frontière.

b. Un ensemble Z est négligeable dans X (*3.24e*) si, pour tout $\varepsilon > 0$, il est inclus dans une réunion finie de cellules (1) disjointes avec la somme des volumes $\leqslant \varepsilon$. La frontière de chaque cellule est un ensemble négligeable, puisque, par exemple, le plan $x_j = \gamma$, pour n'importe quels $\varepsilon > 0$ et $C > 0$, se trouve strictement à l'intérieur de la cellule

$$\{x \in R_n : a_1 \leqslant x_1 \leqslant b_1, \ldots$$
$$\ldots, \gamma - C\varepsilon \leqslant x_j \leqslant \gamma + C\varepsilon, \ldots, a_n \leqslant x_n \leqslant b_n\}$$

dont le volume vaut $(b_1 - a_1) \ldots 2C\varepsilon \ldots (b_n - a_n)$, donc peut être rendu $\leqslant \varepsilon$ par le choix convenable de C. Ainsi, *l'espace X est normalement chargé* (*3.31c*). Il en résulte d'après *3.31e* qu'en définissant un ensemble négligeable Z on peut considérer n'importe quels recouvrements de Z par des cellules de mesure totale $\leqslant \varepsilon$, et non seulement ceux qui contiennent Z strictement à l'intérieur.

c. Montrons que *chaque partie de X définie par une équation de la forme*

$$x_i = \varphi(x'), \quad x' = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$$

avec une fonction continue φ donnée sur un pavé

$$B' = \{x \in R_{n-1} : a_1 \leqslant x_1 \leqslant b_1, \ldots, a_{i-1} \leqslant x_{i-1} \leqslant b_{i-1},$$
$$a_{i+1} \leqslant x_{i+1} \leqslant b_{i+1}, \ldots, a_n \leqslant x_n \leqslant b_n\}$$

ou sur sa partie compacte quelconque K' est un ensemble égligeable dans X.

Fixons un $\varepsilon > 0$ arbitrairement et, compte tenu de la continuité uniforme de la fonction φ sur K', trouvons $\delta > 0$ tel que $|\varphi(x') - \varphi(y')| \leqslant \varepsilon$ du moment que $|x' - y'| \leqslant \delta$, $x' \in K'$, $y' \in K'$. Partageons le pavé B' en des cellules A'_j avec $d(A'_j) \leqslant \delta$.

Dans chaque ensemble non vide $A'_j \cap K'$ choisissons de façon arbitraire un point ξ_j et considérons dans R_n le pavé

$$B_j = \{x \in R_n : \varphi(\xi_j) - \varepsilon \leqslant x_i \leqslant \varphi(\xi_j) + \varepsilon, \ x' \in A'_j\}.$$

Il résulte de la définition du nombre δ que chaque point de notre surface qui se projette dans A'_j appartient au pavé B_j. Par conséquent, cette surface appartient à la réunion de tous les B_j. Or, la somme des mesures de ces pavés ne dépasse pas $2\varepsilon \sum mA'_j \leqslant 2\varepsilon m B'$. Comme ε est arbitraire, la proposition s'avère démontrée.

d. *L'image géométrique dans X qui correspond à la représentation paramétrique*

$$\left.\begin{aligned} x_1 &= f_1(u_1, \ldots, u_k), \\ &\cdots\cdots\cdots\cdots \\ x_n &= f_n(u_1, \ldots, u_k) \end{aligned}\right\} (u_1, \ldots, u_k) = u \in U \subset R_k; \quad k < n, \quad (2)$$

où U est une partie compacte de R_k et les fonctions $f_1(u), \ldots, f_n(u)$ possèdent dans le domaine U leurs dérivées premières continues, est un ensemble négligeable. En effet, d'après le théorème du rang *1.74b*, chaque point $u \in U$ possède un voisinage dans lequel les équations (2) sont équivalentes à une ou plusieurs équations de la forme (1), donc définissent un ensemble négligeable dans R_n. En appliquant le théorème sur le recouvrement fini nous voyons que l'image dans R_n de tout le compact U est un ensemble négligeable. La proposition est démontrée.

e. Il découle de *c* que *chaque ensemble $G \subset X \subset R_n$ défini par des inégalités de la forme*

$$\varphi_i(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \leqslant \\ \leqslant x_i \leqslant \psi_i(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n), \; i = 1, \ldots, n$$

(avec des remplacements éventuels de $\leqslant$ par $<$) *est un ensemble jordanien dans X, donc possède un volume* que nous désignons, comme plus haut, par $|G|$. Ce volume $|G|$ ne dépend pas du choix du pavé $X \subset R_n$ contenant l'ensemble G; pour cette raison nous appellerons la quantité $|G|$ volume de l'ensemble G dans l'espace n-dimensionnel R_n.

f. Signalons quelques propriétés élémentaires des volumes dans R_n. *Si un ensemble $G \subset R_n$ a un volume, alors toute sa translation $G + a$, $a \in R_n$ a le même volume.* Ceci résulte du fait que les cellules dans R_n qui nous servent à mesurer les volumes (les pavés) admettent n'importe quelles translations, et cela sans changer de volume.

Si $E \subset R_n$ est un ensemble et λ un nombre positif, nous désignons par λE l'image de E par l'homothétie de rapport λ et de centre en l'origine des coordonnées. *Si G est un ensemble jordanien, alors λG en est encore un, et les volumes de ces ensembles sont liés par la relation* $|\lambda G| = \lambda^n |G|$ puisque celle-ci est valable pour chaque pavé.

En faisant la combinaison d'une homothétie et d'une translation, nous obtenons la relation $|\lambda(G + a)| = |\lambda G + a| = \lambda^n |G|$ *.

Il est un peu plus difficile de démontrer que les ensembles jordaniens ne changent pas de volume en subissant une transformation

---

* Si les cellules disponibles permettent d'effectuer non pas toutes les translations et homothéties, mais seulement celles qui sont déterminées, par exemple, par les valeurs rationnelles des paramètres a et λ, alors la relation en question se déduit à l'aide d'un passage supplémentaire à la limite.

orthogonale (rotation) ; on le verra dans *i*. Le raisonnement précédent ne s'étend pas directement à ce cas, puisque l'image d'un pavé, c'est-à-dire d'un parallélépipède rectangle dont les arêtes sont parallèles aux axes de coordonnées, par une rotation n'est plus un pavé.

g. Lemme. *Le cube obtenu par une transformation orthogonale d'un pavé cubique Q d'arête* 1 (donc aussi de volume 1) *a lui aussi* 1 *pour volume.*

Démonstration. Soit S la boule de centre à l'origine des coordonnées et de volume 1 ; pour $\varepsilon > 0$ donné et pour h suffisamment petit, on peut indiquer un corps T formé de $N = N(h)$ pavés cubiques d'arête h chacun, sans points intérieurs communs, de sorte que ce corps soit contenu à l'intérieur de la boule $(1 + \varepsilon) S$ et contienne lui-même la boule $(1 - \varepsilon) S$. Le volume du corps T est compris entre les volumes des boules :

$$(1 - \varepsilon)^n \leqslant h^n N(h) \leqslant (1 + \varepsilon)^n. \tag{3}$$

Effectuons une transformation orthogonale τ de l'espace. Le cube Q deviendra un cube τQ dont le volume nous désignons par v. Les pavés cubiques d'arête h qui composent le corps T deviendront des cubes homothétiques au cube τQ avec le rapport h et, en vertu de f, auront $h^n v$ pour volumes. Les boules $(1 - \varepsilon) S$ et $(1 + \varepsilon) S$ se transformeront en elles-mêmes. Le corps T se transformera en un corps τT composé du même nombre $N(h)$ de cubes d'arête h sans points intérieurs communs. Ce nouveau corps sera contenu dans la boule $(1 + \varepsilon) S$ et contiendra lui-même la boule $(1 - \varepsilon) S$. Son volume $N(h) \cdot h^n v$ sera également compris entre ceux des boules :

$$(1 - \varepsilon)^n \leqslant N(h)\, h^n v \leqslant (1 + \varepsilon)^n. \tag{4}$$

De (3) et (4) nous avons

$$\frac{(1-\varepsilon)^n}{(1+\varepsilon)^n} \leqslant v \leqslant \frac{(1+\varepsilon)^n}{(1-\varepsilon)^n}.$$

Comme $\varepsilon > 0$ est arbitraire, $v = 1$, ce qu'il fallait.

h. Lemme. *Le volume d'un parallélépipède rectangle ne change pas de valeur lors d'une transformation orthogonale.*

Démonstration. Si les longueurs des arêtes d'un parallélépipède rectangle donné sont commensurables, il peut être composé de cubes, et la conclusion du lemme découle du lemme *g* (et de *3.33b*). Dans le cas contraire, le parallélépipède peut être remplacé à une précision voulue (au sens de longueurs des arêtes, donc, aussi au sens de volume) par un parallélépipède dont les longueurs des arêtes soient commensurables, ce qui fournit la démonstration du lemme.

i. Théorème. *Le volume $|G|$ d'un ensemble jordanien $G \subset R_n$ ne change pas de valeur lors d'une transformation orthogonale.*

Démonstration. Pour $\varepsilon > 0$ donné, trouvons deux corps $G_\varepsilon^- \subset G$, $G_\varepsilon^+ \supset G$ formés de pavés sans point intérieur commun, de

sorte que

$$|G| - \varepsilon \leqslant |G_\varepsilon^-| \leqslant |G| \leqslant |G_\varepsilon^+| \leqslant |G| + \varepsilon. \qquad (5)$$

Les images des ensembles $G_\varepsilon^- \subset G \subset G_\varepsilon^+$ par une transformation orthogonale τ sont respectivement $\tau G_\varepsilon^- \subset \tau G \subset \tau G_\varepsilon^+$, et, vu le lemme h (et *3.33b*), on a $|\tau G_\varepsilon^+| = |G_\varepsilon^+|$, $|\tau G_\varepsilon^-| = |G_\varepsilon^-|$; par conséquent,

$$|G| - \varepsilon \leqslant |G_\varepsilon^-| = |\tau G_\varepsilon^-| \leqslant |\tau G| \leqslant |\tau G_\varepsilon^+| = |G_\varepsilon^+| \leqslant |G| + \varepsilon. \qquad (6)$$

Il résulte de (6) que $||G| - |\tau G|| \leqslant \varepsilon$. Comme $\varepsilon > 0$ est arbitraire, on a $|\tau G| = |G|$, ce qu'il fallait démontrer.

j. Le lemme qui suit relie les caractéristiques d'un ensemble négligeable dans R_n en termes de mesure et en termes de métrique.

Lemme. *Soit Z un ensemble négligeable dans un pavé $X \subset R_n$. Pour n'importe quels $\varepsilon > 0$ et $\delta > 0$ on peut trouver un ensemble élémentaire $B \subset X$ de mesure $\leqslant \varepsilon$ qui contient l'ensemble Z strictement dans son intérieur et qui est lui-même contenu dans le δ-voisinage de Z.*

Démonstration. D'après la définition, il existe pour $\varepsilon > 0$ donné un ensemble élémentaire $P \subset X$ qui contient l'ensemble Z strictement dans son intérieur et qui a une mesure $\leqslant \varepsilon$. Cet ensemble élémentaire P représente la réunion finie de quelques pavés qu'on peut choisir fermés; donc P est également fermé. Désignons par S la partie de P formée des points x dont la distance de Z est $\geqslant \delta$. L'ensemble S est fermé. Chaque point $x \in S$ appartient à une cellule (ouverte) de diamètre $< \delta/2$; toutes ces cellules forment un recouvrement de S et on peut en extraire un recouvrement fini de S. La réunion de toutes les cellules du recouvrement fini représente un ensemble élémentaire Q. La différence $B = P - QP$ est encore un ensemble élémentaire; il contient Z strictement dans son intérieur et ne contient aucun point de l'ensemble S, c'est-à-dire est contenu dans δ-voisinage de l'ensemble Z, ce qu'il fallait.

3.52. a. Nous avons défini les volumes des ensembles jordaniens n-dimensionnels dans R_n. Considérons un sous-espace k-dimensionnel R_k dans R_n. On peut le munir du produit scalaire emprunté à l'espace R_n et construire une base orthogonale $g_1, \ldots, g_k$; ensuite, en se servant des pavés dont les arêtes sont parallèles aux vecteurs $g_1, \ldots, g_k$, on peut établir une structure d'espace chargé et mesurer les volumes comme c'était indiqué dans *3.51* (pour $k = n$). Convenons de noter $|G|_k$ le volume d'un ensemble jordanien $G \subset R_k$. En vertu de *3.51i*, le nombre $|G|_k$ ne dépend pas du choix de la base orthogonale $g_1, \ldots, g_k$, puisque toute autre base s'obtient de celle-ci par une transformation orthogonale, ce qui ne change pas, on l'a vu, les valeurs des volumes des ensembles jordaniens. Cela fournit une possibilité principale de réduire une intégrale de multiplicité n à des intégrales de multiplicité inférieures. Ce fait est utilisé, par exemple, dans les théorèmes de la géométrie élémentaire sur le calcul des

volumes de certains corps en multipliant la hauteur par l'aire de la base.

b. Considérons un domaine borné $G \subset R_n$ dont la frontière se compose d'un nombre fini de surfaces ayant les équations de la forme $x_i = \varphi_i(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$, où φ_i sont des fonctions continues des variables indépendantes. Un tel domaine représente un corps jordanien (*3.51e*) sur lequel le schéma d'intégration donné dans *3.33* est valable. La projection E du domaine G sur le plan des coordonnées $x_1, \ldots, x_{n-1}$ est encore un corps jordanien (dans R_{n-1}) (puisque chaque point de la frontière de E est la projection d'un point de la frontière de G), et la frontière de E, de même que toute la frontière de G, admet une décomposition en nombre fini de parties dont chacune est décrite par une équation résolue par rapport à une coordonnée (il faut poser $x_n = 0$).

Considérons la projection E du domaine G comme espace chargé X' (dans l'espace R_{n-1}) avec les cellules correspondantes. La projection du domaine G sur l'axe des x_n se trouve sur un segment que nous considérons comme espace chargé X'' avec les cellules ordinaires (intervalles). Le corps G est inclus dans le produit cartésien $X' \times X''$; si l'on transforme ce produit en un espace chargé d'après *3.17*, ses cellules auront la même mesure que dans l'espace R_n.

c. Les sections verticales du corps $G \subset X' \times X''$ sont, en général, d'une forme assez compliquée. Supposons pour le moment que toute droite parallèle à l'axe des x_n et passant par G perce le domaine G suivant *un seul* segment déterminé, par exemple, par les inégalités de la forme

$$\varphi(x') \equiv \varphi(x_1, \ldots, x_{n-1}) \leqslant x_n \leqslant \psi(x_1, \ldots, x_{n-1}) = \psi(x'),$$
$$x' = (x_1, \ldots, x_{n-1})$$

(et dégénérant éventuellement en un point) (fig. 3.5-1).

Théorème. *Dans l'hypothèse formulée, pour toute fonction* $f(x) = f(x_1, \ldots, x_n)$ *continue dans le domaine* G, *on a la relation*

$$\int_G f(x)\,dx = \int_E \left\{ \int_{x_n = \varphi(x_1, \ldots, x_{n-1})}^{\psi(x_1, \ldots, x_{n-1})} f(x_1, \ldots, x_{n-1}, x_n)\,dx_n \right\} dx'. \qquad (1)$$

Ce résultat permet la réduction de l'intégrale sur un domaine n-dimensionnel à une intégration sur un domaine $(n-1)$-dimensionnel suivie d'une intégration unidimensionnelle.

Pour le démontrer, il suffit de poser dans le théorème général *3.34c* $U = G$, $X = E$, $Y = [a_n, b_n]$, $Y_x = \{x_n \in R_1 : \varphi(x') \leqslant x_n \leqslant \psi(x')\}$.

En fait, nous avons déjà vérifié toutes les conditions nécessaires.

d. Si le domaine G est d'une forme plus compliquée (certaines droites parallèles à l'axe des x_n et passant par G percent G suivant

plus qu'un seul segment) mais peut être décomposé en un nombre fini de domaines G_1, G_2, ... de la forme considérée plus haut qui n'aient les points communs que sur la frontière (fig. 3.5-2), alors

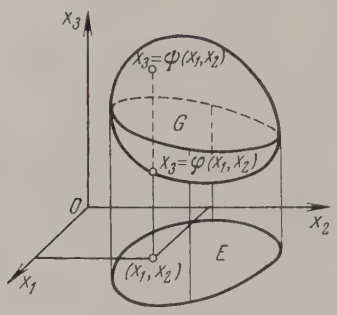


Fig. 3.5-1

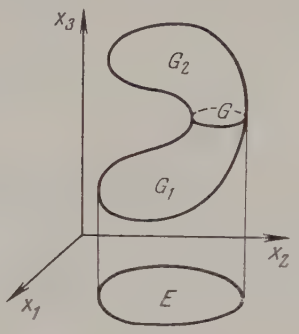


Fig. 3.5-2

l'intégrale sur le domaine G peut être représentée comme somme des intégrales sur les domaines simples, à chacune de ces intégrales on peut appliquer la méthode de calcul décrite plus haut.

3.53. Exemples.

a. Considérons le cas $n = 2$. Le théorème *3.52c* permet alors d'exprimer l'intégrale sur un domaine bidimensionnel G (fig. 3.5-3)

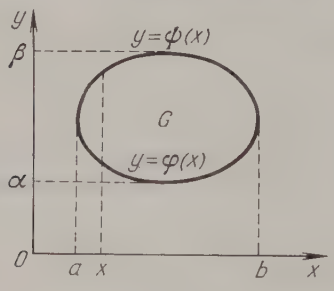


Fig. 3.5-3

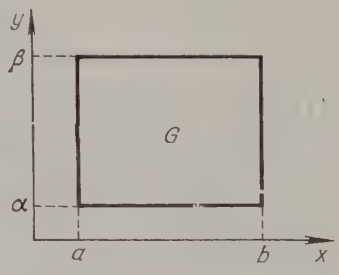


Fig. 3.5-4

par deux intégrales simples:

$$\int_G f(x, y)\,dx\,dy = \int_{x=a}^{b} \left\{ \int_{y=\varphi(x)}^{\psi(x)} f(x, y)\,dy \right\} dx. \qquad (1)$$

Si l'on intègre sur un rectangle dont les côtés sont parallèles aux axes de coordonnées (fig. 3.5-4), alors les limites de l'intégrale inté-

rieure sont constantes: $\varphi(x) = \alpha$, $\psi(x) = \beta$; par conséquent,

$$\int_G f(x, y)\,dx\,dy = \int_{x=a}^{b} \left\{ \int_{y=\alpha}^{\beta} f(x, y)\,dy \right\} dx.$$

Ainsi, lorsque $G = \{x, y: 3 \leqslant x \leqslant 4,\ 1 \leqslant y \leqslant 2\}$, $f(x, y) = \dfrac{1}{(x+y)^2}$, on a

$$\int_G \frac{dx\,dy}{(x+y)^2} = \int_{x=3}^{4} \left\{ \int_{y=1}^{2} \frac{\partial y}{(x+y)^2} \right\} dx = \int_3^4 \left(\frac{1}{x+1} - \frac{1}{x+2} \right) dx =$$

$$= [\ln(x+1) - \ln(x+2)]\,|_3^4 = \ln \frac{5 \cdot 5}{6 \cdot 4} = \ln \frac{25}{24}.$$

b. Si le domaine G n'est pas un rectangle de côtés parallèles aux axes de coordonnées, il faut tenir compte de la dépendance des limites de l'intégration intérieure de x.

En tant qu'exemple calculons l'intégrale de la fonction $f(x, y) = x^2/y^2$ sur le domaine G découpé par les droites $x = 2$, $y = x$ et par l'hyperbole $xy = 1$ (fig. 3.5-5). La projection du domaine G sur l'axe des x coïncide avec l'intervalle $[1, 2]$. En fixant un x sur cette intervalle, la droite parallèle à l'axe des y coupe le domaine G suivant l'intervalle $1/x \leqslant y \leqslant x$. En vertu de la formule (1) on a

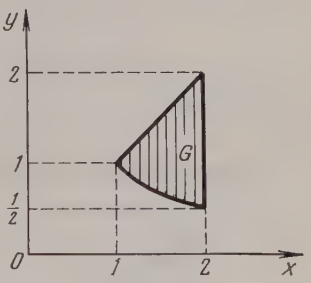


Fig. 3.5-5

$$\int_G \frac{x^2}{y^2}\,dx\,dy = \int_{x=1}^{2} \left\{ \int_{y=\frac{1}{x}}^{x} \frac{x^2}{y^2}\,dy \right\} dx =$$

$$= \int_1^2 x^2 \left[-\frac{1}{y} \right] \Big|_{y=\frac{1}{x}}^{x} dx = \int_1^2 x^2 \left(x - \frac{1}{x} \right) dx =$$

$$= \left(\frac{x^4}{4} - \frac{x^2}{2} \right) \Big|_1^2 = \frac{9}{4}.$$

Remarquons que, en déduisant la formule 3.52 (1), on aurait pu effectuer l'intégration intérieure non pas par rapport à la coordonnée x_n, mais par rapport à l'une quelconque des autres coordonnées $x_1, \ldots, x_{n-1}$. Le choix de tel ou tel ordre d'intégrations, qui est en principe sans importance, peut tout de même conduire à des calculs plus ou moins compliqués.

Si, dans l'exemple précédent, nous avions d'abord fixé non pas x, mais y, la droite horizontale correspondante aurait coupé le domaine G suivant l'in-

tervalle

$$\frac{1}{y} \leqslant x \leqslant 2 \quad \text{pour} \quad \frac{1}{2} \leqslant y \leqslant 1, \quad y \leqslant x \leqslant 2 \quad \text{pour} \quad 1 \leqslant y \leqslant 2;$$

les calculs auraient pris la forme suivante:

$$\int_G \frac{x^2}{y^2}\,dx\,dy = \int_{y=\frac{1}{2}}^{1} \left\{ \int_{x=\frac{1}{y}}^{2} \frac{x^2}{y^2}\,dx \right\} dy +$$

$$+ \int_{y=1}^{2} \left\{ \int_{x=y}^{2} \frac{x^2}{y^2}\,dx \right\} dy = \int_{\frac{1}{2}}^{1} \frac{1}{y^2} \frac{x^3}{3}\Big|_{x=\frac{1}{y}}^{2} dy +$$

$$+ \int_{1}^{2} \frac{1}{y^2} \frac{x^3}{3}\Big|_{x=y}^{2} dy = \frac{1}{3} \int_{\frac{1}{2}}^{1} \frac{1}{y^2} \left(8 - \frac{1}{y^3}\right) dy +$$

$$+ \frac{1}{3} \int_{1}^{2} \frac{1}{y^2} (8 - y^3)\,dy = \frac{1}{3} \left(-\frac{8}{y} + \frac{1}{4y^4} \right)\Big|_{\frac{1}{2}}^{1} +$$

$$+ \frac{1}{3} \left(-\frac{8}{y} - \frac{y^2}{2} \right)\Big|_{1}^{2} = \frac{17}{12} + \frac{5}{6} = \frac{9}{4}.$$

c. Donnons à present un exemple de calcul d'une intégrale triple en la ramenant à une chaîne d'intégrales unidimensionnelles.

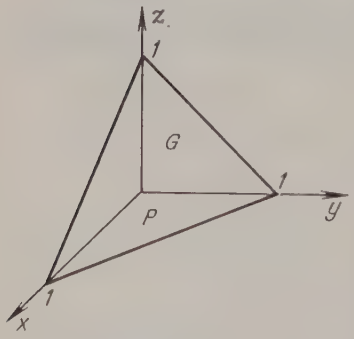


Fig. 3.5-6

Calculer l'intégrale

$$\int_G \frac{dx\,dy\,dz}{(1+x+y+z)^3},$$

où le domaine G est le tétraèdre limité par les plans $x = 0$, $y = 0$, $z = 0$, $x + y + z = 1$ (fig. 3.5-6).

La projection P du tétraèdre G sur le plan x, y représente un triangle qui est découpé dans ce plan par les droites $x = 0$, $y = 0$, $x + y = 1$. La droite verticale qui passe par un point fixe (x, y) coupe le domaine G suivant l'intervalle $0 \leqslant z \leqslant 1 - x - y$. Nous avons donc, d'après la formule (1):

$$\int_G \frac{dx\,dy\,dz}{(1+x+y+z)^3} = \int_P \left\{ \int_{z=0}^{1-x-y} \frac{dz}{(1+x+y+z)^3} \right\} dx\,dy.$$

Calculons l'intégrale intérieure:

$$\int_0^{1-x-y} \frac{dz}{(1+x+y+z)^3} = \frac{1}{2} \frac{1}{(1+x+y+z)^2}\Big|_{z=1-x-y}^{0} =$$

$$= \frac{1}{2} \left(\frac{1}{(1+x+y)^2} - \frac{1}{4} \right).$$

Nous avons à intégrer la fonction obtenue des variables x et y sur le domaine P. La projection de P sur l'axe des x est l'intervalle $0 \leqslant x \leqslant 1$. Fixons x sur cet intervalle; la droite correspondante qui est parallèle à l'axe des y coupe P suivant l'intervalle $0 \leqslant y \leqslant 1 - x$. Par conséquent

$$\int\limits_P \frac{1}{2}\left(\frac{1}{(1+x+y)^2}-\frac{1}{4}\right) dx\,dy = \int\limits_{x=0}^{1}\left\{\int\limits_{y=0}^{1-x}\frac{1}{2}\left(\frac{1}{(1+x+y)^2}-\frac{1}{4}\right) dy\right\} dx =$$

$$= \int\limits_0^1 \frac{1}{2}\left(\frac{1}{x+1}-\frac{3-x}{4}\right) dx = \frac{1}{2}\left(\ln 2-\frac{5}{8}\right).$$

d. Intégrale de Dirichlet. Etablissons la formule de Dirichlet

$$\int\limits_{x_n=0}^{x}\ldots\int\limits_{x_2=0}^{x_3}\int\limits_{x_1=0}^{x_2} f(x_1)\,dx_1\,dx_2\ldots dx_n = \frac{1}{\Gamma(n)}\int\limits_0^x (x-\xi)^{n-1} f(\xi)\,d\xi$$

qui réduit l'intégrale n-uple indéfinie d'une fonction d'une seule variable à une intégrale simple. Pour $n = 1$, la formule est immédiate. Supposons qu'elle soit juste pour un n naturel et démontrons-la pour $n + 1$. A cette fin calculons l'intégrale

$$I = \int\limits_{x_{n+1}=0}^{x}\left\{\frac{1}{\Gamma(n)}\int\limits_{\xi=0}^{x_{n+1}}(x_{n+1}-\xi)^{n-1} f(\xi)\,d\xi\right\} dx_{n+1}.$$

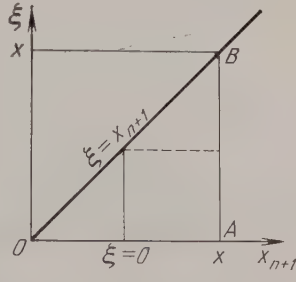


Fig. 3.5-7

L'intégration extérieure se fait par rapport à la variable x_{n+1} de 0 à x (fig. 3.5-7). L'intégration intérieure, pour x_{n+1} fixe, se fait par rapport à la variable ξ de 0 à x_{n+1}. En fin des comptes, nous arrivons à une intégrale sur l'intérieur du triangle OAB. Réduisons-la à deux intégrations simples, mais à condition que l'intégration extérieure soit effectuée par rapport à la variable ξ qui varie de 0 à x, et l'intégration intérieure, par rapport à x_{n+1} qui, dans les limites du triangle OAB, varie de $x_{n+1} = \xi$ à $x_{n+1} = x$, ξ étant fixe. Ainsi, nous écrivons

$$I = \int\limits_{\xi=0}^{x}\left\{\int\limits_{x_{n+1}=\xi}^{x}\frac{1}{\Gamma(n)} f(\xi)\,(x_{n+1}-\xi)^{n-1}\,dx_{n+1}\right\} d\xi =$$

$$= \frac{1}{\Gamma(n)}\int\limits_{\xi=0}^{x} f(\xi)\left\{\int\limits_{x_{n+1}=\xi}^{x}(x_{n+1}-\xi)^{n-1}\,dx_{n+1}\right\} d\xi.$$

Le calcul de l'intégrale intérieure fournit :

$$\int\limits_{x_{n+1}=\xi}^{x} (x_{n+1}-\xi)^{n-1}\,dx_{n+1} = \frac{(x_{n+1}-\xi)^n}{n}\Big|_{x_{n+1}=\xi}^{x} = \frac{(x-\xi)^n}{n},$$

donc

$$I = \frac{1}{n\Gamma(n)} \int\limits_0^x f(\xi)\,(x-\xi)^n\,d\xi = \frac{1}{\Gamma(n+1)} \int\limits_0^x f(\xi)\,(x-\xi)^n\,d\xi,$$

ce qu'il fallait.

3.54. a. Principe de Cavalieri. Si l'on utilise les « sections horizontales » d'un corps jordanien G, on peut arriver à une autre méthode de transformation d'une intégrale n-uple en une composition d'une intégrale $(n-1)$-uple et d'une intégrale simple. Supposons que le domaine G dans *3.52b* soit tel que la projection sur E de son intersection E_y avec tout plan $x_n = y$ soit une partie jordanienne du corps jordanien E (fig. 3.5-8). La formule *3.34* (4) donne alors :

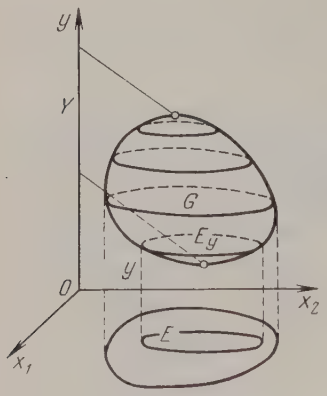


Fig. 3.5-8

$$\int\limits_G f(x)\,dx =$$

$$= \int\limits_Y \left\{ \int\limits_{E_y} f(x_1, \ldots, x_{n-1}, y)\,dx_1 \ldots dx_{n-1} \right\} dy,$$

Y étant la projection du domaine G sur l'axe des x_n.

L'intégrale intérieure s'étend sur la section E_y. En particulier, pour $f(x) \equiv 1$, nous obtenons une condition suffisante de l'égalité des volumes de deux corps (donnée par Cavalieri) :

Si deux corps $G^{(1)}$ et $G^{(2)}$ ont, pour tout y, les sections horizontales respectives $E_y^{(1)}$ et $E_y^{(2)}$ d'une même aire, alors les volumes de ces corps sont égaux.

La situation est particulièrement simple lorsque toutes les sections ont une même aire S. On a alors

$$|G| = \int\limits_G 1 \cdot dx = SmY.$$

b. Plan équilibrant et centre de gravité. Soit un plan ω dans l'espace R_3 ; à deux demi-espaces en lesquels il partage R_3 attribuons d'une façon quelconque les signes $+$ et $-$. Dans la mécanique, pour un point matériel $M(x, y, z) \in R_3$ qui porte une masse m, le produit de m par la distance $\rho(M, \omega)$ du point M au plan ω muni du signe $\varepsilon(M)$ du demi-espace contenant le point M est appelé *moment statique du point M par rapport au plan*

ω. Pour un corps $G \subset R_3$ dont la densité de masse est $\mu(M)$ *, on appelle *moment statique par rapport au plan* ω la quantité

$$P(G, \omega) = \iiint_G \varepsilon(M)\, \rho(M, \omega)\, \mu(M)\, dx\, dy\, dz. \tag{1}$$

Le plan ω est dit *équilibrant pour un corps G* si $P(G, \omega) = 0$.

Etant donné un corps G, trouvons son plan équilibrant ω de la forme $z = z_0$. Pour un point $M(x, y, z)$ et pour n'importe quel plan $z = z_0$, si l'on attribue au demi-espace $z > z_0$ le signe +, on a $\varepsilon(M)\, \rho(M, \omega) = z - z_0$, de sorte que

$$P(G, \omega) = \iiint_G (z - z_0)\, \mu(M)\, dx\, dy\, dz.$$

En annulant cette expression nous trouvons pour z_0 l'équation

$$\iiint_G \mu z\, dx\, dy\, dz = z_0 \iiint_G \mu\, dx\, dy\, dz,$$

d'où

$$z_0 = \frac{\iiint_G \mu z\, dx\, dy\, dz}{\iiint_G \mu\, dx\, dy\, dz}.$$

La quantité $\iiint_G \mu\, dx\, dy\, dz$ est la masse totale du corps G. On la suppose toujours positive.

Nous voyons en particulier que, *dans la famille des plans parallèles* $z = \text{const}$, *il existe un et un seul plan équilibrant*. De façon analogue, on peut trouver un plan équilibrant dans toute autre famille de plans parallèles. Il y a, par exemple, les plans équilibrants $x = x_0$ et $y = y_0$ avec

$$x_0 = \frac{\iiint_G \mu x\, dx\, dy\, dz}{\iiint_G \mu\, dx\, dy\, dz}, \qquad y_0 = \frac{\iiint_G \mu y\, dx\, dy\, dz}{\iiint_G \mu\, dx\, dy\, dz}. \tag{2}$$

---

* Du point de vue mathématique, la masse $m(Q)$ contenue dans un domaine Q est une fonction additive spéciale du domaine Q, et la densité $\mu(M)$ de la masse $m(Q)$ est la densité de cette fonction additive au sens de *3.42b*:

$$\mu(M) = \lim_{Q \to M} \frac{m(Q)}{|Q|}.$$

En vertu du théorème *3.41* la masse $m(Q)$ s'exprime par sa densité $\mu(M)$, à condition que cette dernière soit continue, par la formule

$$m(Q) = \int_G \mu(M)\, dv.$$

Pour un corps homogène ($\mu(x, y, z) = \text{const}$), les formules deviennent plus simples :

$$\left.\begin{aligned} x_0 &= \frac{\iiint\limits_G x\,dx\,dy\,dz}{\iiint\limits_G dx\,dy\,dz}, \quad y_0 = \frac{\iiint\limits_G y\,dx\,dy\,dz}{\iiint\limits_G dx\,dy\,dz}, \\ z_0 &= \frac{\iiint\limits_G z\,dx\,dy\,dz}{\iiint\limits_G dx\,dy\,dz}. \end{aligned}\right\} \tag{3}$$

Le dénominateur dans ces formules coïncide évidemment avec le volume du corps G.

Le point des coordonnées x_0, y_0, z_0 s'appelle *centre de gravité du corps G*. Il est à noter que *tout plan équilibrant passe par ce point*. Afin de le prouver supposons que $x_0 = y_0 = z_0 = 0$ (sinon on fera une translation du corps), autrement dit, que

$$\iiint\limits_G \mu x\,dx\,dy\,dz = \iiint\limits_G \mu y\,dx\,dy\,dz = \iiint\limits_G \mu z\,dx\,dy\,dz = 0. \tag{4}$$

Maintenant il est à vérifier que tout plan ω passant par l'origine des coordonnées soit équilibrant.

Un tel plan ω peut être donné par son vecteur unitaire normal $m = (\cos\alpha, \cos\beta, \cos\gamma)$. Attribuons aux demi-espaces déterminés par le plan ω les signes $+$ et $-$ d'une telle façon que le vecteur m soit dirigé vers le sous-espace muni de $+$. Si l'on en tient compte, la distance d'un point $M(x, y, z)$ au plan ω s'exprime par la formule

$$\varepsilon(M)\,\rho(M, \omega) = (M, m) = x\cos\alpha + y\cos\beta + z\cos\gamma.$$

C'est pourquoi le moment du corps G par rapport au plan ω sera

$$P(G, \omega) = \iiint\limits_G (x\cos\alpha + y\cos\beta + z\cos\gamma)\,\mu\,dx\,dy\,dz = 0$$

d'après les relations (4).

Les raisonnements ci-dessus restent évidemment valables pour les corps dans l'espace R_n, quel que soit $n = 1, 2, \ldots$

Pour le calcul des expressions (2) et (3) qui déterminent les coordonnées du centre de gravité, il est naturel d'appliquer le principe de Cavalieri. Ainsi, en effectuant l'intégration intérieure sur une section horizontale du corps G dans la dernière expression (3), nous aboutissons à l'aire $S(z)$ de cette section après quoi l'expression de la coordonnée z_0 se transforme en un rapport de deux intégrales simples

$$z_0 = \frac{\int\limits_{z_1}^{z_2} zS(z)\,dz}{\int\limits_{z_1}^{z_2} S(z)\,dz},$$

où z_1 et z_2 sont les troisièmes coordonnées des points inférieur et supérieur respectivement du corps G. Il va de soi qu'on peut procéder d'une manière analogue avec les autres expressions (2) et (3).

3.55. Volume k-dimensionnel d'un parallélépipède k-dimensionnel.

a. On appelle *parallélépipède k-dimensionnel engendré par les vecteurs* $g_1, \ldots, g_k$ (dans n'importe quel espace vectoriel) l'ensemble P_k de vecteurs x décrit par la formule

$$x = \alpha_1 g_1 + \ldots + \alpha_k g_k, \quad 0 \leqslant \alpha_i \leqslant 1, \quad i = 1, \ldots, k$$

(voir fig. 3.5-9 pour $k = 3$).

Si l'on munit l'espace k-dimensionnel R_k engendré par les vecteurs $g_1, \ldots, g_k$ d'un produit scalaire, alors, d'après *3.51a*, on peut introduire dans n'importe quel son pavé une structure d'espace chargé; en vertu de *3.51e*, le parallélépipède P_k devient ensemble jordanien, donc on peut lui faire correspondre un nombre $V_k(P_k)$, son volume k-dimensionnel.

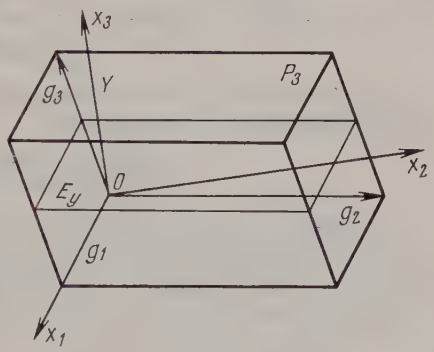


Fig. 3.5-9

Soit $R_1 \subset R_2 \subset \ldots \subset R_n$ un système d'espaces de dimension finie engendrés respectivement par un, deux, ..., n premiers vecteurs d'un ensemble donné de n vecteurs linéairement indépendants, $g_1, \ldots, g_n$ ($\in R_n$), et supposons que, dans chacun de ces espaces, le produit scalaire soit celui du plus grand espace R_n. Il y a des liens entre les volumes $V_1, \ldots, V_n$ des parallélépipèdes $P_1, \ldots, P_n$ engendrés respectivement par les vecteurs g_1; g_1, g_2; ...; $g_1, g_2, \ldots, g_n$ que nous allons mettre en évidence. Comme les volumes ne changent pas lors d'une transformation orthogonale (*3.51i*), on peut supposer que les axes dans l'espace R_n soient disposés de façon que les vecteurs $g_1, \ldots, g_{n-1}$ se trouvent dans l'hyperplan $x_n = 0$.

La section horizontale du parallélépipède

$$P_n = \{x \in R_n : x = \alpha_1 g_1 + \ldots + \alpha_{n-1} g_{n-1} + \alpha_n g_n, \quad 0 \leqslant \alpha_i \leqslant 1, \ i = 1, \ldots, n\}$$

par $x_n = y$ a la forme

$$E_y = \{x \in R_n : x = \alpha_1 g_1 + \ldots + \alpha_{n-1} g_{n-1} + \alpha_n (y)\, g_n, \quad 0 \leqslant \alpha_i \leqslant 1, \ i = 1, \ldots, n-1\}$$

et représente le translaté de la base inférieure

$$E_0 = \{x \in R_n : x = \alpha_1 g_1 + \ldots + \alpha_{n-1} g_{n-1}\}$$

d'un vecteur constant $\alpha_n(y)\, g_n$. C'est pourquoi toutes les sections E_y ont une même aire (égale à V_{n-1}). Ainsi donc,

$$V_n = V_{n-1} mY = V_{n-1} h_n,$$

où $mY = h_n$ est la longueur de la projection du vecteur g_n sur l'axe $y = x_n$, i.e. la hauteur du parallélépipède P_n. Définitivement, *le volume d'un parallélépipède est égal au produit de l'aire de n'importe quelle sa section horizontale, en particulier, de la base, par la longueur de sa hauteur*. En itérant ce résultat et appliquant la relation évidente $V_1 = h_1 = |g_1|$, nous avons

$$V_n = V_{n-1} h_n = V_{n-2} h_{n-1} h_n = \ldots$$
$$\ldots = V_1 h_2 \ldots h_n = h_1 h_2 \ldots h_n. \qquad (1)$$

b. Avec la formule (1) on peut démontrer d'une façon déjà purement algébrique que (*L8.73*)

$$V_n = \sqrt{\det \| (g_i, g_j) \|} = |\det \| \xi_i^{(j)} \||, \qquad (2)$$

où $\xi_i^{(j)}$ est la i-ème coordonnée du vecteur g_j dans une base orthonormée quelconque (mais fixe) de l'espace R_n.

c. D'après *3.52a*, le volume k-dimensionnel du parallélépipède k-dimensionnel construit sur les vecteurs $g_1, \ldots, g_k$ dans R_n peut être calculé sans sortir de l'espace k-dimensionnel R_k engendré par les vecteurs $g_1, \ldots, g_k$, avec le produit scalaire emprunté à R_n. Cela conduit à la formule

$$V_k = \sqrt{\det \| (g_i, g_j) \|_{i,\, j=1,\, \ldots,\, k}}. \qquad (3)$$

Rappelons ici qu'a lieu l'égalité

$$\det \| (g_i, g_j) \|_{i,\, j=1,\, \ldots,\, k} = \sqrt{\sum_{i_1 < \ldots < i_k} [M^{1 \ldots k}_{i_1 \ldots i_k} (g_1, \ldots, g_k)]^2}, \qquad (4)$$

où $M^{1 \ldots k}_{i_1 \ldots i_k}$ est le mineur d'ordre k de la matrice $\| \xi_i^{(j)} \|$ ($i = 1, \ldots, n$; $j = 1, \ldots, k$) correspondant aux lignes de numéros $i_1, \ldots, i_k$ (*L8.73*).

Nous allons appliquer la formule (4) à la déduction de certaines propriétés du volume k-dimensionnel dont nous aurons besoin par la suite.

d. Le volume du parallélépipède construit sur les vecteurs $g_1, \ldots, g_k$ sera également désigné par $|[g_1, \ldots, g_k]|$. (Le symbole $[g_1, \ldots, g_k]$ a aussi un sens que nous n'explicitons pas pour le moment.)

Si les vecteurs $g_1, \ldots, g_i$ sont orthogonaux à chacun des vecteurs $g_{i+1}, \ldots, g_k$, alors

$$|[g_1, \ldots, g_k]| = |[g_1, \ldots, g_i]|\, |[g_{i+1}, \ldots, g_k]|.$$

En effet, on a

$$|[g_1, \ldots, g_k]| =$$

$$= \begin{vmatrix} (g_1, g_1) \ldots (g_1, g_i) & 0 & \ldots & 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots & & & \\ (g_i, g_1) \ldots (g_i, g_i) & 0 & \ldots & 0 \\ 0 \quad \ldots \quad 0 & (g_{i+1}, g_{i+1}) & \ldots & (g_{i+1}, g_k) \\ \ldots\ldots\ldots\ldots\ldots\ldots & & & \\ 0 \quad \ldots \quad 0 & (g_k, g_{i+1}) & \ldots & (g_k, g_k) \end{vmatrix} =$$

$$= \begin{vmatrix} (g_1, g_1) \ldots (g_1, g_i) \\ \ldots\ldots\ldots\ldots \\ (g_i, g_1) \ldots (g_i, g_i) \end{vmatrix} \cdot \begin{vmatrix} (g_{i+1}, g_{i+1}) \ldots (g_{i+1}, g_k) \\ \ldots\ldots\ldots\ldots\ldots \\ (g_k, g_{i+1}) \ldots (g_k, g_k) \end{vmatrix} =$$

$$= |[g_1, \ldots, g_i]| \cdot |[g_{i+1}, \ldots, g_k]|,$$

ce qu'il fallait.

e. *Si c est un nombre fixe, alors*

$$|[g_1 + cg_2, g_2, \ldots, g_k]| = |[g_1, g_2, \ldots, g_k]|,$$

ce qui résulte immédiatement de (4) et des propriétés des mineurs. Il en découle aisément que *le volume d'un parallélépipède k-dimensionnel ne change pas lorsqu'on ajoute à n'importe lequel des vecteurs qui l'engendrent g_i $(i = 1, \ldots, k)$ une combinaison linéaire des autres.*

f. Pour $|g_1| \leqslant M, \ldots, |g_k| \leqslant M, |h_1| \leqslant m, \ldots, |h_k| \leqslant m$ $(m \leqslant 1, M \geqslant m)$, *les volumes des parallélépipèdes engendrés respectivement par les vecteurs $g_1 + h_1, \ldots, g_k + h_k$ et $g_1, \ldots, g_k$ sont liés par la formule*

$$|[g_1 + h_1, \ldots, g_k + h_k]|^2 =$$
$$= |[g_1, \ldots, g_k]|^2 + C_n m (M + m)^{2k-1}\theta, \quad |\theta| \leqslant 1, \quad (5)$$

où C_n est une constante $(C_n \leqslant 3n^2 (n!)^3)$.

En effet, soient $g_j = \sum_{i=1}^{n} a_{ij}e_i$, $h_j = \sum \lambda_{ij}e_i$ les développements des vecteurs g_j et h_j suivant une base orthonormée $\{e_i\}$. Désignons par (i) l'indice composé $(i_1, \ldots, i_k)$ où $i_1 < \ldots < i_k$; soit $j = (j_1, \ldots, j_k)$ n'importe quelle permutation des nombres $i_1, \ldots, i_k$ et $\varepsilon(j)$ son signe. Alors

$$|[g_1 + h_1, g_2, \ldots, g_k]|^2 = \sum_{(i)} [\sum_{(j)} \varepsilon(j) (a_{1j_1} + \lambda_{1j_1}) a_{2j_2} \ldots a_{kj_k}]^2 =$$
$$= \sum_{(i)} [\sum_{(j)} \varepsilon(j) a_{1j_1} a_{2j_2} \ldots a_{kj_k} + \sum_{(j)} \varepsilon(j) \lambda_{1j_1} a_{2j_2} \ldots a_{kj_k}]^2 =$$
$$= \sum_{(i)} [\sum_{(j)} \varepsilon(j) a_{1j_1} a_{2j_2} \ldots a_{kj_k}]^2 + R = |[g_1, g_2, \ldots, g_k]|^2 + R$$

où, en désignant toujours par ν une permutation arbitraire des nombres $i_1, \ldots, i_k$, on a

$$|R| \leqslant \sum_{(i)} \Big| 2 \sum_{(j)(\nu)} \varepsilon(j)\,\varepsilon(\nu)\, a_{1j_1} \ldots a_{kj_k} \lambda_{1\nu_1} a_{2\nu_2} \ldots a_{k\nu_k} + \\ + \Big[\sum_{(j)} \varepsilon(j)\, \lambda_{1j_1} a_{2j_2} \ldots a_{kj_k}\Big]^2 \Big| \leqslant \\ \leqslant \sum_{(i)} \Big[2 \sum_{(j)(\nu)} M^{2k-1} m + \Big(\sum_{(j)} m M^{k-1}\Big)^2\Big] \leqslant \\ \leqslant 3 m M^{2k-1} (k!)^3 \leqslant A_n m M^{2k-1},$$

A_n désignant une constante :

$$A_n = 3 \max_k (k!)^3 \leqslant 3 (n!)^3.$$

Ainsi, nous avons

$$|\,[g_1 + h_1, g_2, \ldots, g_k]\,|^2 = |\,[g_1, g_2, \ldots, g_k]\,|^2 = \\ = A_n m M^{2k-1} \theta_1, \quad |\,\theta_1\,| \leqslant 1.$$

D'une manière analogue :

$$|\,[g_1 + h_1, g_2 + h_2, g_3, \ldots, g_k]\,|^2 = \\ = |\,[g_1 + h_1, \ldots, g_2, \ldots, g_k]\,|^2 + A_n m (M + m)^{2k-1} \theta_2, \\ |\,\theta_2\,| \leqslant 1,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$|\,[g_1 + h_1, \ldots, g_k + h_k]\,|^2 = |\,[g_1 + h_1, \ldots, g_{k-1} + \\ + h_{k-1}, g_k]\,|^2 + A_n m (M + m)^{2k-1} \theta_k, \quad |\,\theta_k\,| \leqslant 1.$$

En additionnant les égalités obtenues nous retrouvons (5).

3.56. Autres exemples.

a. Volume d'un simplexe. On appelle simplexe k-dimensionnel engendré par les vecteurs $g_1, \ldots, g_k$ (dans n'importe quel espace vectoriel) l'ensemble Σ_k de vecteurs x décrit par la formule

$$x = \alpha_1 g_1 + \ldots + \alpha_k g_k, \quad 0 \leqslant \alpha_i \leqslant 1, \quad \sum_{i=1}^{k} \alpha_i \leqslant 1 \qquad (1)$$

(enveloppe convexe des points $g_1, \ldots, g_k$ et 0). D'une façon analogue à *3.55a*, les simplexes k-dimensionnels dans l'espace euclidien R_n possèdent leurs volumes k-dimensionnels. Il existe un lien entre les volumes $|\,\Sigma_1\,|, |\,\Sigma_2\,|, \ldots, |\,\Sigma_n\,|$ des simplexes $\Sigma_1, \Sigma_2, \ldots, \Sigma_n$ engendrés respectivement par les vecteurs g_1 ; g_1, g_2 ; $\ldots$; $g_1, \ldots, g_n$ que nous allons établir.

Appelons hauteur d'un simplexe Σ_k la longueur h_k de la perpendiculaire abaissée de l'extrémité du vecteur g_k sur le sous-espace R_{k-1}

engendré par les vecteurs $g_1, \ldots, g_{k-1}$ (fig. 3.5-10). Le simplexe Σ_k est donné par les conditions (1); sa section par le plan parallèle

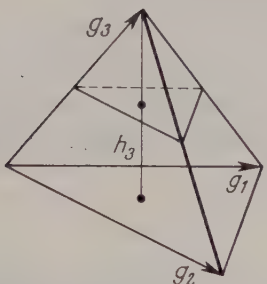


Fig. 3.5-10

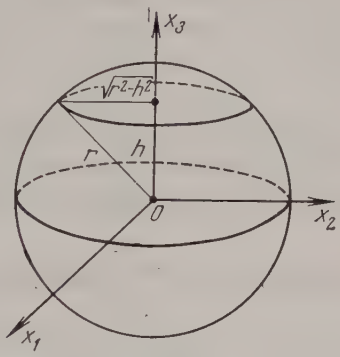


Fig. 3.5-11

à R_{k-1} et éloigné de Σ_{k-1} d'une distance s est déterminée par les conditions

$$x = \frac{s}{h_k}(\alpha_1 g_1 + \ldots + \alpha_{k-1} g_{k-1} + g_k),$$

$$0 \leqslant \alpha_i \leqslant 1 \quad (i = 1, \ldots, k-1), \quad \sum_{i=1}^{k-1} \alpha_i \leqslant 1.$$

Il est évident que cette section représente elle aussi un simplexe obtenu de Σ_{k-1} par une translation et une homothétie de rapport s/h_k. D'après *3.51f*, l'aire de la section vaut $\left(\frac{s}{h_k}\right)^{k-1} |\Sigma_{k-1}|$.

Il en résulte en vertu du principe de Cavalieri *3.54*:

$$|\Sigma_k| = \int_0^{h_k} \left(\frac{s}{h_k}\right)^{k-1} |\Sigma_{k-1}|\, ds = \frac{h^k}{k} |\Sigma_{k-1}|.$$

Etant donnée la relation évidente $|\Sigma_1| = h_1 = |g_1|$, on a

$$|\Sigma_n| = |\Sigma_{n-1}| \frac{h_n}{n} = \ldots = \frac{h_1 \ldots h_n}{n!}.$$

En rapprochant ceci et le volume du parallélépipède *3.55* (1) nous obtenons:

$$|\Sigma_n| = \frac{1}{n!} |[g_1, \ldots, g_n]|.$$

b. Volume d'une boule n-dimensionnelle. Soit

$$S_n(r) = \{x \in R_n : |x| \leqslant r\}.$$

Deux boules d'une même dimension étant homothétiques, on a $| S_n(r) | = r^n S_n(1) = C_n r^n$ où C_n est une constante à trouver. La section de la boule $S_n(r)$ par un hyperplan $x_n = h$, $| h | \leqslant r$ est une boule $(n-1)$-dimensionnelle de rayon $\sqrt{r^2 - h^2}$ (fig. 3.5-11). Conformément au principe de Cavalieri *3.54a*, on a

$$| S_n(r) | = \int_{-r}^{r} | S_{n-1}(\sqrt{r^2 - h^2}) | \, dh = C_{n-1} \int_{-r}^{r} (r^2 - h^2)^{\frac{n-1}{2}} \, dh.$$

Effectuons la substitution $h = r \sin \theta$; nous aurons (*U11.54*) :

$$C_n r^n = | S_n(r) | = C_{n-1} r^n \int_{-\pi/2}^{\pi/2} \cos^n \theta \, d\theta =$$

$$= C_{n-1} r^n B\left(\frac{1}{2}, \frac{n+1}{2}\right) = C_{n-1} r^n \frac{\Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}+1\right)},$$

d'où

$$\frac{C_n}{C_{n-1}} = \frac{\Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}+1\right)}.$$

Comme $C_1 = S_1(1) = 2$, on a $C_2 = 2 \dfrac{\Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{3}{2}\right)}{\Gamma(2)} = \pi$, $C_3 = \frac{4}{3}\pi$, et en général :

$$C_n = 2 \frac{\Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{3}{2}\right)}{\Gamma(2)} \cdot \frac{\Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{4}{2}\right)}{\Gamma\left(\frac{5}{2}\right)} \cdots \frac{\Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{n}{2}+\frac{1}{2}\right)}{\Gamma\left(\frac{n}{2}+1\right)} =$$

$$= 2 \frac{\Gamma^{n-1}\left(\frac{1}{2}\right) \Gamma\left(\frac{3}{2}\right)}{\Gamma\left(\frac{n}{2}+1\right)} = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)}.$$

Définitivement :

$$| S_n(r) | = \frac{\pi^{n/2} r^n}{\Gamma\left(\frac{n}{2}+1\right)}. \qquad (2)$$

3.57. Coefficient de distorsion d'un volume par une application dérivable.

a. Soit $x = x(u)$ une application dérivable d'un domaine borné $U \subset R_n$ dans un domaine borné $X \subset R_n$; sous une forme plus déve-

loppée, l'application $x(u)$ s'écrit :

$$\left.\begin{aligned} x_1 &= x_1(u_1, \ldots, u_n), \\ &\cdots\cdots\cdots\cdots \\ x_n &= x_n(u_1, \ldots, u_n). \end{aligned}\right\} \tag{1}$$

En vertu de *3.51b*, l'image de la frontière de chaque pavé $B \subset U$ est un ensemble négligeable dans le domaine X ; par conséquent, l'application $x(u)$ transforme chaque pavé B en un ensemble jordanien $x(B)$. Nous tenons à calculer le coefficient de distorsion de la mesure (*3.45*) :

$$\varphi(u) = \lim_{B \to u} \frac{|x(B)|}{|B|}.$$

Si $x(u)$ est une application linéaire, alors $x(B)$ est le parallélépipède dont les arêtes sont les images par $x(u)$ de celles du pavé B. Soient

$$\begin{aligned} &(\Delta u_1, \ 0, \ \ldots, \ 0), \\ &(0, \ \Delta u_2, \ \ldots, \ 0), \\ &\cdots\cdots\cdots\cdots \\ &(0, \ 0, \ \ldots, \ \Delta u_n) \end{aligned}$$

les arêtes du pavé B et

$$\left.\begin{aligned} x_1 &= a_{11}u_1 + \ldots + a_{1n}u_n, \\ &\cdots\cdots\cdots\cdots \\ x_n &= a_{n1}u_1 + \ldots + a_{nn}u_n \end{aligned}\right\}$$

l'application $x = x(u)$. Alors les images des arêtes du pavé B seront respectivement

$$\begin{aligned} &(a_{11}, a_{21}, \ldots, a_{n1})\,\Delta u_1, \\ &(a_{12}, a_{22}, \ldots, a_{n2})\,\Delta u_2, \\ &\cdots\cdots\cdots\cdots \\ &(a_{1n}, a_{2n}, \ldots, a_{nn})\,\Delta u_n, \end{aligned}$$

et le volume du parallélépipède $x(B)$ se trouvera égal à (*3.55c*)

$$|x(B)| = |\det \| a_{ij} \| \, | \Delta u_1 \ldots \Delta u_n = |\det \| a_{ij} \| \, | \, |B|.$$

Donc,

$$\frac{|x(B)|}{|B|} = |\det \| a_{ij} \| \, |, \tag{2}$$

et c'est bien la valeur du coefficient de distorsion de la mesure.

b. Considérons à présent une application dérivable arbitraire $x(u)$. A un point donné $a \in U$ dégageons la partie linéaire principale

$$x(a + \Delta u) = x'(a)\,\Delta u + o(\Delta u). \tag{3}$$

L'opérateur $x'(a)$ $(R_n \to R_n)$ est donné par la matrice de Jacobi (*1.25a*)

$$x'(a) \cong \left\| \begin{matrix} \frac{\partial x_1(a)}{\partial u_1} & \cdots & \frac{\partial x_n(a)}{\partial u_1} \\ \cdots & \cdots & \cdots \\ \frac{\partial x_1(a)}{\partial u_n} & \cdots & \frac{\partial x_n(a)}{\partial u_n} \end{matrix} \right\|.$$

Théorème. *Pour une application dérivable $x(u)$, le coefficient de distorsion du volume peut être calculé d'après la formule*

$$\varphi(a) \equiv \lim_{B \to a} \frac{|x(B)|}{|B|} = |\det x'(a)|. \tag{4}$$

Démonstration sera effectuée en plusieurs étapes. Nous démontrerons d'abord la relation limite (4) pour le cas où les pavés B se contractant vers le point a sont des cubes. Nous verrons ensuite qu'elle reste valable pour n'importe quels pavés et même pour n'importe quels ensembles jordaniens se contractant vers le point a (*3.58b*).

c. Supposons d'abord que l'application $x'(a)$ ne soit pas dégénérée, de sorte que $\det x'(a) \neq 0$. D'après le théorème *1.56* sur la fonction inverse, il existe un domaine V contenant le point a et qui s'applique de façon difféomorphe sur un domaine $W \subset X$ contenant le point $b = x(a)$.

Lorsque le vecteur $h \in R_n$ parcourt le cube $Q = \left\{ -\frac{1}{2} \leqslant x_i \leqslant \frac{1}{2},\ i = 1, \ldots, n \right\}$, le vecteur $x'(a)\,h$ parcourt un parallélépipède non dégénéré S de volume $|S| = |\det x'(a)|$. Comme il est non dégénéré, le parallélépipède S contient à son intérieur une boule d'un rayon $c > 0$. Fixons arbitrairement un $\varepsilon > 0$ et considérons le $c\varepsilon$-voisinage $S_{c\varepsilon}$ du parallélépipède S. Tout le voisinage est recouvert par la réunion du parallélépipède S avec les translatés du parallélépipède εS, chacun de ces derniers ayant des points communs avec S, puisque contenant une boule de rayon $c\varepsilon$. Ainsi, tout le $c\varepsilon$-voisinage est recouvert par le parallélépipède $(1 + \varepsilon)\,S$. Nous obtenons pour le volume de l'ensemble $S_{c\varepsilon}$ l'inégalité suivante:

$$|S_{c\varepsilon}| \leqslant (1 + \varepsilon)^n |S| = (1 + \varepsilon)^n |\det x'(a)|.$$

Quel que soit un pavé cubique A d'arête δ, le volume du $c\varepsilon\delta$-voisinage du parallélépipède $x'(a)\,A$ est estimé, pour la raison d'homothétie, comme suit:

$$|(x'(a)\,A)_{c\varepsilon\delta}| \leqslant (1 + \varepsilon)^n |x'(a)\,A| =$$
$$= (1 + \varepsilon)^n,\ \delta^n |\det x'(a)| = (1 + \varepsilon)^n |A| |\det x'(a)|.$$

D'après $\varepsilon > 0$ donné, trouvons $\delta > 0$ tel que la quantité $o(\Delta u)$ dans (3) admette pour tout $|\Delta u| < \delta \sqrt{n}$ la majoration $|(o(\Delta u)| \leqslant \leqslant c\varepsilon |\Delta u|$. Soient $A \subset V$ un pavé cubique d'arête δ contenant le point 0 à l'intérieur ou sur la frontière et $B = a + A$. Pour $h \in A$ le point $a + h$ est éloigné de a de $\delta \sqrt{n}$ au plus, donc la distance du vecteur $x(a + h)$ au vecteur $x(a) + x'(a) h$ ne dépasse pas $c\varepsilon\delta$, c'est-à-dire le premier d'eux appartient au $c\varepsilon\delta$-voisinage du parallélépipède $b + x'(a) A$. Par conséquent, on peut trouver un majorant de $|x(B)|$ exprimé par $|B|$:

$$|x(B)| \leqslant |(b + x'(a) A)_{c\varepsilon\delta}| \equiv |(x'(a) A)_{c\varepsilon\delta}| \leqslant$$

$$\leqslant (1 + \varepsilon)^n |A| |\det x'(a)| = |(1 + \varepsilon)^n |B| \cdot |\det x'(a)|. \quad (5)$$

Afin de minorer $|x(B)|$ à l'aide de $|B|$, raisonnons comme suit.

Désignons par $b + D$ le parallélépipède concentrique et homothétique au parallélépipède $b + x'(a) A$, avec $1 - \varepsilon$ pour rapport d'homothétie. Il est clair que $D = x'(a) C$ où $a + C$ est le pavé cubique concentrique et homothétique au pavé cubique $a + A$, avec le même rapport d'homothétie $1 - \varepsilon$.

Soit $u = u(x)$ l'inverse de la fonction $x(u)$, de sorte que $a = u(b)$. La fonction $u(x)$ est dérivable dans le domaine W, et l'on a

$$u(b + k) = a + u'(b) k + o(k). \quad (6)$$

Pour $\varepsilon > 0$ donné, trouvons $\delta > 0$ tel que la quantité $o(k)$ dans (6) admette la majoration $|o(k)| \leqslant \varepsilon |k|$ pour tout $|k| < \delta$. Certainement, on peut admettre que ce soit le même δ qui amène à l'inégalité (5). L'opérateur $u'(b)$ applique le parallélépipède S inversement sur le pavé cubique Q, et le parallélépipède D sur le pavé cubique C. L'image du parallélépipède $b + D$ par l'application $x \to u(x)$ est contenue dans le $\varepsilon\delta$-voisinage du cube $a + C$ et, à fortiori, dans le cube B; en d'autres termes,

$$u(b + x'(a) C) \subset B$$

ou, ce qui revient au même,

$$x(B) \supset b + x'(a) C.$$

Il en découle l'inégalité suivante pour les volumes:

$$|x(B)| \geqslant |b + x'(a) C| = |x'(a) C| =$$

$$= (1 - \varepsilon)^n |x'(a) A| = (1 - \varepsilon)^n |B| |\det x'(a)|. \quad (7)$$

En rapprochant les inégalités (5) et (7) nous pouvons donc affirmer que, pour tout $\varepsilon > 0$ et pour tout pavé cubique B d'arête $\delta \leqslant \delta(\varepsilon)$ contenant le point a, a lieu l'estimation double

$$(1 - \varepsilon)^n |\det x'(a)| \leqslant \frac{|x(B)|}{|B|} \leqslant (1 + \varepsilon)^n |\det x'(a)|.$$

Cela démontre le théorème dans le cas où $\det x'(a) \neq 0$.

d. A présent, soit $\det x'(a) = 0$, mais $x'(a) \neq 0$. Dans ce cas l'opérateur $x'(a)$ transforme le cube Q en le parallélépipède dégénéré $S = x'(a)\,Q$. Désignons par r, $0 < r < n$, le rang de la matrice $x'(a)$. Le parallélépipède S se trouve dans un plan r-dimensionnel γ et contient une boule r-dimensionnelle de centre à l'origine des coordonnées et d'un rayon c. D'après $\varepsilon > 0$ trouvons $\delta > 0$ tel que la quantité $o(\Delta u)$ de l'égalité (3) admette l'estimation $|o(h)| \leqslant c\varepsilon\,|h|$ pour tout $\Delta u = h$, $|h| < \delta$; le vecteur $x(a+h)$, pour tout $h \in A$, appartient au $c\varepsilon\delta$-voisinage du parallélépipède $b + x'(a)\,A$. Or, l'intersection du plan γ et de ce voisinage, de même que dans c, se trouve dans un parallélépipède T de volume (r-dimensionnel) $\leqslant (1+\varepsilon)^r\,|B|\,|S|_r$. En outre, le $c\varepsilon\delta$-voisinage mentionné contient certains points en dehors du plan γ; ils appartiennent tous au produit du parallélépipède T par une boule de rayon $\varepsilon\delta$ dans le sous-espace $(n-r)$-dimensionnel, orthogonal au plan γ, de l'espace R_n. Pour cette raison, le volume du domaine $x(B)$ ne dépasse pas la quantité

$$|T|_r\,(\varepsilon\delta)^{n-r} \leqslant (1+\varepsilon)^r\,\varepsilon^{n-r}\delta^{n-r}\delta^r\,|S_r| = (1+\varepsilon)^r\,\varepsilon^{n-r}\,|B|\,|S_r|.$$

Ainsi, dans le cas considéré nous avons

$$\frac{|x(B)|}{|B|} \leqslant (1+\varepsilon)^r\,\varepsilon^{n-r}\,|S_r|.$$

Pour $B \to a$ ou, ce qui est le même, $\varepsilon \to 0$, nous obtenons

$$\lim_{B \to a} \frac{|x(B)|}{|B|} = 0.$$

e. Il nous reste un seul cas où $x'(a) = 0$, de sorte que l'opérateur $x'(a)$ transforme le cube Q en un point, origine des coordonnées de l'espace X. D'après $\varepsilon > 0$ donné, trouvons $\delta > 0$ tel que la quantité $o(h)$ dans (3) admette l'estimation $|o(h)| \leqslant \varepsilon\,|h|$ pour tout $|h| < \delta$. Ainsi, la distance entre le vecteur $x(a+h)$ et le point $x(a) = b$ ne dépasse pas, pour $|h| < \delta$, le nombre $\varepsilon\delta$. Par conséquent, l'image $x(B)$ du pavé B se trouve dans la boule de rayon $\varepsilon\delta$ centrée au point b. Donc,

$$\frac{|x(B)|}{|B|} \leqslant C\,\frac{(\varepsilon\delta)^n}{\delta^n} = C\varepsilon^n,$$

et, lorsque $B \to a$ ($\varepsilon \to 0$), la quantité $\frac{|x(B)|}{|B|}$ a 0 pour limite. Définitivement, *dans tous les cas possibles*, on a

$$\lim_{B \to a} \frac{|x(B)|}{|B|} = |\det x'(a)|,$$

et le théorème b est complètement démontré.

3.58. Changement de variables dans une intégrale multiple. Nous sommes maintenant en mesure,

en appliquant le théorème *3.45*, de formuler la règle fondamentale de changement des variables dans une intégrale multiple.

a. Théorème. *Soit* $x = x(u)$ *une application jordanienne* (*3.45*) *dérivable d'un corps jordanien compact* $U \subset R_n$ *sur un corps jordanien compact* $X \subset R_n$, *telle que l'intérieur du corps* U *s'applique sur l'intérieur du corps* X *et la frontière de* U *sur celle de* X. *Supposons qu'une fonction* $f(x)$ *soit continue sur l'ensemble* X. *Alors la formule suivante a lieu :*

$$\int_X f(x)\,dx = \int_U g(u) \det x'(u)\,du, \qquad (1)$$

où $g(u) = f(x(u))$.

Démonstration. Si U est un cube dans R_n, le résultat cherché découle des théorèmes *3.45* et *3.57b*, compte tenu du fait que, dans le cas en question, on peut choisir pour cellules une famille de cubes inclus dans U (*3.33g*). Si U est un ensemble élémentaire, i.e. une réunion finie de cubes sans points intérieurs communs, le résultat cherché se déduit en additionnant les égalités du type (1) écrites pour chacun de ces cubes.

Envisageons le cas général où U est un corps jordanien quelconque. Pour $\varepsilon > 0$ donné, trouvons un ensemble ouvert élémentaire $Q \subset R_n$ qui contienne la frontière Z du corps X strictement dans son intérieur et ait une mesure $\leqslant \varepsilon$. L'image réciproque complète $F \subset U$ de l'ensemble $X - QX$ est une partie fermée de U qui n'a aucun point commun avec la frontière W du corps U, de sorte que $\delta = d(F, W) > 0$ (*3.23d*). Ensuite, en nous servant du lemme *3.51j*, trouvons un ensemble élémentaire P de mesure $\leqslant \varepsilon$ qui contienne W strictement dans son intérieur et soit lui-même contenu dans le δ-voisinage de W. La différence $U - P$ est un ensemble élémentaire qui se transforme en un ensemble jordanien $Y \subset X$ contenant l'ensemble $X - Q$, de sorte que $m(X - Y) \leqslant mQ \leqslant \varepsilon$. On peut maintenant écrire :

$$\int_X f(x)\,dx = \int_Y f(x)\,dx + \int_{X-Y} f(x)\,dx,$$

$$\int_U g(u) \det x'(u)\,du = \int_{U-P} g(u) \det x'(u)\,du + \int_{P \cap U} g(u) \det x'(u)\,du.$$

Comme $U - P$ est un ensemble élémentaire, les faits démontrés impliquent :

$$\int_{U-P} g(u) \det x'(u)\,du = \int_Y f(x)\,dx.$$

Puisque $m(X - Y) \leqslant \varepsilon$, $mP \leqslant \varepsilon$ on a

$$\left| \int\limits_{X-Y} f(x)\, dx \right| \leqslant \max_X |f(x)|\, \varepsilon,$$

$$\left| \int\limits_{P \cap U} g(u) \det x'(u)\, du \right| \leqslant \max_X |f(x)| \cdot \max_U |\det x'(u)| \cdot \varepsilon.$$

Il en découle

$$\left| \int\limits_X f(x)\, dx - \int\limits_U g(u) \det x'(u)\, du \right| \leqslant \max_X |f(x)|\, (1 + \max_U |\det x'(u)|)\, \varepsilon,$$

et, $\varepsilon > 0$ étant arbitraire, on a l'égalité (1), ce qu'il fallait prouver.

b. Revenons à la formule *3.57* (4):

$$\lim_{B \to a} \frac{|x(B)|}{|B|} = |\det x'(a)|, \qquad (2)$$

qu'on a démontrée plus haut pour le cas des pavés cubiques B. Soit $B_1, B_2, \ldots, B_\nu, \ldots$ une suite quelconque de corps jordaniens de mesures positives $mB_\nu = |B_\nu|$ qui se contractent vers le point a. D'après ce qu'on a démontré dans a, nous avons

$$|x(B_\nu)| = \int\limits_{x(B_\nu)} dx = \int\limits_{B_\nu} |\det x'(u)|\, du;$$

donc, en raison de la continuité de $x'(u)$,

$$\frac{|x(B_\nu)|}{|B_\nu|} = \frac{1}{|B_\nu|} \int\limits_{B_\nu} |\det x'(u)|\, du \to |\det x'(a)|;$$

la formule (2) s'avère ainsi valable non seulement pour les cubes, mais pour n'importe quels corps jordaniens B qui se contractent vers le point a. Dans ce qui suit, nous écrirons $|A|$ à la place de $|\det A|$.

3.59. Exemples.

a. Aire d'une figure plane en coordonnées polaires. Soit $X \subset R_2 = \{x, y\}$ une figure plane (domaine jordanien). Son aire s'exprime par l'intégrale

$$S = |X| = \int\limits_X \int dx\, dy, \qquad (1)$$

Si l'on passe aux coordonnées polaires

$$x = r \cos \varphi, \quad y = r \sin \varphi,$$

on aura

$$\left| \frac{\partial(x, y)}{\partial(r, \varphi)} \right| = \begin{vmatrix} x_r & x_\varphi \\ y_r & y_\varphi \end{vmatrix} = \begin{vmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{vmatrix} = r.$$

Par conséquent, pour le domaine correspondant U dans le plan r, φ, on obtiendra :

$$S = \iint\limits_X dx\,dy = \iint\limits_U r\,dr\,d\varphi. \qquad (2)$$

Supposons que le domaine X soit limité par deux demi-droites $\varphi = \alpha$ et $\varphi = \beta$ et une courbe $r = r(\varphi)$ (fig. 3.5-12). Alors, en mettant l'intégrale (2) sous la forme d'une intégrale itérée, avec l'intégration intérieure par rapport à r, on trouve :

$$S = \iint\limits_U r\,dr\,d\varphi = \int\limits_{\varphi=\alpha}^{\beta} \left\{ \int\limits_{r=0}^{r(\varphi)} r\,dr \right\} d\varphi =$$

$$= \frac{1}{2} \int\limits_{\alpha}^{\beta} r^2(\varphi)\,d\varphi.$$

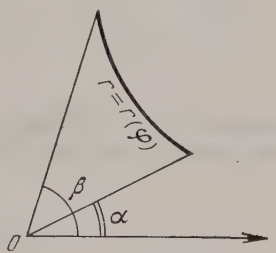


Fig. 3.5-12

C'est la formule qu'on a déduite directement (*U9.62*).

b. Aire d'une figure G limitée par une courbe $x^{2/\lambda} + y^{2/\mu} = 1$ $(x \geqslant 0,\ y \geqslant 0)$ et par les axes de coordonnées. On suppose que les exposants $2/\lambda$ et $2/\mu$ soient des nombres réels positifs.

Passons d'abord aux variables $x^{1/\lambda} = X$, $y^{1/\mu} = Y$ ayant en vue d'utiliser par la suite les coordonnées polaires. Le domaine correspondant Q dans le plan X, Y est découpé lui aussi par les axes de coordonnées et par un quart de la circonférence $X^2 + Y^2 = 1$. De plus, on a

$$\left| \frac{\partial(x, y)}{\partial(X, Y)} \right| = \begin{vmatrix} \lambda X^{\lambda-1} & \cdots & 0 \\ 0 & & \mu Y^{\mu-1} \end{vmatrix} = \lambda\mu X^{\lambda-1} Y^{\mu-1};$$

donc,

$$S = \iint\limits_G dx\,dy = \lambda\mu \iint\limits_Q X^{\lambda-1} Y^{\mu-1}\,dX\,dY.$$

En passant aux coordonnées polaires $X = r\cos\varphi$, $Y = r\sin\varphi$ et compte tenu de *U11.54*, on a :

$$S = \lambda\mu \int\limits_{\varphi=0}^{\pi/2} \int\limits_{r=0}^{1} r^{\lambda+\mu-1} \cos^{\lambda-1}\varphi \sin^{\mu-1}\varphi\,d\varphi\,dr =$$

$$= \frac{\lambda\mu}{\lambda+\mu} \int\limits_0^{\pi/2} \cos^{\lambda-1}\varphi \sin^{\mu-1}\varphi\,d\varphi = \frac{\lambda\mu}{\lambda+\mu} \frac{1}{2} \mathrm{B}\left(\frac{\lambda}{2}, \frac{\mu}{2}\right) =$$

$$= \frac{\lambda\mu}{2(\lambda+\mu)} \frac{\Gamma\left(\frac{\lambda}{2}\right)\Gamma\left(\frac{\mu}{2}\right)}{\Gamma\left(\frac{\lambda+\mu}{2}\right)},$$

où $\mathrm{B}(p, q)$ est la fonction bêta d'Euler, $\Gamma(s)$ la fonction gamma d'Euler (*U11.54*).

Cas particuliers (fig. 3.5-13):

$\lambda=\mu=1$ (arc de circonférence): $S=\frac{1}{4}\frac{\Gamma^2\left(\frac{1}{2}\right)}{\Gamma(1)}=\frac{\pi}{4}$;

$\lambda=\mu=2$ (segment de droite): $S=\frac{1}{2}\frac{\Gamma^2(1)}{\Gamma(2)}=\frac{1}{2}$;

$\lambda=\mu=3$ (arc d'astroïde): $S=\frac{3}{4}\frac{\Gamma^2\left(\frac{3}{2}\right)}{\Gamma(3)}=\frac{3\pi}{32}$.

c. Coordonnées sphériques dans R_3. Dans l'espace euclidien tridimensionnel R_3 avec x, y, z pour coordonnées rectangulaires, le système des coordonnées sphériques est formé des quantités suivantes (fig. 3.5-14):

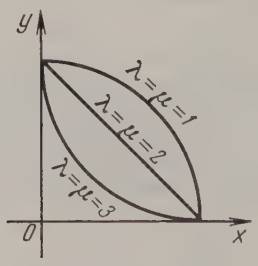


Fig. 3.5-13

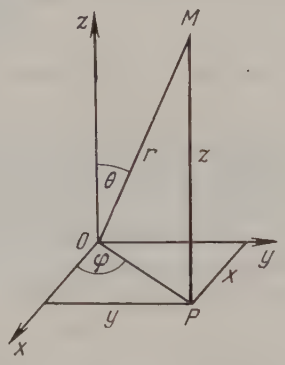


Fig. 3.5-14

la distance r du point O au point $M(x, y, z)$;
l'angle θ que la demi-droite OM fait avec le demi-axe positif des z;
l'angle φ que la projection OP de la demi-droite OM sur le plan x, y fait avec le demi-axe positif des x.
Il est évident que

$$\left.\begin{aligned} z&=r\cos\theta,\\ x&=r\sin\theta\cos\varphi,\\ y&=r\sin\theta\sin\varphi. \end{aligned}\right\} \tag{3}$$

De plus:

$$\left|\frac{\partial(x, y, z)}{\partial(r, \theta, \varphi)}\right|=\begin{vmatrix} \sin\theta\cos\varphi & r\cos\theta\cos\varphi & -r\sin\theta\sin\varphi\\ \sin\theta\sin\varphi & r\cos\theta\sin\varphi & r\sin\theta\cos\varphi\\ \cos\theta & -r\sin\theta & 0 \end{vmatrix}=r^2\sin\theta.$$

Par conséquent, pour deux domaines G et U correspondant l'un à l'autre, en coordonnées cartésiennes et sphériques respectivement on a

$$\iiint\limits_G f(x, y, z)\,dx\,dy\,dz=\iiint\limits_U f(x, y, z)\,r^2\sin\theta\,dr\,d\theta\,d\varphi,$$

où les variables x, y, z dans le second membre sont à remplacer par leurs expressions (3).

d. Calculer le volume d'un corps $G \subset R_3$ en sachant ses sections S_φ par les demi-plans $\varphi = \text{const} \in [0, 2\pi]$ (φ est la troisième coordonnée sphérique, *c*) (cf. fig. 3.5-15).

En coordonnées sphériques on a

$$|G| = \iiint_G dx\,dy\,dz = \iiint_U r^2 \sin\theta\, dr\, d\theta\, d\varphi = \int_{\varphi=0}^{2\pi} \left\{ \iint_{S_\varphi} r^2 \sin\theta\, dr\, d\theta \right\} d\varphi.$$

Introduisons dans le demi-plan S_φ les coordonnées cartésiennes $z = r\cos\theta$,

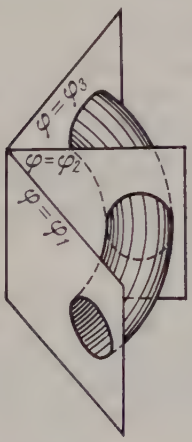


Fig. 3.5-15

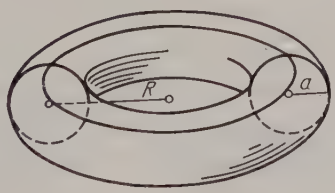


Fig. 3.5-16

$\rho = r\sin\theta$. Alors, pour le domaine correspondant G_φ dans le plan ρ, z, on a :

$$\iint_{S_\varphi} r^2 \sin\theta\, dr\, d\theta = \iint_{S_\varphi} r\sin\theta\, dr\, d\theta = \iint_{G_\varphi} \rho\, dz\, d\rho.$$

L'intégrale dans le membre droit est la coordonnée horizontale $\rho_C(\varphi)$ du centre de gravité de la figure C_φ (*3.54b*) multipliée par l'aire $|G_\varphi|$ de cette figure. Nous obtenons ainsi

$$|G| = \int_0^{2\pi} \rho_C(\varphi)\,|G_\varphi|\,d\varphi.$$

Notre problème sera résolu dès que nous sachions, pour tout $\varphi \in [0, 2\pi]$ l'aire de la figure G_φ et la distance $\rho_C(\varphi)$ de son centre de gravité à l'axe des z.

Pour le tore montré fig. 3.5-16 nous avons $|G_\varphi| = \pi a^2$, $\rho_C(\varphi) = R$; par conséquent,

$$|T| = 2\pi \cdot \pi a^2 R = 2\pi^2 a^2 R.$$

§ 3. 6. Intégrale de surface

3.61. Définition d'une intégrale de surface. **a.** Soit $x = \varphi(u)$ $(R_k \to R_n)$ une application définie et continûment dérivable dans un domaine $U \subset R_k$. Soit $G \subset U$ un ensemble jordanien fermé. L'ensemble $S = \varphi(G) \subset R_n$ sera appelé *surface k-dimensionnelle dans* R_n.

Nous tenons à définir *l'intégrale de surface*

$$\int_S f(x)\, dS \qquad (1)$$

d'une fonction $f(x)$ donnée et continue sur S.

L'intégrale correspondante sur un domaine plan S peut être interprétée comme masse totale de ce domaine, si $f(x)$ représente la densité de la masse au point x. D'une façon analogue, l'intégrale (1) doit avoir le sens de la masse de toute la surface S, si $f(x)$ est la densité. En particulier, pour $f(x) \equiv 1$, notre définition sera celle de l'*aire de la surface* S.

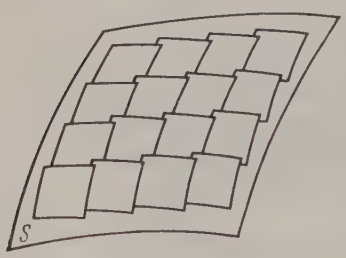


Fig. 3.6-1

Afin d'éviter la superposition des parties essentielles de la surface S, nous supposons que l'application $x = \varphi(u)$ soit biunivoque, à l'exception éventuelle d'un ensemble négligeable $Z \subset G$.

b. Passons aux définitions rigoureuses. Considérons une partition jordanienne quelconque $\Pi = \{E_i\}$ de l'ensemble G; soit $d(\Pi)$, comme d'ordinaire, le diamètre maximal des ensembles E_i. Dans tout ensemble E_i fixons un point quelconque ξ_i et considérons, dans les limites de l'ensemble E_i, l'application linéaire « tangente »

$$\xi_i + \Delta u \to \varphi(\xi_i) + \varphi'(\xi_i)\, \Delta u.$$

L'image E_i de l'ensemble E_i par cette application est un sous-ensemble jordanien de la variété linéaire k-dimensionnelle correspondante. Les ensembles E_i forment un recouvrement « en tuiles » de la surface S (fig. 3.6-1). En attribuant à chaque « tuile » E_i une « densité » uniforme $f(\varphi(\xi_i))$, nous voyons que le « poids » de tout le recouvrement « en tuiles » vaut

$$\sum_i f(\varphi(\xi_i))\, |\mathrm{E}_i|_k. \qquad (2)$$

A présent trouvons la valeur $|\,E_i\,|_k$. L'application $x = \varphi(u)$ écrite en coordonnées a la forme

$$\begin{aligned} x_1 &= \varphi_1(u_1, \ldots, u_k), \\ &\cdots\cdots\cdots\cdots \\ x_n &= \varphi_n(u_1, \ldots, u_k). \end{aligned}$$

La dérivée $\varphi'(u)$ est un opérateur linéaire $(R_k \to R_n)$ avec la matrice

$$\varphi'(u) \cong \left\| \begin{matrix} \frac{\partial \varphi_1}{\partial u_1} & \cdots & \frac{\partial \varphi_1}{\partial u_k} \\ \cdots\cdots\cdots & & \\ \frac{\partial \varphi_n}{\partial u_1}, & \cdots & \frac{\partial \varphi_n}{\partial u_k} \end{matrix} \right\|.$$

Soit d'abord E_i un parallélépipède rectangle B_i d'arêtes

$$\left.\begin{aligned} g_1 &= (\Delta u_1, 0, \ldots, 0), \\ &\cdots\cdots\cdots\cdots \\ g_k &= (0, 0, \ldots, \Delta u_k). \end{aligned}\right\}$$

Alors $\varphi'(\xi_i)\, B_i = E_i$ est le parallélépipède construit sur les vecteurs

$$\begin{aligned} \varphi'(\xi_i)\, g_1 &= \left(\frac{\partial \varphi_1}{\partial u_1}, \ldots, \frac{\partial \varphi_n}{\partial u_1}\right) \Delta u_1, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ \varphi'(\xi_i)\, g_k &= \left(\frac{\partial \varphi_1}{\partial u_k}, \ldots, \frac{\partial \varphi_n}{\partial u_k}\right) \Delta u_k. \end{aligned}$$

Son volume k-dimensionnel $|\,E_i\,|_k$ vaut, conformément *3.55*, la racine carrée de la somme des carrés de tous les mineurs d'ordre k de la matrice $\varphi'(u)$ multipliée par $\Delta u_1 \ldots \Delta u_k$. D'après *3.55d*, désignons cette quantité par

$$|\,E_i\,|_k = \left| \left[\frac{\partial \varphi(\xi_i)}{\partial u_1}, \ldots, \frac{\partial \varphi(\xi_i)}{\partial u_k} \right] \right| \Delta u_1 \ldots \Delta u_k. \tag{3}$$

N'importe quel ensemble jordanien E, selon *3.31c*, est formé d'une certaine façon des cellules B_j ; la quantité $\varphi'(\xi_i)$ est constante pour i fixe ; c'est pourquoi, en effectuant le procédé correspondant pour E_i, nous voyons que la formule (3) reste valable pour n'importe quels ensembles jordaniens :

$$|\,E_i\,|_k = \left| \left[\frac{\partial \varphi(\xi_i)}{\partial u_1}, \ldots, \frac{\partial \varphi(\xi_i)}{\partial u_k} \right] \right| |\,E_i\,|. \tag{4}$$

La somme (2) se met donc sous la forme

$$\sum_i f(\varphi(\xi_i)) \left| \left[\frac{\partial \varphi(\xi_i)}{\partial u_1}, \ldots, \frac{\partial \varphi(\xi_i)}{\partial u_k} \right] \right| |\,E_i\,|;$$

elle représente une somme intégrale pour l'intégrale

$$\int_G f(\varphi(u)) \left| \left[\frac{\partial \varphi(u)}{\partial u_1}, \ldots, \frac{\partial \varphi(u)}{\partial u_k} \right] \right| du. \tag{5}$$

Comme $f(\varphi(u))$ et $\left|\left[\frac{\partial\varphi(u)}{\partial u_1}, \ldots, \frac{\partial\varphi(u)}{\partial u_k}\right]\right|$ sont continus sur G, les sommes intégrales (4) tendent vers l'intégrale (5) lorsque $d(\Pi)\to 0$. Nous sommes ainsi amenés à la définition :

$$\int_S f(x)\,dS = \int_G f(x(u))\left|\left[\frac{\partial\varphi(u)}{\partial u_1}, \ldots, \frac{\partial\varphi(u)}{\partial u_k}\right]\right| du. \tag{6}$$

En particulier, nous appelons *mesure* k-dimensionnelle ou, tout simplement, *aire* de la surface S l'intégrale (6) avec $f(x) \equiv 1$, c'est-à-dire la quantité

$$|S|_k = \int_S dS = \int_G \left|\left[\frac{\partial\varphi(u)}{\partial u_1}, \ldots, \frac{\partial\varphi(u)}{\partial u_k}\right]\right| du. \tag{7}$$

c. Il faut prouver que la définition ci-dessus soit correcte, i.e. ne dépend pas du choix des paramètres déterminant la surface S. Supposons que, à côté de la représentation $x = \varphi(u)$, $u \in G$, il y ait une autre représentation $x = \psi(v)$ de la même surface S, le paramètre v parcourant un domaine $V \subset R_k$, et les paramètres u, v étant liés par les relations mutuellement inverses

$$u = u(v), \quad v = v(u) \tag{8}$$

avec les fonctions $u(v)$ et $v(u)$ dérivables. De telles représentations $x = \varphi(u)$ et $x = \psi(v)$ seront dites *équivalentes*.

D'après la définition, on a

$$\int_S f(x)\,dS = \int_G f(\varphi(u))\left|\left[\frac{\partial\varphi(u)}{\partial u_1}, \ldots, \frac{\partial\varphi(u)}{\partial u_k}\right]\right| du. \tag{9}$$

Effectuons dans le second membre de (9) la substitution (8) d'après la règle *3.58a*. Nous obtenons :

$$\int_S f(x)\,dS = \int_V f(\psi(v))\left|\left[\frac{\partial\varphi(u)}{\partial u_1}, \ldots, \frac{\partial\varphi(u)}{\partial u_k}\right]\right|\left|\frac{\partial(u_1, \ldots, u_k)}{\partial(v_1, \ldots, v_k)}\right| dv.$$

Or, en vertu de *1.33b*,

$$\left|\left[\frac{\partial\varphi(u)}{\partial u_1}, \ldots, \frac{\partial\varphi(u)}{\partial u_k}\right]\right|\left|\frac{\partial(u_1, \ldots, u_k)}{\partial(v_1, \ldots, v_k)}\right| =$$

$$= \sqrt{\sum_{(i)} \det^2 \frac{\partial(x_{i_1}, \ldots, x_{i_k})}{\partial(u_1, \ldots, u_k)} \det^2 \frac{\partial(u_1, \ldots, u_k)}{\partial(v_1, \ldots, v_k)}} =$$

$$= \sqrt{\sum_{(i)} \det^2 \frac{\partial(x_{i_1}, \ldots, x_{i_k})}{\partial(v_1, \ldots, v_k)}} = \left|\left[\frac{\partial\psi(v)}{\partial v_1}, \ldots, \frac{\partial\psi(v)}{\partial v_k}\right]\right|.$$

Il en résulte :

$$\int_G f(\varphi(u)) \left| \left[\frac{\partial \varphi(u)}{\partial u_1}, \ldots, \frac{\partial \varphi(u)}{\partial u_k} \right] \right| du =$$

$$= \int_V f(\psi(v)) \left| \left[\frac{\partial \psi(v)}{\partial v_1}, \ldots, \frac{\partial \psi(v)}{\partial v_k} \right] \right| dv,$$

ce qui démontre l'invariance de la valeur de l'intégrale dans le premier membre par rapport aux représentations équivalentes de la surface S.

3.62. Cas particuliers.

a. Soit $k = 1$; on a affaire à une courbe $L = \{x = x(t)\ (R_1 \to R_n),\ a \leqslant t \leqslant b\}$ dans l'espace n-dimensionnel R_n. La matrice $x'(t)$ n'a qu'une seule colonne $\{x_1'(t), \ldots, x_n'(t)\}$ et

$$|[x'(t)]| = |x'(t)| = \sqrt{\sum_{i=1}^{n} [x_i'(t)]^2}.$$

Les formules *3.61* (6) et (7) prennent les formes respectives

$$\int_L f(x)\, dL = \int_a^b f(x(t)) \sqrt{\sum_{i=1}^{n} [x_i'(t)]^2}\, dt, \tag{1}$$

$$|L| = \int_L dL = \int_a^b \sqrt{\sum_{i=1}^{n} [x_i'(t)]^2}\, dt. \tag{2}$$

Les deux formules, l'une pour l'intégrale de la fonction $f(x)$ le long de la courbe L et l'autre pour la longueur de la courbe L, nous sont familières depuis *U9.91* ($g = s$) et *U9.63*.

b. Si $k = 2$, on a affaire à une surface bidimensionnelle

$$S = \{x = \varphi(u, v)\ (R_2 \to R_n),\ (u, v) \in G \subset R_2\}.$$

La matrice $x'(u, v)$ a en l'occurrence deux colonnes

$$\frac{\partial x}{\partial u} = \left\{ \frac{\partial x_1}{\partial u}, \ldots, \frac{\partial x_n}{\partial u} \right\} \quad \text{et} \quad \frac{\partial x}{\partial v} = \left\{ \frac{\partial x_1}{\partial v}, \ldots, \frac{\partial x_n}{\partial v} \right\}. \tag{3}$$

La quantité $\left| \left[\frac{\partial x}{\partial u}, \frac{\partial x}{\partial v} \right] \right|$ est la racine carrée de la somme de $\frac{n(n+1)}{2}$ carrés des mineurs d'ordre 2 de la matrice $x'(u, v)$. D'ailleurs, l'expression $\left| \left[\frac{\partial x}{\partial u}, \frac{\partial x}{\partial v} \right] \right|$ qui représente l'aire du parallélogramme construit sur les vecteurs (3) peut être mise sous une

autre forme: si ω est l'angle des vecteurs $\frac{\partial x}{\partial u}$ et $\frac{\partial x}{\partial v}$, on a

$$\left|\left[\frac{\partial x}{\partial u}, \frac{\partial x}{\partial v}\right]\right| = \left|\frac{\partial x}{\partial u}\right| \left|\frac{\partial x}{\partial v}\right| \sin\omega = \left|\frac{\partial x}{\partial u}\right| \left|\frac{\partial x}{\partial v}\right| \sqrt{1-\cos^2\omega} =$$

$$= \sqrt{\left|\frac{\partial x}{\partial u}\right|^2 \left|\frac{\partial x}{\partial v}\right|^2 - \left(\frac{\partial x}{\partial u}, \frac{\partial x}{\partial v}\right)^2} = \sqrt{EG-F^2},$$

où

$$E = \left|\frac{\partial x}{\partial u}\right|^2 = \sum_{i=1}^{n} \left(\frac{\partial x_i}{\partial u}\right)^2,$$

$$F = \left(\frac{\partial x}{\partial u}, \frac{\partial x}{\partial v}\right) = \sum_{i=1}^{n} \frac{\partial x_i}{\partial u} \frac{\partial x_i}{\partial v},$$

$$G = \left|\frac{\partial x}{\partial v}\right|^2 = \sum_{i=1}^{n} \left(\frac{\partial x_i}{\partial v}\right)^2.$$

Ainsi, l'intégrale *3.61* (6) s'écrit, pour $k=2$, comme suit:

$$\int_S f(x)\, dS = \int_G f(x(u,v)) \sqrt{EG-F^2}\, du\, dv, \tag{4}$$

et l'aire de la surface S (*3.61* (7)) comme suit:

$$|S| = \int_S dS = \int_G \sqrt{EG-F^2}\, du\, dv. \tag{5}$$

c. Si $k=n-1$, on a affaire à une surface $(n-1)$-dimensionnelle $S=\{x=\varphi(u_1, \ldots, u_{n-1})\,(R_{n-1} \to R_n),\ (u_1, \ldots, u_{n-1}) \in G \subset \subset R_{n-1}\}$. La matrice $\varphi'(u)$ a $n-1$ colonnes et n lignes:

$$\varphi'(u) \cong \left\| \begin{matrix} \frac{\partial \varphi_1(u)}{\partial u_1} & \cdots & \frac{\partial \varphi_1(u)}{\partial u_{n-1}} \\ \cdots & \cdots & \cdots \\ \frac{\partial \varphi_n(u)}{\partial u_1} & \cdots & \frac{\partial \varphi_n(u)}{\partial u_{n-1}} \end{matrix} \right\|.$$

Le vecteur

$$N = \begin{vmatrix} e_1 & \cdots & e_n \\ \frac{\partial \varphi_1(u)}{\partial u_1} & \cdots & \frac{\partial \varphi_n(u)}{\partial u_1} \\ \cdots & \cdots & \cdots \\ \frac{\partial \varphi_1(u)}{\partial u_{n-1}} & \cdots & \frac{\partial \varphi_n(u)}{\partial u_{n-1}} \end{vmatrix} = \left[\frac{\partial \varphi}{\partial u_1}, \ldots, \frac{\partial \varphi}{\partial u_{n-1}}\right]$$

s'appelle *produit vectoriel des vecteurs* $\frac{\partial \varphi}{\partial u_1}, \ldots, \frac{\partial \varphi}{\partial u_{n-1}}$. Il est, évidemment, normal à tout vecteur $\frac{\partial \varphi(u)}{\partial u_i}$ (i.e. normal dans R_n

à la surface même S), et sa longueur est juste la quantité $\left|\left[\frac{\partial\varphi}{\partial u_1}, \ldots, \frac{\partial\varphi}{\partial u_{n-1}}\right]\right|$. Ainsi, l'intégrale de la fonction $f(x)$ sur la surface est dans ce cas calculée selon la formule

$$\int_S f(x)\,dS = \int_G f(\varphi(u))\left|\left[\frac{\partial\varphi}{\partial u_1}, \ldots, \frac{\partial\varphi}{\partial u_{n-1}}\right]\right| du = \int_G f(\varphi(u))\,|N|\,du. \tag{6}$$

Comme d'habitude, l'aire de la surface S s'en déduit en posant $f \equiv 1$:

$$|S| = \int_S dS = \int_G |N|\,du = \int_G \left|\left[\frac{\partial\varphi}{\partial u_1}, \ldots, \frac{\partial\varphi}{\partial u_{n-1}}\right]\right| du. \tag{7}$$

d. Signalons un cas bien commun où la surface S est donnée par une équation de la forme

$$x_n = \varphi(x_1, \ldots, x_{n-1}), \quad (x_1, \ldots, x_{n-1}) \in Q \subset R_{n-1}$$

ou, ce qui revient au même, par un système de la forme

$$\begin{aligned} x_1 &= x_1, \\ x_2 &= x_2, \\ &\cdots\cdots\cdots\cdots \\ x_{n-1} &= x_{n-1}, \\ x_n &= \varphi(x_1, \ldots, x_{n-1}). \end{aligned}$$

On a alors

$$N = \begin{vmatrix} e_1 & e_2 & \ldots & e_{n-1} & e_n \\ 1 & 0 & \ldots & 0 & \frac{\partial\varphi}{\partial x_1} \\ 0 & 1 & \ldots & 0 & \frac{\partial\varphi}{\partial x_2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \ldots & 1 & \frac{\partial\varphi}{\partial x_{n-1}} \end{vmatrix} = (-1)^n \left(e_1 \frac{\partial\varphi}{\partial x_1} + \ldots + e_{n-1}\frac{\partial\varphi}{\partial x_{n-1}} - e_n\right),$$

d'où

$$|N| = \sqrt{\left(\frac{\partial\varphi}{\partial x_1}\right)^2 + \ldots + \left(\frac{\partial\varphi}{\partial x_{n-1}}\right)^2 + 1}.$$

Pour l'intégrale *3.61* (6) de la fonction $f(x)$ sur la surface S, on obtient l'expression

$$\int_S f(x)\,dS = \int_Q f(x_1, \ldots, x_{n-1}, \varphi(x_1, \ldots, x_{n-1})) \times$$

$$\times \sqrt{\left(\frac{\partial\varphi}{\partial x_1}\right)^2 + \ldots + \left(\frac{\partial\varphi}{\partial x_{n-1}}\right)^2 + 1}\, dx_1 \ldots dx_{n-1}. \tag{8}$$

e. Soit, enfin, $k = n$, de sorte que le rôle de la surface S soit joué par un domaine dans l'espace n-dimensionnel. La matrice $\varphi'(u)$ a n lignes et n colonnes; le volume du parallélépipède construit sur les vecteurs $\frac{\partial \varphi}{\partial u_1}, \ldots, \frac{\partial \varphi}{\partial u_n}$ est égal au module de son déterminant. Les formules *3.61* (6), (7)

$$\int_S f(x)\,dS = \int_G f(x(u)) \left| \frac{\partial(\varphi_1, \ldots, \varphi_n)}{\partial(u_1, \ldots, u_n)} \right| du,$$

$$|S| = \int_S dS = \int_G \left| \frac{\partial(\varphi_1, \ldots, \varphi_n)}{\partial(u_1, \ldots, u_n)} \right| du$$

se transforment en les formules usuelles de changement des variables dans l'intégrale n-uple *3.58* (1).

3.63. Exemples.

a. Trouvons dans R_3 l'aire de la partie de la sphère de rayon R qui est découpée par deux méridiennes $\varphi_1 \leqslant \varphi \leqslant \varphi_2$ et deux parallèles $\theta_1 \leqslant \theta \leqslant \theta_2$ (fig. 3.6-2).

En coordonnées sphériques $x = R \sin\theta \cos\varphi$, $y = R \sin\theta \sin\varphi$, $z = R \cos\theta$ (*3.59c*) nous avons

$$\left\| \begin{matrix} \frac{\partial x}{\partial \varphi} & \frac{\partial y}{\partial \varphi} & \frac{\partial z}{\partial \varphi} \\ \frac{\partial x}{\partial \theta} & \frac{\partial y}{\partial \theta} & \frac{\partial z}{\partial \theta} \end{matrix} \right\| = \left\| \begin{matrix} -R\sin\theta\sin\varphi & R\sin\theta\cos\varphi & 0 \\ R\cos\theta\cos\varphi & R\cos\theta\sin\varphi & -R\sin\theta \end{matrix} \right\|,$$

de sorte que

$$E = \left(\frac{\partial x}{\partial \varphi}\right)^2 + \left(\frac{\partial y}{\partial \varphi}\right)^2 + \left(\frac{\partial z}{\partial \varphi}\right)^2 = R^2 \sin^2\theta,$$

$$F = \frac{\partial x}{\partial \varphi}\frac{\partial x}{\partial \theta} + \frac{\partial y}{\partial \varphi}\frac{\partial y}{\partial \theta} + \frac{\partial z}{\partial \varphi}\frac{\partial z}{\partial \theta} = 0,$$

$$G = \left(\frac{\partial x}{\partial \theta}\right)^2 + \left(\frac{\partial y}{\partial \theta}\right)^2 + \left(\frac{\partial z}{\partial \theta}\right)^2 = R^2.$$

Il en résulte:

$$|S| = \int_{\varphi_1}^{\varphi_2}\int_{\theta_1}^{\theta_2} \sqrt{EG - F^2}\,d\varphi\,d\theta = R^2 \int_{\varphi_1}^{\varphi_2}\int_{\theta_1}^{\theta_2} \sin\theta\,d\varphi\,d\theta =$$

$$= R^2(\varphi_2 - \varphi_1)(\cos\theta_1 - \cos\theta_2).$$

En particulier, pour toute la sphère $0 \leqslant \varphi \leqslant 2\pi$, $0 \leqslant \theta \leqslant \pi$, nous obtenons

$$|S| = R^2 \cdot 2\pi \cdot 2 = 4\pi R^2.$$

b. Calculons l'aire du tore produit par la rotation autour de l'axe des z de la circonférence de rayon a qui se trouve au début dans le plan x, z où elle a pour centre le point $x = b$ $(b > a)$, $y = 0$, $z = 0$.

Une représentation paramétrique du tore est immédiate (voir fig. 3.6-3):

$$x = (b + a\cos\psi)\cos\varphi, \quad y = (b + a\cos\psi)\sin\varphi, \quad z = a\sin\psi;$$

les angles φ et ψ variant de 0 à 2π. Vu que

$$\left\| \begin{matrix} \frac{\partial x}{\partial \varphi} & \frac{\partial y}{\partial \varphi} & \frac{\partial z}{\partial \varphi} \\ \frac{\partial x}{\partial \psi} & \frac{\partial y}{\partial \psi} & \frac{\partial z}{\partial \psi} \end{matrix} \right\| = \left\| \begin{matrix} -(b + a\cos\psi)\sin\varphi & (b + a\cos\psi)\cos\varphi & 0 \\ -a\sin\psi\cos\varphi & -a\sin\psi\sin\varphi & a\cos\psi \end{matrix} \right\|,$$

on trouve

$$E = (b + a\cos\psi)^2, \quad F = 0, \quad G = a^2.$$

Définitivement :

$$|S| = \int_0^{2\pi}\int_0^{2\pi} \sqrt{EG - F^2}\, d\varphi\, d\psi = a\int_0^{2\pi}\int_0^{2\pi} (b + a\cos\psi)\, d\varphi\, d\psi = 4\pi^2 ab.$$

c. Aires de surfaces homothétiques. Deux surfaces k-dimensionnelles S_1 et S_2 que l'on puisse donner par les applications dérivables respectives

$$x = \varphi(u),$$
$$x = b\varphi(u)$$

($x \in R_n$, $b > 0$), la variable u parcourant, comme d'habitude, un domaine

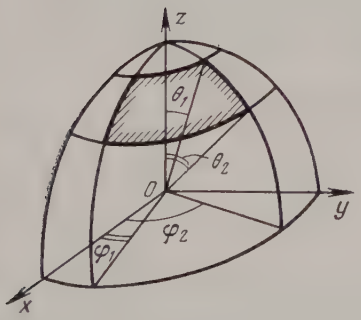


Fig. 3.6-2

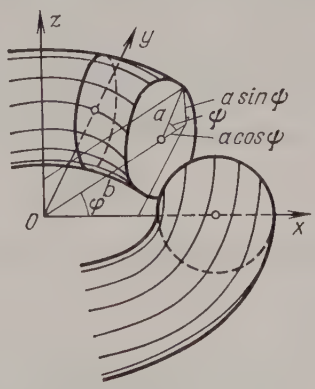


Fig. 3.6-3

jordanien $G \subset R_k$, sont dites *homothétiques, de rapport d'homothétie* b. Trouvons la relation entre leurs aires. Ces dernières s'expriment d'après la formule *3.61* (7)

$$|S_1| = \int_G \left| \left[\frac{\partial \varphi}{\partial u_1}, \ldots, \frac{\partial \varphi}{\partial u_n} \right] \right| du, \tag{1}$$

$$|S_2| = \int_G \left| \left[b\frac{\partial \varphi}{\partial u_1}, \ldots, b\frac{\partial \varphi}{\partial u_k} \right] \right| du. \tag{2}$$

En mettant le coefficient b en facteur pour chacune de k colonnes des mineurs figurant sous le signe somme dans (2), nous voyons que

$$| S_2 | = b^k | S_1 |.$$

Ainsi, *le rapport des aires de deux surfaces k-dimensionnelles homothétiques est égal à la puissance k-ème de leur rapport d'homothétie.*

3.64. Aire d'une surface comme limite des aires de surfaces polyèdres inscrites.

a. Nous avons, à son temps, défini la longueur d'une courbe comme limite des longueurs de lignes polygonales inscrites. Il s'avère que l'aire d'une surface peut être définie d'une manière analogue, comme limite des aires de surfaces polyèdres inscrites lorsque les faces de ces dernières son indéfiniment morcelées. Mais ce n'est pas toute suite de surfaces polyèdres inscrites dans une surface donnée qui convient à obtenir l'aire de cette dernière par le passage à la limite lorsque les faces sont indéfiniment morcelées (cf. exercice 5). Nous allons indiquer certains types de suites de surfaces polyèdres inscrites dont les aires ont pour limite celle de la surface donnée.

Soit $S = \{x = \varphi(u),\ u \in U \subset R_k\}$ une surface k-dimensionnelle dans l'espace n-dimensionnel. On suppose que l'application $\varphi(u)$ possède dans le domaine fermé U la dérivée $\varphi'(u)$ continue et non dégénérée, de sorte que la somme des carrés de tous les mineurs d'ordre k de la matrice $\varphi'(u)$ soit supérieure ou égale à un nombre positif c^2. Cette dernière hypothèse s'exprime également comme suit : si $g_1, \ldots, g_k$ sont les vecteurs de base dans l'espace R_k, alors, dans tout le domaine U a lieu l'inégalité

$$\left| \left[\frac{\partial \varphi}{\partial u_1}, \ldots, \frac{\partial \varphi}{\partial u_k} \right] \right| \geqslant c ;$$

ou encore : pour tout $\delta > 0$, on a

$$| [\varphi'(u)\, \delta g_1, \ldots, \varphi'(u)\, \delta g_k] | \geqslant c\delta^k. \qquad (1)$$

b. Soient $Q = \{\alpha_1 \leqslant u_1 \leqslant \beta_1, \ldots, \alpha_k \leqslant u_k \leqslant \beta_k\}$ un cube k-dimensionnel inclus dans le domaine U, et $\beta_1 - \alpha_1 = \ldots = \beta_k - \alpha_k = h$. On peut partager le cube en $k!$ simplexes k-dimensionnels d'une même aire (*3.56a*) d'après la règle suivante. Soient $g_1, \ldots, g_k$ les arêtes du cube Q issues d'un sommet fixe P, par exemple du sommet $u_1 = \alpha_1, \ldots, u_k = \alpha_k$. Soit $i_1, \ldots, i_k$ une permutation quelconque des nombres $1, \ldots, k$. Considérons les vecteurs

$$\lambda_1 g_{i_1} + \lambda_2 (g_{i_1} + g_{i_2}) + \ldots + \lambda_k (g_{i_1} + \ldots + g_{i_k}),$$

$$\lambda_1 \geqslant 0, \ldots, \lambda_k \geqslant 0,\ \lambda_1 + \ldots + \lambda_k \leqslant 1.$$

Ils remplissent un simplexe $Q_{i_1 \ldots i_k}$ de sommets aux points P, $P + g_{i_1}$, $P + g_{i_1} + g_{i_2}$, $\ldots$, $P + g_{i_1} + \ldots + g_{i_k}$. Le volume

de ce simplexe vaut, d'après *3.56a*,

$$\frac{1}{k!}|[g_{i_1}, g_{i_1}+g_{i_2}, \ldots, g_{i_1}+\ldots+g_{i_k}]| =$$

$$= \frac{1}{k!}|[g_{i_1}, \ldots, g_{i_k}]| = \frac{1}{k!}|[g_1, \ldots, g_k]| = \frac{h^k}{k!},$$

de sorte que tous les simplexes $Q_{i_1 \ldots i_k}$ ont une même aire. Montrons que chaque point du cube Q appartient à un au moins de ces simplexes. Pour un point donné $x \in Q$ déterminons les indices $i_1, \ldots, i_k$ de façon à ce que, dans la représentation

$$x - P = c_1 g_{i_1} + \ldots + c_k g_{i_k}, \quad c_i \geqslant 0, \quad \sum_{i=1}^{n} c_i \leqslant 1, \tag{2}$$

les nombres c_i ne croissent pas : $c_1 \geqslant c_2 \geqslant \ldots \geqslant c_k$. Nous affirmons alors que le point $x - P$ appartient au simplexe $Q_{i_1 \ldots i_k}$, c'est-à-dire qu'a lieu la représentation

$$x - P = \lambda_1 g_{i_1} + \lambda_2 (g_{i_1} + g_{i_2}) + \ldots + \lambda_k (g_{i_1} + \ldots + g_{i_k}), \tag{3}$$

où $\lambda_i \geqslant 0$ $(i = 1, \ldots, k)$, $\lambda_1 + \ldots + \lambda_k \leqslant 1$. En effet, on peut mettre (3) sous la forme

$$x - P = (\lambda_1 + \ldots + \lambda_k) g_{i_1} + (\lambda_2 + \ldots + \lambda_k) g_{i_2} + \ldots + \lambda_k e_{i_k};$$

en comparant à (2) nous voyons qu'on peut poser

$$c_1 = \lambda_1 + \ldots + \lambda_k, \; c_2 = \lambda_2 + \ldots + \lambda_k, \; \ldots, \; c_k = \lambda_k.$$

Nous en tirons

$$\lambda_1 = c_1 - c_2 \geqslant 0, \; \lambda_2 = c_2 - c_3 \geqslant 0, \; \ldots, \; \lambda_k = c_k \geqslant 0$$

avec

$$\lambda_1 + \ldots + \lambda_k = c_1 \leqslant 1,$$

de sorte que $x - P$ appartient effectivement à $Q_{i_1 \ldots i_k}$. En particulier, nous voyons que les simplexes $Q_{i_1 \ldots i_k}$ n'ont deux à deux aucun point intérieur commun, puisque la somme de leurs volumes est h^k, c'est-à-dire vaut le volume de tout le cube Q.

c. Considérons maintenant le simplexe $Q_{i_1 \ldots i_k}$. Les images de ses sommets P, $P + g_{i_1}$, $\ldots$, $P + g_{i_1} + \ldots + g_{i_k}$ par l'application $\varphi(u)$ représentent $k+1$ points sur la surface S que nous désignons $\varphi(P)$, $\varphi(P_{i_1})$, $\ldots$, $\varphi(P_{i_1 \ldots i_k})$ respectivement. Ces $k+1$ points déterminent, à leur tour, un simplexe $\varphi(Q_{i_1 \ldots i_k})$ dans l'espace R_n qui est inscrit dans la surface S à ce sens que tous ses sommets appartiennent à la surface. On peut donc inscrire dans la surface S $k!$ simplexes $\varphi(Q_{i_1 \ldots i_k})$. Si l'on considère maintenant une

partition Π de l'espace R_k en cubes $Q^{(j)}$ d'arête h, on en sépare ceux qui sont contenus entièrement dans le domaine U, on partage chacun de ces derniers cubes en $k!$ simplexes $Q^{(j)}_{i_1 \ldots i_k}$ d'une même aire et on construit les simplexes correspondants $\varphi(Q^{(j)}_{i_1 \ldots i_k})$, alors cette construction, prise en entier, fournit une surface polyèdre Π_h dont tous les sommets appartiennent à la surface S, i.e. une surface inscrite dans S.

d. Calculons l'aire de la surface polyèdre Π_h. Le volume du simplexe $\varphi(Q^{(j)}_{i_1 \ldots i_k})$ de sommets $\varphi(P)$, $\varphi(P_{i_1})$, $\ldots$, $\varphi(P_{i_1 \ldots i_k})$ est

$$\frac{1}{k!} \left| [\varphi(P+g_{i_1}) - \varphi(P), \quad \varphi(P+g_{i_1}+g_{i_2}) - \varphi(P), \ldots \right.$$
$$\left. \ldots, \varphi(P+g_{i_1}+\ldots+g_{i_k}) - \varphi(P)] \right|. \qquad (4)$$

Pour $\varepsilon > 0$ donné, trouvons $\delta > 0$ tel que l'inégalité

$$|\varphi'(u') - \varphi'(u'')| \leqslant \varepsilon$$

ait lieu dès que $|u' - u''| \leqslant \delta$. Si les diamètres des cubes $Q^{(j)}$ ne dépassent pas δ, alors, d'après le théorème de la moyenne *1.42* (6), on a :

$$\varphi(P_j + g^{(j)}_{i_1}) - \varphi(P_j) = \varphi'(P_j)\, g^{(j)}_{i_1} + \varepsilon^{(j)}_1 \delta,$$
$$\varphi(P_j + g^{(j)}_{i_1} + g^{(j)}_{i_2}) - \varphi(P_j) = \varphi'(P_j)\,(g^{(j)}_{i_1} + g^{(j)}_{i_2}) + \varepsilon^{(j)}_2 \delta,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\varphi(P_j + g^{(j)}_{i_1} + \ldots + g^{(j)}_{i_k}) - \varphi(P_j) =$$
$$= \varphi'(P_j)\,(g^{(j)}_{i_1} + \ldots + g^{(j)}_{i_k}) + \varepsilon^{(j)}_k \delta,$$

où $|\varepsilon^{(j)}_i| \leqslant \varepsilon$, $i = 1, \ldots, k$. En portant ces expressions dans (4) et appliquant *3.55e,f*, nous obtenons

$$|\varphi(Q^{(j)}_{i_1 \ldots i_k})|^2 = \frac{1}{(k!)^2} |[\varphi'(P_j)\, g^{(j)}_{i_1} + \varepsilon^{(j)}_1 \delta, \ldots$$
$$\ldots, \varphi'(P_j)\,(g^{(j)}_{i_1} + \ldots + g^{(j)}_{i_k}) + \varepsilon^{(j)}_k \delta]|^2 =$$
$$= \frac{1}{(k!)^2} |[\varphi'(P_j)\, g^{(j)}_{i_1} + \varepsilon^{(j)}_1 \delta, \ldots, \varphi'(P_j)\, g^{(j)}_{i_k} + \varepsilon^{(j)}_k \delta]|^2 =$$
$$= \frac{1}{(k!)^2} |[\varphi'(P_j)\, g^{(j)}_{i_1}, \ldots, \varphi'(P_j)\, g^{(j)}_{i_k}]|^2 +$$
$$+ \frac{C_n}{(k!)^2} \varepsilon\delta\,(M + \varepsilon\delta)^{2k-1}\theta, \qquad |\theta| \leqslant 1,$$

où

$$M = \sup_{i,j} |\varphi'(P_j)\, g^{(j)}_{i_1}| \leqslant D\delta, \qquad D = \sup \|\varphi'(u)\|.$$

L'expression $C_n\varepsilon\delta\,(M+\varepsilon\delta)^{2k-1}\theta$ peut être écrite comme suit :

$$C_n\varepsilon\delta^{2k}\,(D+\varepsilon)^{2k-1}\theta = C'_n\varepsilon\delta^{2k}\theta_1, \qquad |\theta_1|\leqslant|\theta|\leqslant 1.$$

En permutant les vecteurs dans le premier terme et le mettant en facteur nous trouvons

$$|\varphi(Q^{(j)}_{i_1\ldots i_k})|^2 = \frac{1}{(k!)^2}\,|[\varphi'(P_j)\,g^{(j)}_1,\ \ldots,\ \varphi'(P_j)\,g^{(j)}_k]|^2 \times$$

$$\times\left(1+\frac{C'_n\varepsilon\delta^{2k}\theta_1}{|[\varphi'(P_j)\,g^{(j)}_1,\ \ldots,\ \varphi'(P_j)\,g^{(j)}_k]|^2}\right),$$

ce qui implique, vu que l'opérateur $\varphi'(u)$ (1) n'est pas dégénéré,

$$|\varphi(Q^{(j)}_{i_1\ldots i_k})|^2 = \frac{1}{(k!)^2}\,|[\varphi'(P_j)\,g^{(j)}_1,\ \ldots$$

$$\ldots,\ \varphi'(P_j)\,g^{(j)}_k]|^2\left(1+\frac{C'_n\varepsilon\theta_1}{c^2}\right),$$

soit

$$|\varphi(Q^{(j)}_{i_1\ldots i_k})| = \frac{1}{k!}\,|[\varphi'(P_j)\,g^{(j)}_1,\ \ldots,\ \varphi'(P_j)\,g^{(j)}_k]|\sqrt{1+\frac{C''_n\,\varepsilon\,\theta_1}{c^2}}.$$

Etant données les inégalités $\sqrt{1+\mu}\leqslant 1+\frac{\mu}{2}$ et $|[\varphi(P_j)\,g^{(j)}_1,\ \ldots$ $\ldots,\ \varphi(P_j)\,g^{(j)}_k]|\leqslant D^k\delta^k = D^k\,|\,Q^{(j)}_{i_1\ldots i_k}|\,k!$, nous avons :

$$|\varphi(Q^{(j)}_{i_1\ldots i_k})| = \frac{1}{k!}\,|[\varphi'(P_j)\,g^{(j)}_1,\ \ldots,\ \varphi'(P_j)\,g^{(j)}_k]|+$$

$$+\frac{1}{2c^2}\,C''_n\varepsilon\theta_1 D^k k!\,|\,Q^{(j)}_{i_1\ldots i_k}|. \qquad (5)$$

En sommant ces relations pour tous les simplexes $Q^{(j)}_{i_1\ldots i_k}$ (pour j fixe) on obtient

$$\sum_{(i)}|\varphi(Q^{(j)}_{i_1\ldots i_k})| = |[\varphi'(P_j)\,g^{(j)}_1,\ \ldots,\ \varphi'(P_j)\,g^{(j)}_k]|+C'''_n\varepsilon\theta_2\,|\,Q^{(j)}\,|. \qquad (6)$$

En faisant la somme sur j on aboutit à la relation

$$\sum_j\sum_{(i)}|\varphi(Q^{(j)}_{i_1\ldots i_k})| = \sum_j|[\varphi'(P_j)\,g^{(j)}_1,\ \ldots,\ \varphi'(P_j)\,g^{(j)}_k]|+C'''_n\varepsilon\theta_3\,|\,U\,|. \qquad (7)$$

Le premier terme à droite n'est rien d'autre que la somme intégrale figurant dans la définition de l'aire d'une surface *3.61b* (cf. aussi *3.33c*). Lorsque $\varepsilon\to 0$, cette somme tend vers l'aire de la surface ; le second terme tend, évidemment, vers 0. Ainsi, donc, *la limite des aires des surfaces polyèdres que nous avons construites est égale à l'aire de la surface initiale*, ce qu'on affirmait.

e. Supposons que, sur la surface S et dans un de ses voisinages, soit donnée une fonction $f(x)$ continue par morceaux et bornée en

module par un nombre M. Si, avant d'additionner les relations (6), on multiplie la j-ème d'elles par le nombre $f(P^{(j)})$, on aura à la place de (7):

$$\sum_j f(P^{(j)}) \sum_{(i)} |\varphi(Q^{(j)}_{i_1\ldots i_k})| = \sum_j f(P^{(j)}) |[\varphi'(P_j) g_1^{(j)}, \ldots$$

$$\ldots, \varphi'(P_j) g_k^{(j)}]| + C_n''' \varepsilon \theta_3 M |U|,$$

où $|\theta_3| \leqslant 1$. Le premier terme à droite a pour limite, lorsque $\varepsilon \to 0$, l'intégrale de surface

$$\int_S f(x)\, dS.$$

Le premier membre diffère de l'intégrale de la fonction $f(x)$ sur la surface polyèdre Π_h par un infiniment petit, puisque, pour la fonction uniformément continue $f(x)$, nous avons:

$$\left| \int_{\Pi_h} f(x)\, dS - \sum_j f(P^{(j)}) \sum_{(i)} |\varphi(Q^{(j)}_{i_1\ldots i_k})| \right| =$$

$$= \left| \sum_{(i),\, j} \int_{\varphi(Q^{(j)}_{i_1\ldots i_k})} f(x)\, dS - \sum_{(i),\, j} f(P^{(j)}) |\varphi(Q^{(j)}_{i_1\ldots i_k})| \right| =$$

$$= \left| \sum_{(i),\, j} \sum_{\varphi(Q^{(j)}_{i_1\ldots i_k})} [f(x) - f(P^{(j)})]\, dS \right| < \varepsilon$$

si h est suffisamment petit ; pour une fonction continue par morceaux, on peut réaliser le raisonnement ci-dessus sur chaque morceau de la continuité uniforme de $f(x)$, puis additionner les résultats.

Ainsi, *l'intégrale sur la surface S de toute fonction $f(x)$ continue par morceaux peut être obtenue comme limite des intégrales de la même fonction sur certaines surfaces polyèdres approchant la surface S.*

3.65. Couche engendrée par une surface k-dimensionnelle.

a. Soit S une surface bidimensionnelle dérivable dans R_3. Menons la normale (*1.26b*) par tout son point et portons sur elle, de chaque côté de la surface, un segment de longueur h. Le corps tridimensionnel ainsi obtenu $V_h(S)$ s'appelle *couche d'épaisseur $2h$ engendrée par la surface S.*

Soit L une courbe dérivable dans R_3. Menons par tout son point x le plan normal et considérons dans ce plan le cercle de rayon h centré au point x. Le corps tridimensionnel ainsi obtenu $V_h(L)$ s'appelle *couche d'épaisseur $2h$ engendrée par la courbe L.*

b. Ces définitions élémentaires se généralisent au cas de surface k-dimensionnelle dans l'espace R_n de la manière suivante.

Soit S une surface k-dimensionnelle dans R_n, $S = \{x \in R_n, x = \varphi(u), u \in U \subset R_n\}$, ou, plus en détail

$$\left.\begin{aligned} x_1 &= \varphi_1(u_1, \ldots, u_k), \\ &\cdots\cdots\cdots\cdots \\ x_n &= \varphi_n(u_1, \ldots, u_k). \end{aligned}\right\} \tag{1}$$

On suppose que la fonction $x = \varphi(u)$ ou, ce qui revient au même, les fonctions $\varphi_1(u_1, \ldots, u_k), \ldots, \varphi_n(u_1, \ldots, u_k)$ possèdent les dérivées continues du premier et du second ordre dans le domaine U.

En tout point x de la surface S il existe un plan tangent k-dimensionnel Π et un plan normal $(n - k)$-dimensionnel (complément orthogonal à Π). Considérons dans le plan normal la boule de rayon h centrée au point x. La réunion de toutes ces boules représente un ensemble n-dimensionnel $V_h(S)$ que l'on appelle *couche n-dimensionnelle d'épaisseur* $2h$ *engendrée par la surface* S.

c. Dans certains cas, la couche n-dimensionnelle engendrée par la surface S admet une paramétrisation « canonique », à savoir une représentation de la forme :

$$\left.\begin{aligned} x_1 &= f_1(u_1, \ldots, u_k, v_{k+1}, \ldots, v_n) \equiv f_1(u, v), \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_n &= f_n(u_1, \ldots, u_k, v_{k+1}, \ldots, v_n) \equiv f_n(u, v), \end{aligned}\right\} \tag{2}$$

où $f_1, \ldots, f_n$ sont des fonctions à dérivées premières continues dans le domaine $U \times Q_h \subset R_n$. Ici $U \subset R_k$ est le domaine initial de variation des paramètres $u = (u_1, \ldots, u_k)$, et Q_h, la boule $(n - k)$-dimensionnelle de rayon h et de centre au point $v = 0$, $v = (v_{k+1}, \ldots, v_n)$. En outre, le système des fonctions (2) doit satisfaire aux conditions suivantes : 1) pour $v = 0$, les fonctions $f_1(u, 0), \ldots, f_n(u, 0)$ coïncident respectivement avec les fonctions $\varphi_1(u), \ldots, \varphi_n(u)$ qui donnent la représentation paramétrique de la surface S ; 2) pour $u \in U$ fixe, les fonctions (2) fournissent une représentation isométrique de la boule $(n - k)$-dimensionnelle centrée au point $x = \varphi(u) \in S$.

Etant donnée l'application (2), il est aisé de calculer le module du déterminant $\frac{\partial(f_1, \ldots, f_n)}{\partial(u_1, \ldots, v_n)}$ aux points de la surface S, i.e. pour les valeurs $v_{k+1} = \ldots = v_n = 0$. Ce module, nous le connaissons (*3.56*), est égal au volume de l'image de la cellule $a_1 \leqslant u_1 \leqslant a_1 + \delta, \ldots, a_k \leqslant u_k \leqslant a_k + \delta, 0 \leqslant v_{k+1} \leqslant \delta, \ldots, 0 \leqslant v_n \leqslant \delta$ par l'application $x = f'(u, v)$ divisé par le volume de cette cellule même qui vaut δ^n. L'application $f'(u, v)$, pour $(u, v) = (a_1, \ldots, a_k, 0, \ldots, 0)$ transforme la cellule $(a_1 \leqslant u_1 \leqslant a_1 + \delta, \ldots, a_k \leqslant u_k \leqslant a_k + \delta) \subset U$ en le parallélépipède dans le plan tangent à la surface S dont les arêtes sont $\frac{\partial \varphi}{\partial u_1}\delta, \ldots, \frac{\partial \varphi}{\partial u_k}\delta$, et la cellule $(0 \leqslant v_{k+1} \leqslant \delta, \ldots, 0 \leqslant v_n \leqslant \delta) \subset V_h$ de volume δ^{n-k} en une cellule dans le plan normal ayant le même volume δ^{n-k}.

Pour le rapport des volumes, nous obtenons (*3.55d*):

$$\left|\frac{\partial(f_1, \ldots, f_n)}{\partial(u_1, \ldots, v_n)}\right| = \frac{\left|\left[\frac{\partial\varphi}{\partial u_1}\delta, \ldots, \frac{\partial\varphi}{\partial u_k}\delta\right]\right| \delta^{n-k}}{\delta^n} = \left|\left[\frac{\partial\varphi}{\partial u_1}, \ldots, \frac{\partial\varphi}{\partial u_k}\right]\right|.$$

Il en résulte en particulier que l'application (2), du moment qu'elle existe, est biunivoque (et bidérivable) dans un voisinage de chaque point ordinaire de la surface S, i.e. de chaque point où $\left|\left[\frac{\partial\varphi}{\partial u_1}, \ldots, \frac{\partial\varphi}{\partial u_k}\right]\right| \neq 0$. Nous avons donc, aux points de la surface S:

$$\left|\frac{\partial(f_1, \ldots, f_n)}{\partial(u_1, \ldots, v_n)}\right| du_1 \ldots du_k = \left|\frac{\partial(\varphi_1, \ldots, \varphi_k)}{\partial(u_1, \ldots, u_k)}\right| du_1 \ldots du_k,$$

c'est-à-dire que *le produit du jacobien* $\left|\frac{\partial(f_1, \ldots, f_n)}{\partial(u_1, \ldots, v_n)}\right|$ *par* $du_1 \ldots du_k$ *est égal à l'aire de l'élément de la surface* S.

d. Si la valeur $v = (v_{k+1}, \ldots, v_n) \neq 0$ est fixe et $u = (u_1, \ldots, u_k)$ parcourt l'ensemble U, alors les fonctions (2) décrivent une surface k-dimensionnelle S_v qu'il est naturel d'appeler « parallèle » à S. Comme les images de la cellule $(a_1 \leqslant u_1 \leqslant a_1 + \delta, \ldots, a_k \leqslant u_k \leqslant a_k + \delta) \in U$ et de la cellule $(b_{k+1} \leqslant v_{k+1} \leqslant b_{k+1} + \delta, \ldots, b_n \leqslant v_n \leqslant b_n + \delta) \in V$ par l'application f' n'appartiennent, en général, plus aux plans orthogonaux (k-dimensionnel et $(n - k)$-dimensionnel), le produit $\frac{\partial(f_1, \ldots, f_n)}{\partial(u_1, \ldots, v_n)} du_1 \ldots du_k$ n'est pas un élément de la surface S_v. Mais, si l'application $x = \varphi(u)$ (1) est telle que les images des cellules indiquées plus haut appartiennent aux plans orthogonaux pour $v \neq 0$, alors le raisonnement précédent reste en vigueur, et l'expression $\frac{\partial(f_1, \ldots, f_n)}{\partial(u_1, \ldots, v_n)} du_1 \ldots du_k$ est l'aire d'un élément de la surface S_v.

e. L'intégration d'une fonction $F(x)$ sur la couche $W_h(S)$ se fait, en général, d'après la règle *3.58a*:

$$\int_{W_h(S)} F(x)\, dx = \int_{U \times Q_h} F(f(u, v)) \left|\frac{\partial(f_1, \ldots, f_n)}{\partial(u_1, \ldots, v_n)}\right| du_1 \ldots dv_n. \quad (3)$$

Effectuons l'intégration par rapport aux variables u, pour v fixes; nous obtenons la fonction

$$\Phi(v) = \int_U F(f(u, v)) \left|\frac{\partial(f_1, \ldots, f_n)}{\partial(u_1, \ldots, v_n)}\right| du_1 \ldots du_k. \quad (4)$$

Si l'application f conserve l'orthogonalité des surfaces S_v aux boules correspondantes $u = \text{const}$, alors l'intégrale (4) est étendue

à la surface S_v, comme on l'a vu plus haut. Pour l'intégrale (3) nous avons

$$\int\limits_{W_h(S)} F(x)\,dx = \int\limits_{Q_h} \Phi(v)\,dv, \tag{5}$$

ce qu'on peut traiter comme suit : on trouve l'intégrale de la fonction $F(x)$ sur la surface S_v, puis on intègre le résultat sur la boule Q_h. Dans le cas général où la surface S_v n'est pas orthogonale aux boules $u=\text{const}$, la relation (5) reste, naturellement, valable, mais on ne peut plus dire que la quantité $\Phi(v)$ soit intégrale de la fonction $F(x)$ sur la surface S_v. On ne peut qu'affirmer, d'après le théorème *3.35a* sur la continuité d'une intégrale par rapport au paramètre, que la fonction $\Phi(v)$ dans (5) est continue, donc, en particulier, pour $v \to 0$, tend vers la quantité $\Phi(0)$ qui est l'intégrale de la fonction $F(x)$ sur la surface S.

f. Dans certains cas, l'existence d'une représentation canonique (2) de la couche $W_h(S)$ est évidente. Soit, par exemple, S la sphère $(n-1)$-dimensionnelle d'un rayon r dans R_n, centrée au point 0, ou bien une partie de cette sphère d'aire positive. Alors la couche n-dimensionnelle $W_h(S)$ $(h < r)$ engendrée par la sphère est le domaine obtenu en enlevant à la boule de rayon $r+h$ la boule concentrique de rayon $r-h$. Une paramétrisation canonique (2) se déduit de n'importe quelle paramétrisation de la sphère

$$\begin{aligned} x_1 &= \varphi_1(u_1, \ldots, u_{n-1}), \\ &\cdots\cdots\cdots \\ x_n &= \varphi_n(u_1, \ldots, u_{n-1}) \end{aligned}$$

en additionnant au second membre de la i-ème ligne le terme $v_n\varphi_i/r$ $(|v_n| \leqslant h)$. Les boules $u=\text{const}$ représentent des segments de longueur $2h$ des demi-droites radiales ; les surfaces S_v $(-h \leqslant v \leqslant h < r)$ sont des parties des sphères de rayon $r+v$ orthogonales aux demi-droites radiales. Par conséquent, l'égalité (5) signifie, dans le présent cas, que *l'intégrale de la fonction $F(x)$ sur la couche $W_h(S)$ peut être obtenue en intégrant la fonction $F(x)$ sur la partie convenable de la sphère S_v, puis en intégrant par rapport à v de $r-h$ à $r+h$.*

g. Etablissons ici une condition suffisante d'existence locale d'une représentation canonique (2) pour la surface k-dimensionnelle arbitraire :

Théorème. *S'il est possible, pour une surface k-dimensionnelle $S \subset R_n$, d'indiquer, à tout plan normal $\gamma(x)$, une base orthonormée des vecteurs $g_{k+1}(x), \ldots, g_n(x)$ qui possèdent les dérivées continues en x, alors chaque point $x_0 \in S$ admet un voisinage dans lequel la couche $W_h(S)$ possède une représentation canonique.*

Démonstration. Soient $e_1, \ldots, e_n$ la base initiale orthonormée de l'espace R_n et

$$x_1 = \varphi_1(u_1, \ldots, u_k), \ldots, x_n = \varphi_n(u_1, \ldots, u_k)$$

une représentation paramétrique, par rapport à cette base, de la surface S au voisinage d'un point x_0. Considérons le système de fonctions

$$\left.\begin{aligned} \bar{x}_1 &= \varphi_1(u_1, \ldots, u_k) + v_{k+1}(g_{k+1}, e_1) + \ldots + v_n(g_n, e_1), \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \bar{x}_n &= \varphi_n(u_1, \ldots, u_k) + v_{k+1}(g_{k+1}, e_n) + \ldots + v_n(g_n, e_n). \end{aligned}\right\} \quad (6)$$

Il est évident que, pour $v_{k+1} = \ldots = v_n = 0$, nous avons un point x de la surface S. En fixant $u_1, \ldots, u_k$ on additionne le vecteur allant au point x sur la surface S et le vecteur $\rho = v_{k+1}g_{k+1} + \ldots + v_ng_n$ qui se trouve dans le plan normal $\gamma(x)$; lorsque les paramètres $v_{k+1}, \ldots, v_n$ parcourent une boule de rayon r dans R_{n-k}, le vecteur $x + \rho$ parcourt une boule isométrique de rayon r et de centre x dans le plan $\gamma(x)$. D'après c, le déterminant de la matrice $\frac{\partial(x_1, \ldots, x_n)}{\partial(u_1, \ldots, v_n)}$ est non nul aux points ordinaires de la surface S, de sorte que l'application (6) est biunivoque dans un voisinage de chaque tel point, ce qu'il nous fallait.

h. Montrons qu'au voisinage de chaque point ordinaire x de la surface S il est possible de choisir une base demandée par le théorème g dans le plan normal $\gamma(x)$.

Théorème. *Pour tout point ordinaire $x_0 = \varphi(u_0)$ de la surface S, il existe un voisinage $U_0 \subset R_k$ dans lequel on peut définir les fonctions vectorielles orthonormées $g_j(\varphi(u))$, $j = k+1, \ldots, n$ ayant les dérivées continues et les valeurs dans le plan normal $\gamma(x)$.*

Démonstration. Choisissons une base orthonormale quelconque $g_{k+1}(x_0), \ldots, g_n(x_0)$ dans le plan $\gamma(x_0)$. Ensuite, pour tout autre point $x \in S$, trouvons dans le plan $\gamma(x)$ les vecteurs $q_{k+1}(x), \ldots, q_n(x)$ dont les projections sur le plan $\gamma(x_0)$ coïncident avec les vecteurs $g_{k+1}(x_0), \ldots, g_n(x_0)$ respectivement. Il est affirmé que, dans un voisinage du point x_0, de tels vecteurs existent et sont définis d'une façon unique. L'orthonormalisation des vecteurs $q_{k+1}(x), \ldots, q_n(x)$ nous fournira les vecteurs cherchés $g_{k+1}(x), \ldots, g_n(x)$. Pour déterminer le vecteur $q_{k+j}(x)$, nous avons à écrire le système d'équations :

$$\left.\begin{aligned} &(q_{k+j}(x), \psi_i(x)) = 0, \quad \psi_i(x) = \frac{\partial\varphi(x)}{\partial u_i}, \quad i = 1, \ldots, k; \\ &(q_{k+1}(x) - g_{k+j}(x_0), g_r(x_0)) = 0 \\ &\qquad (r = k+1, \ldots, n; \quad i = 1, \ldots, n-k). \end{aligned}\right\} \quad (7)$$

Etant donnés $u \in U \subset R_k$, $y^j \in R_n$, considérons l'application du couple (u, y^j) dans R_n réalisée par les formules

$$\left.\begin{aligned} \xi_1 &= (y^j, \psi_1(\varphi(u))), \\ &\cdots\cdots\cdots \\ \xi_k &= (y^j, \psi_k(\varphi(u))), \\ \xi_{k+1} &= (y^j - g_{k+j}(x_0), g_{k+1}(x_0)), \\ &\cdots\cdots\cdots\cdots \\ \xi_n &= (y^j - g_{k+j}(x_0), g_n(x_0)). \end{aligned}\right\} \tag{8}$$

Cette application est linéaire en y, sa dérivée par rapport à y est déterminée par la matrice des coordonnées des vecteurs $\psi_1(\varphi(u)), \ldots, \psi_k(\varphi(u)), g_{k+1}(x_0), \ldots, g_n(x_0)$. Cette matrice est non dégénérée, pour $u = u_0$ au moins; la valeur $u = u_0$ correspond à un point ordinaire de la surface, où les vecteurs $\psi_1(\varphi(u_0)), \ldots, \psi_k(\varphi(u_0))$ sont linéairement indépendants et les vecteurs $g_{k+1}(x_0), \ldots, g_n(x_0)$ forment une base orthonormale dans le plan normal $\gamma(x_0)$, par conséquent,

$$|[\psi_1(\varphi(u_0)), \ldots, \psi_k(\varphi(u_0)), g_{k+1}(x_0), \ldots, g_n(x_0)]| = \\ = |[\psi_1(\varphi(u_0)), \ldots, \psi_k(\varphi(u_0))]| \neq 0.$$

En outre, les premiers membres du système (8) s'annulent pour $u = u_0$ et $y^j = g_{k+j}(x_0)$. D'après le théorème *1.53* sur la fonction implicite, dans un voisinage du point u_0 il existe l'unique fonction vectorielle $y^j(u)$ qui annule identiquement les premiers membres du système (8) et vaut $g_{k+j}(x_0)$ pour $u = u_0$. En posant $q_{k+j}(u) = y^j(u)$ nous obtenons la solution du système (7), ce qu'il nous fallait.

Les fonctions $q_{k+j}(u)$ sont continues et dérivables en u en raison de la dérivabilité des fonctions $\psi_i(\varphi(u))$ (la surface S est supposée deux fois dérivable) et du théorème sur la fonction implicite.

L'orthonormalisation du système $q_{k+j}(u)$ fournit les fonctions $g_{k+j}(u)$ qui sont toujours continues et dérivables, puisque les coefficients d'orthonormalisation s'expriment linéairement par les rapports des produits scalaires des vecteurs $q_{k+j}(u)$ aux déterminants de la forme

$$|[q_k(u)]|^2, \quad |[q_{k+1}(u), q_{k+2}(u)]|^2, \ldots, |[q_{k+1}(u), \ldots, q_n(u)]|^2$$

(*L7.51*); ces déterminants sont continus en u et valent 1 pour $u = u_0$, donc sont non nuls dans un voisinage du point u_0. Le théorème est démontré.

3.66. Aire d'une surface comme limite de l'aire moyenne de la couche n-dimensionnelle qu'elle engendre.

a. Soient S une surface bidimensionnelle dans R_3 et $W_h(S)$ la couche d'épaisseur $2h$ qu'elle engendre (*3.65a*). Le volume de la couche divisé par $2h$ s'appelle *aire moyenne de la couche*; une conjecture assez naturelle, qui se réalise au moins dans le cas où $S \subset R_3$ est un plan, consiste en ce que, *lorsque* $h \to 0$, *l'aire moyenne de la couche ait pour limite l'aire de la surface* S.

D'une façon analogue, soient L une courbe dérivable dans R_3 et $W_h(L)$ la couche d'épaisseur $2h$ qu'elle engendre (*3.65a*). Le volume de la couche divisé par πh^2 s'appelle *longueur moyenne de la couche*; on peut conjecturer que, de même que dans le cas où $L \subset R_3$ est une droite, *la longueur moyenne de la couche tende vers la longueur de la courbe* L *lorsque* $h \to 0$.

b. Les deux conjectures s'avèrent justes; nous allons formuler un théorème général pour une surface k-dimensionnelle dans R_n.

Soient $S \subset R_n$ une surface k-dimensionnelle et $W_h(S)$ la couche d'épaisseur $2h$ qu'elle engendre. Le volume de la couche divisé par le volume de la boule $(h-k)$-dimensionnelle de rayon h s'appelle *aire moyenne de la couche* (pour $k > 1$; si $k = 1$ on l'appelle *longueur moyenne de la couche*).

Théorème. *Si la couche n-dimensionnelle* $W_h(S)$ *d'épaisseur* $h \leqslant h_0$ *engendrée par la surface* S *admet une représentation canonique 3.65* (2), *alors l'aire moyenne de la couche tend, pour* $h \to 0$, *vers l'aire de la surface* S.

Démonstration. Le volume de la couche $W_h(S)$ est donné par l'intégrale *3.65* (3) ($F(x) \equiv 1$):

$$|W_h(S)| = \int\limits_{U \times V_h} \left| \frac{\partial(f_1, \ldots, f_n)}{\partial(u_1, \ldots, v_n)} \right| du_1 \ldots dv_n.$$

Effectuons l'intégration par rapport aux variables u, pour v fixes; nous aurons la fonction:

$$\Phi(v_{k+1}, \ldots, v_n) = \int\limits_U \left| \frac{\partial(f_1, \ldots, f_n)}{\partial(u_1, \ldots, v_n)} \right| du_1 \ldots du_k.$$

La fonction $\Phi(v_{k+1}, \ldots, v_n)$ est continue par rapport à ses variables. Si l'on intègre Φ sur la boule V_h, puis on divise le résultat par le volume de la boule, on obtient l'aire moyenne de la couche S_h qui est donc égale à la valeur moyenne de la fonction $\Phi(v_{k+1}, \ldots, v_n)$ dans la boule V_h. Cette valeur moyenne tend, pour $h \to 0$, vers la valeur de la fonction $\Phi(v_{k+1}, \ldots, v_n)$ au point $v_{k+1} = \ldots = v_n = 0$. Ainsi, l'aire moyenne de la couche a pour limite

$$\Phi(0, \ldots, 0) = \int\limits_U \left| \frac{\partial(f_1, \ldots, f_n)}{\partial(u_1, \ldots, v_n)} \right|_{v=0} du_1 \ldots du_k.$$

En utilisant la valeur du jacobien $\left|\frac{\partial(f_1, \ldots, f_n)}{\partial(u_1, \ldots, v_n)}\right|$ sur la surface S obtenue plus haut (*3.65c*) nous voyons que

$$\lim_{h\to 0} \frac{|W_h(S)|}{|Q_h|} = \int_U \left|\left[\frac{\partial\varphi}{\partial u_1}, \ldots, \frac{\partial\varphi}{\partial u_k}\right]\right| du,$$

ce qui représente l'aire de la surface S.

c. Exemple. Trouvons l'aire de la sphère $(n-1)$-dimensionnelle $S = S_{n-1}(r)$ d'un rayon r dans R_n. Comme nous l'avons vu dans *3.65f*, la couche n-dimensionnelle $W_h(S)$ engendrée par cette sphère est le domaine obtenu en enlevant à la boule Q_{r+h} de rayon $r+h$ la boule concentrique Q_{r-h} de rayon $r-h$. Le volume $|Q_r|$ d'une boule de rayon r dans l'espace n-dimensionnelle vaut, d'après *3.56b*,

$$|Q_r| = \frac{\pi^{n/2} r^n}{\Gamma\left(\frac{n}{2}+1\right)}.$$

En appliquant le théorème *b* nous voyons que l'aire de la sphère $(n-1)$-dimensionnelle de rayon r dans R_n est

$$\lim_{h\to 0} \frac{|W_h(S)|}{2h} = \lim_{h\to 0} \frac{|Q_{r+h}| - |Q_{r-h}|}{2h} = \frac{d|Q_r|}{dr} = \frac{n\pi^{n/2} r^{n-1}}{\Gamma\left(\frac{n}{2}+1\right)}.$$

d. En combinant les résultats *b* et *3.65h* nous concluons:

Si la surface $S = \{x = \varphi(u),\ u \in U \subset R_k\}$ *est donnée par une fonction deux fois dérivable* $\varphi(u)$ $(R_k \to R_n)$, *et si* x_0 *est un point ordinaire de cette surface, alors il existe un voisinage* G *du point* x_0 *dans lequel il est possible de construire une couche* n*-dimensionnelle* $W_h(S)$ *d'épaisseur* $2h$ *engendrée par la surface* S; *l'aire de la partie de la surface qui se trouve dans le voisinage* G *peut être calculée comme limite de l'aire moyenne de la couche pour* $h \to 0$.

La question se pose: est-ce qu'un théorème analogue a lieu au sens non local, i.e. pour la surface S toute entière? Cette question est liée à des considérations topologiques, par exemple à l'absence de points d'auto-intersection de la couche $W_h(S)$, et s'avère trop compliquée pour être considérée dans notre cours.

§ 3.7. Intégrales impropres

3.71. Définitions de base. Nous avons établi la définition de l'intégrale d'une fonction bornée sur un domaine borné (ensemble jordanien) de l'espace R_n. Ici nous étendons cette définition aux cas suivants: a) d'une fonction localement bornée dans un domaine fermé non borné (intégrale de première espèce); b) d'une fonction non bornée dans un domaine fermé borné (intégrale de deuxième espèce) et c) d'une fonction non bornée dans un domaine

fermé non borné (intégrale de troisième espèce). Ici nous appelons *domaine fermé* un ensemble fermé qui est la fermeture d'un ensemble ouvert.

a. Supposons qu'il y ait un domaine fermé non borné $G \subset R_n$ sur lequel soit définie une fonction $f(x)$ admissible, i.e. bornée et continue par morceaux sur chaque ensemble borné; nous tenons à attribuer un sens à l'intégrale impropre de première espèce

$$If = \int_G f(x)\,dx. \tag{1}$$

Choisissons arbitrairement une suite de domaines fermés bornés $G_1 \subset G_2 \subset \ldots \subset G_m \subset \ldots \subset G$ vérifiant la condition suivante: pour toute boule $V_\rho = \{x \in R_n : |x| \leqslant \rho\}$, il existe un m tel que le domaine G_m (donc aussi chaque domaine suivant) contient l'ensemble $V_\rho \cap G$. Une telle suite de domaines sera dite *exhaustive*. Les intégrales

$$I_m f = \int_{G_m} f(x)\,dx \tag{2}$$

existent. Si, pour $m \to \infty$, la suite $I_m f$ tend vers une limite (finie) qui ne dépend pas du choix de la suite exhaustive G_m, alors nous disons que l'intégrale (1) *existe* (ou *converge*) et nous posons par définition:

$$If \equiv \int_G f(x)\,dx = \lim_{m\to\infty} \int_{G_m} f(x)\,dx. \tag{3}$$

Si les intégrales $I_m f$ n'ont pas de limite pour $m \to \infty$, nous disons que l'intégrale (1) *diverge*.

b. Supposons que dans un domaine fermé borné $G \subset R_m$ soit donnée une fonction *admissible* $f(x)$; cela signifie, cette fois, qu'il existe un ensemble négligeable $Z \subset G$ en dehors de chaque voisinage (*3.23b*) duquel la fonction $f(x)$ est bornée et continue par morceaux.

Nous tenons à définir l'intégrale impropre de deuxième espèce

$$If = \int_G f(x)\,dx. \tag{4}$$

Choisissons arbitrairement une suite d'ensembles jordaniens $G_1 \subset G_2 \subset \ldots \subset G$ telle que chaque ensemble complémentaire $G - G_m$ contient Z strictement à son intérieur, étant contenu lui-même dans un ε_m-voisinage de l'ensemble Z, et cela avec $\varepsilon_m \to 0$ pour $m \to \infty$. Une telle suite $G_1 \subset G_2 \subset \ldots$ sera dite *exhaustive*. Si les intégrales

$$I_m f = \int_{G_m} f(x)\,dx \tag{5}$$

tendent, pour $m \to \infty$, vers une limite qui ne dépend pas du choix de la suite G_m, nous disons que l'intégrale (4) *existe* ou *converge* (nous disons *diverge* dans le cas contraire), et nous posons par définition :

$$If \equiv \int_G f(x)\,dx = \lim_{m\to\infty} \int_{G_m} f(x)\,dx. \tag{6}$$

c. Supposons que, dans un domaine fermé non borné $G \subset R_n$, soit donnée une fonction $f(x)$, éventuellement non bornée. Plus exactement, nous supposons que chaque boule $V_\rho = \{x \in R_n : |x| \leqslant \rho\}$ contienne un ensemble négligeable Z_ρ tel que la fonction $f(x)$ soit bornée et continue par morceaux sur la différence de $V_\rho \cap G$ et de tout voisinage de Z_ρ. De telles fonctions sont dites ici *admissibles*. La définition de l'« intégrale impropre de troisième espèce »

$$If = \int_G f(x)\,dx \tag{7}$$

est construite comme suit. Appelons *exhaustive* chaque suite $G_1 \subset G_2 \subset \ldots \subset G$ d'ensembles jordaniens bornés si, pour toute boule $V_\rho = \{x \in R_n : |x| \leqslant \rho\}$ et pour tout $\varepsilon > 0$, il existe un m tel que le domaine G_m (donc aussi chaque domaine suivant) contient l'ensemble $V_\rho \cap G$ à l'exception de ε-voisinage de l'ensemble Z_ρ, et que, du moment que ce domaine G_m, pour un $\rho_1 > \rho$, soit contenu dans la boule V_{ρ_1}, il ne contient pas non plus le ε-voisinage de l'ensemble Z_{ρ_1}. Alors les intégrales

$$I_m f = \int_{G_m} f(x)\,dx \tag{8}$$

sont définies ; si elles tendent, pour $m \to \infty$, vers une limite If qui ne dépend pas du choix de la suite exhaustive G_m, nous disons que l'intégrale (7) *existe* ou *converge* (*diverge* dans le cas contraire), et nous posons par définition :

$$If \equiv \int_G f(x)\,dx = \lim_{m\to\infty} \int_{G_m} f(x)\,dx. \tag{9}$$

d. La construction des suites exhaustives dans tous les trois cas a, b, c possède une même propriété : si $G_1 \subset G_2 \subset \ldots$ et $G'_1 \subset G'_2 \subset \ldots$ sont deux suites exhaustives (pour un même type d'intégrale impropre), alors chaque domaine G_m de la première suite est contenu dans un domaine $G'_{m'}$ de la deuxième et vice versa. Par conséquent, étant données deux suites exhaustives, il est toujours possible de construire la suite exhaustive mixte :

$$G_{i_1} \subset G'_{i'_1} \subset G_{i_2} \subset G'_{i'_2} \subset \ldots \tag{10}$$

Il en résulte que l'existence d'une limite des intégrales du type (2) suivant *chaque* suite exhaustive implique déjà l'identité de ces limites: la limite suivant la suite mixte (10) doit coïncider avec celle suivant la suite $G_1 \subset G_2 \subset \ldots$ et avec celle suivant la suite $G_1' \subset G_2' \subset \ldots$, d'où l'égalité de ces dernières limites.

e. Nous voyons donc que, dans tous les cas, la définition de l'intégrale impropre If d'une fonction admissible $f(x)$ est ramenée à la définition de la limite de la suite des intégrales propres de la fonction $f(x)$ suivant une suite exhaustive arbitraire. Il en découle, en particulier, que l'intégrale impropre, de même que l'intégrale propre, possède la propriété linéaire: si elle converge pour deux fonctions $f_1(x)$ et $f_2(x)$, elle converge également pour toute combinaison linéaire $c_1 f_1(x) + c_2 f_2(x)$ (c_1 et c_2 étant des nombres), et l'on a:

$$I(c_1 f_1 + c_2 f_2) = c_1 I f_1 + c_2 I f_2.$$

3.72. Intégrales impropres des fonctions non négatives et convergence absolue.

a. Si, en plus des hypothèses *3.71*, la fonction $f(x)$ est non négative, $f(x) \geqslant 0$, alors les définitions de toutes les intégrales impropres *3.71a-c* deviennent plus simples. Il suffit de considérer les intégrales propres du type (2) pour une seule quelconque suite exhaustive G_m et de reconnaître si la suite de ces intégrales soit bornée. En vertu de la relation signalée dans *3.71d* entre les suites exhaustives différentes, du moment que les intégrales (2) soient bornées sur une suite G_m et que la fonction $f(x)$ soit non négative, les intégrales (2) sont bornées sur toute autre suite exhaustive. Or, dès que la suite des intégrales $I_m f$ soit bornée et que $I_1 f \leqslant I_2 f \leqslant \ldots$, la limite des nombres $I_m f$, suivant chaque suite exhaustive, existe, d'où, d'après *3.71d*, l'existence de l'intégrale impropre correspondante.

b. Supposons que $0 \leqslant f(x) \leqslant g(x)$ et que les deux fonctions soient admissibles (*3.71c*). *Si l'intégrale sur le domaine G de $g(x)$ existe, il en est de même de l'intégrale de $f(x)$, et l'on a $If \leqslant Ig$.* En effet, les intégrales de $g(x)$ pour une suite exhaustive tendent vers une limite, donc sont majorées par le nombre Ig; par conséquent, les intégrales de $f(x)$ pour la même suite exhaustive sont majorées par le même nombre Ig; vu *a*, il en résulte que l'intégrale If existe et que la relation $If \leqslant Ig$ a lieu.

En conséquence, nous obtenons: *si $0 \leqslant f(x) \leqslant g(x)$ sont des fonctions admissibles et l'intégrale de $f(x)$ sur le domaine G diverge, alors l'intégrale sur le même domaine de la fonction $g(x)$ diverge elle aussi.*

Les deux résultats ci-formulés représentent le *critère de comparaison pour les intégrales impropres.*

c. *Si $f(x)$ est une fonction admissible non négative et si l'intégrale $If = I_G f$ sur un domaine G converge, alors l'intégrale $I_Q f$ sur tout domaine $Q \subset G$ converge elle aussi, et l'on a $I_Q f \leqslant I_G f$.* En effet, du

moment que $\chi_Q(x)$ soit la fonction caractéristique du domaine Q, la fonction $\chi_Q(x) f(x)$ est admissible avec $f(x)$ et satisfait à l'inégalité $0 \leqslant \chi_Q(x) f(x) \leqslant f(x)$; en appliquant a nous avons

$$I_G \chi_Q f = I_Q f \leqslant I_G f,$$

ce qu'il nous fallait.

En conséquence nous obtenons: *si l'intégrale sur un domaine Q d'une fonction admissible $f(x) \geqslant 0$ diverge, alors l'intégrale de $f(x)$ sur tout domaine $G \supset Q$ diverge elle aussi.*

d. *Si, pour une fonction admissible $f(x) \geqslant 0$, les intégrales sur deux ensembles jordaniens G' et G'' convergent, alors son intégrale sur l'ensemble $G' \cup G'' = G$ converge elle aussi.*

En remplaçant G'' par $G'' - G' \cap G''$ on peut réduire l'affirmation au cas où les ensembles jordaniens G' et G'' sont disjoints. Soient G'_m une suite exhaustive d'ensembles jordaniens pour le domaine G' et G''_m pour le domaine G''. Alors, évidemment, les ensembles $G'_m \cup G''_m$ forment une suite exhaustive pour l'ensemble G.

Inversement, toute suite exhaustive G_m pour l'ensemble G engendre deux suites $G'_m = G' \cap G_m$ et $G''_m = G'' \cap G_m$ qui sont, évidemment, exhaustives pour les ensembles G' et G'' respectivement.

D'après *3.33b*, nous avons

$$\int\limits_{G'_m \cup G''_m} f(x)dx = \int\limits_{G'_m} f(x)\,dx + \int\limits_{G''_m} f(x)\,dx \qquad (1)$$

et, comme les intégrales dans le second membre tendent vers leurs limites respectives, l'intégrale dans le premier membre a aussi une limite. Il en résulte, d'après *3.71d*, l'existence de l'intégrale de la fonction $f(x)$ sur le domaine $G = G' \cup G''$. En particulier, si les domaines G' et G'' sont disjoints, le passage à la limite dans (1) fournit:

$$\int\limits_{G' \cup G''} f(x)\,dx = \int\limits_{G'} f(x)\,dx + \int\limits_{G''} f(x)\,dx.$$

Il va de soi que le résultat obtenu s'étend aisément à un nombre arbitraire (fini) de domaines jordaniens $G', G'', \ldots, G^{(k)}$.

e. Principe de localisation. *Soit $f(x) \geqslant 0$ une fonction admissible non bornée dans un domaine fermé borné $G \subset R_n$. Si, pour tout point $a \in G$, il existe un voisinage $V(a)$ où la fonction $f(x)$ est intégrable (i.e. l'intégrale impropre de deuxième espèce sur $V(a)$ de $f(x)$ converge), alors la fonction $f(x)$ est intégrable sur tout le domaine G; et si, pour un point $b \in G$ au moins, il existe un voisinage $V(b)$ où la fonction $f(x)$ n'est pas intégrable, alors elle n'est pas intégrable non plus sur tout le domaine G.*

Démonstration. En fixant pour tout point $a \in G$ un voisinage $V(a)$ où la fonction $f(x)$ est intégrable, choisissons-en un

recouvrement fini $V(a_1), \ldots, V(a_k)$ du domaine G. En vertu de d, notre fonction sera intégrable sur la réunion G de ces derniers voisinages, ce qu'il fallait démontrer.

Si la fonction $f(x)$ n'est pas intégrable sur le voisinage $V(b)$, elle ne l'est pas non plus sur tout le domaine G d'après c.

f. Intégrales absolument convergentes. Soit $f(x)$ une fonction admissible sur un domaine $G \subset R_n$. Supposons qu'il y ait une fonction admissible non négative $g(x)$ dont l'intégrale sur G converge. *Alors, si* $|f(x)| \leqslant g(x)$, *les fonctions* $f(x)$ *et* $|f(x)|$ *sont elles aussi intégrables sur le domaine* G, *et l'on a*:

$$\left|\int\limits_G f(x)\,dx\right| \leqslant \int\limits_G |f(x)|\,dx \leqslant \int\limits_G g(x)\,dx. \tag{2}$$

Pour le prouver, considérons une suite exhaustive quelconque $G_1 \subset G_2 \subset \ldots \subset G$ d'ensembles jordaniens: nous avons pour $k < m$:

$$\left|\int\limits_{G_m} f(x)\,dx - \int\limits_{G_k} f(x)\,dx\right| = \left|\int\limits_{G_m - G_k} f(x)\,dx\right| \leqslant$$

$$\leqslant \int\limits_{G_m - G_k} \left|f(x)\right| dx \leqslant \int\limits_{G_m - G_k} g(x)\,dx = \int\limits_{G_m} g(x)\,dx - \int\limits_{G_k} g(x)\,dx.$$

Ici le dernier membre est non négatif et tend vers zéro pour $k \to \infty$, $m \to \infty$ en vertu de la convergence de l'intégrale de $g(x)$. Vu le critère de Cauchy, il existe également la limite des intégrales de la fonction $f(x)$ sur les domaines G_m ; cela veut dire, d'après *3.71c*, qu'il existe l'intégrale de $f(x)$ sur G. L'existence de l'intégrale de $|f(x)|$ résulte de b. En passant à la limite pour $m \to \infty$ dans l'inégalité

$$\left|\int\limits_{G_m} f(x)\,dx\right| \leqslant \int\limits_{G_m} |f(x)|\,dx \leqslant \int\limits_{G_m} g(x)\,dx,$$

nous arrivons à la relation cherchée (2).

L'intégrale d'une fonction $f(x)$ qui satisfait à toutes les conditions formulées est dite *absolument convergente*. Il est intéressant de remarquer qu'il n'existe en général pas, dans R_n, d'intégrale qui converge non absolument (voir l'exercice 6).

3.73. Exemples.

a. Soit $f(r) \geqslant 0$ une fonction, donnée sur la demi-droite $0 < a \leqslant r < \infty$, continue par morceaux sur tout intervalle fini $a \leqslant r \leqslant b$. En posant $r^2 = \sum\limits_{i=1}^{n} x_i^2 = |x|^2$ nous obtenons la fonction admissible $f(r) = f\left(\sqrt{\sum\limits_{i=1}^{n} x_i^2}\right)$ définie dans R_n. Considérons-la dans le domaine

$G = \{x \in R_n : |x| \geqslant a, \frac{x}{|x|} \in \Sigma\}$ où Σ est un ensemble donné d'aire positive sur la sphère unité de l'espace R_n. Discutons la convergence de l'intégrale impropre de première espèce

$$\int_G f(r)\,dx. \tag{1}$$

En tant que suite exhaustive, choisissons les domaines $G_m = \{x \in R_n : a \leqslant |x| \leqslant m, \frac{x}{|x|} \in \Sigma\}$. L'intégrale de la fonction $f(r)$ sur le domaine G_m est calculée d'après la règle *3.65a*, à savoir

$$\int_{G_m} f(r)\,dx = \int_{r=a}^{m} \left\{ \int_{r\Sigma} f(r)\,d(r\Sigma) \right\} dr,$$

où l'ensemble $r\Sigma$ se trouve sur la sphère de rayon r. Puisque la fonction $f(r)$ est constante sur cette sphère, l'intégrale intérieure vaut, selon *3.63c*:

$$f(r)\,|\,r\Sigma\,| = f(r)\,r^{n-1}\,|\,\Sigma\,|.$$

Définitivement, nous avons

$$\int_{G_m} f(r)\,dx = |\,\Sigma\,| \int_{r=a}^{m} f(r)\,r^{n-1}\,dr.$$

Pour juger sur la convergence de l'intégrale (1), nous avons à faire tendre m vers $+\infty$. La question de la convergence de cette intégrale se réduit donc à la convergence dans R_1 de l'intégrale impropre

$$\int_a^{\infty} f(r)\,r^{n-1}\,dr.$$

Par exemple, en utilisant le résultat *U11.11a*, on peut dire que l'intégrale

$$\int_G \frac{dx}{r^{\alpha}}$$

converge pour $\alpha > n$ et diverge pour $\alpha \leqslant n$. Notons que le résultat ne dépend pas du choix de l'ensemble Σ sur la sphère unité dans R_n qui détermine le domaine G (pourvu que l'ensemble Σ ait une aire positive).

Selon le critère de comparaison *3.72b*, c'est dans les mêmes conditions que converge ou diverge l'intégrale

$$\int_G \theta(x)\,\frac{dx}{r^{\alpha}}. \tag{2}$$

où $\theta(x)$ est une fonction admissible ayant une limite positive pour $|x| \to \infty$. Si, pour $|x| \to \infty$, la fonction $\theta(x)$ n'est que bornée, on ne peut affirmer d'après le critère de comparaison *3.72b* que l'intégrale (2) converge pour $\alpha > n$.

b. Soit $f(r) \geqslant 0$ une fonction donnée sur un intervalle $0 < r \leqslant b$, continue par morceaux sur tout intervalle $a \leqslant r \leqslant b$ $(a > 0)$, et éventuellement non bornée pour $r \to 0$. En posant $r^2 = \sum_{i=1}^{n} x_i^2 = |x|^2$ nous obtenons une fonction admissible $f(r) = f\left(\sqrt{\sum_{i=1}^{n} x_i^2}\right)$ définie pour $0 < |x| \leqslant b$ dans R_n. Considérons-la dans le domaine $G = \{x \in R_n,\ 0 < |x| \leqslant b,\ \frac{x}{|x|} \in \Sigma\}$ où Σ est un ensemble de mesure positive sur la sphère unité dans R_n, et discutons la convergence de l'intégrale impropre de deuxième espèce

$$\int_G f(r)\, dx. \tag{3}$$

Choisissons pour suite exhaustive les domaines

$$G_m = \left\{x \in R_n : \frac{1}{m} \leqslant |x| \leqslant b,\ \frac{x}{|x|} \in \Sigma\right\}.$$

D'une manière analogue à a, nous avons

$$\int_{G_m} f(r)\, dx = |\Sigma| \int_{r=1/m}^{b} f(r)\, r^{n-1}\, dr.$$

Ainsi, le problème de convergence de l'intégrale (3) se réduit à celui de convergence dans R_1 de l'intégrale impropre de deuxième espèce

$$\int_0^b f(r)\, r^{n-1}\, dr.$$

Par exemple, en utilisant les résultats *U11.22a* nous voyons que l'intégrale

$$\int_G \frac{dx}{r^\alpha}$$

converge pour $\alpha < n$ et diverge pour $\alpha \geqslant n$.

Selon le critère de comparaison *3.72b*, c'est dans les mêmes conditions que converge ou diverge l'intégrale

$$\int_G \theta(x)\, \frac{dx}{r^\alpha}, \tag{4}$$

où $\theta(x)$ est une fonction admissible ayant une limite positive pour $x \to 0$. Si la fonction $\theta(x)$ n'est que bornée pour $x \to 0$, alors le critère de comparaison *3.72b* ne sert qu'à établir la convergence de l'intégrale (4) pour $\alpha < n$.

c. Comme conséquence de *a* et *b* nous obtenons: *l'intégrale de troisième espèce de la fonction* $1/r^\alpha$ *sur un « angle solide »* $G = \{x \in R_n : \frac{x}{|x|} \in \Sigma\}$ (où Σ, comme avant, est un ensemble de mesure positive sur la sphère unité dans R_n) *n'existe pour aucun* α.

d. Il s'agit de discuter la convergence de l'intégrale

$$\iint\limits_{G_i} \frac{x^2 - y^2}{r^4} dx\, dy \qquad (5)$$

sur chacun des domaines G_i ($i = 1, 2, 3$) de la fig. 3.7-1.

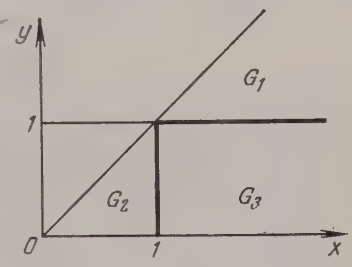


Fig. 3.7-1

Considérons le domaine G_1.

Pour tout ε positif et pour ρ suffisamment grand, le secteur $\{x,\ y \in R_2 :\ \varepsilon < y/x < 1 - \varepsilon,\ x > \rho\}$ est inclus dans ce domaine. Dans le secteur mentionné, la fonction à intégrer est au moins égale à C/r^2, avec un $C > 0$. Par conséquent, l'intégrale (5) sur ce secteur et, a fortiori, sur le domaine G_1 diverge en vertu de *a* et du critère de comparaison *3.72b*. Le domaine G_2 contient le secteur $\{x,\ y \in R_2 :\ 0 \leqslant y/x \leqslant 1 - \varepsilon,\ r \leqslant 1\}$; dans ce secteur, la fonction à intégrer est toujours au moins égale à C/r^2, avec un $C > 0$, et l'intégrale (5) sur G_2 diverge. Quant au domaine G_3, on peut choisir là-dedans une suite exhaustive de rectangles de plus en plus grands $G_m = \{x,\ y \in R_2 :\ 1 \leqslant x \leqslant m,\ 0 \leqslant y \leqslant 1\}$. On a alors:

$$\iint\limits_{G_m} \frac{x^2 - y^2}{r^4} dx\, dy \leqslant \iint\limits_{G_m} \frac{x^2}{r^4} dx\, dy \leqslant \iint\limits_{G_m} \frac{x^2}{x^4} dx\, dy =$$

$$= \iint\limits_{G_m} \frac{dx\, dy}{x^2} = \int\limits_0^1 \left\{ \int\limits_1^m \frac{dx}{x^2} \right\} dy = 1 - \frac{1}{m}.$$

Le dernier membre étant borné pour $m \to \infty$, l'intégrale (5) sur le domaine G_3 s'avère convergente.

e. Considérons, dans un domaine fermé borné $G \subset R_n$, une surface $S = \{x = \varphi(u),\ u \in U \subset R_k\}$ déterminée par une fonction deux fois dérivable et sans singularités $\varphi(u)$ (le dernier veut dire que la matrice $\varphi'(u)$ n'est pas dégénérée). Désignons par $\rho = \rho(x, S)$ la distance d'un point x à la surface S. On se demande, pour quels α la fonction $1/\rho^\alpha(x, S)$ est intégrable sur le domaine G.

Appliquons le principe de localisation *3.72e*, selon lequel il suffit de discuter l'existence de l'intégrale sur les voisinages arbitrairement petits de chaque point $x \in G$. Tout point $a \notin S$ admet un

voisinage sur lequel la fonction $1/\rho^{\alpha}(x, S)$ est bornée, donc intégrable. Considérons un point $a \in S$. D'après *3.65g*, il existe une couche $W_h(S)$ contenant un voisinage $V(a)$ et décrite par les paramètree

$$(u, v) = (u_1, \ldots, u_k; v_{k+1}, \ldots, v_n), \quad u \in U(a) \subset U,$$
$$v \in Q_h = \{v : |v|^2 = \Sigma v_j^2 \leqslant h^2\},$$

de façon que les fonctions

$$x_1 = x_1(u, v),$$
$$\ldots\ldots\ldots$$
$$x_n = x_n(u, v)$$

fournissent, pour $v_{k+1} = \ldots = v_n = 0$, la représentation d'une partie $S(a)$ de la surface S (dans le voisinage mentionné du point a) et, pour $u = (u_1, \ldots, u_k)$ fixe, une application isométrique de la boule $(n-k)$-dimensionnelle $|v| \leqslant h^2$ sur la boule de rayon h et de centre au point a disposée dans le sous-espace $(n-k)$-dimensionnel normal à la surface S. L'intégrale de n'importe quelle fonction (bornée) $f(x)$ sur cette couche s'écrit d'après la règle *3.58*:

$$\int\limits_{W_h(S)} f(x)\, dx = \int\limits_{U(a)\times Q_h} f(x(u, v)) \left| \frac{\partial(x_1, \ldots, x_n)}{\partial(u_1, \ldots, v_n)} \right| du\, dv.$$

La quantité $\rho^2(x, S)$ coïncide dans notre cas avec $\sum\limits_{k+1}^{n} v_j^2$. Choisissons pour suite exhaustive les domaines G_m de la forme

$$G_m = \left\{ x \in W_h(S) : x = x(u, v),\ u \in U(a),\ \frac{1}{m} \leqslant |v| \leqslant h \right\}.$$

Nous avons:

$$\int\limits_{G_m} \frac{dx}{\rho^{\alpha}(x, S)} = \int\limits_{\frac{1}{m} \leqslant |v| \leqslant h} \frac{1}{\rho^{\alpha}(x, S)} \left\{ \int\limits_{U(a)} \left| \frac{\partial(x_1, \ldots, x_n)}{\partial(u_1, \ldots, v_n)} \right| du \right\} dv =$$

$$= \int\limits_{\frac{1}{m} \leqslant |v| \leqslant h} \frac{\Phi(v)}{|v|^{\alpha}}\, dv, \qquad (6)$$

où

$$\Phi(v) = \int\limits_{U(a)} \left| \frac{\partial(x_1, \ldots, x_n)}{\partial(u_1, \ldots, v_n)} \right| du.$$

La fonction $\Phi(v)$ a pour limite, lorsque $|v| = \sqrt{\sum\limits_{k+1}^{n} v_j^2} \to 0$, l'aire de la surface $S(a)$ (*3.65e*). Conformément à *3.73b*, l'intégrale (6) a une limite pour $m \to \infty$ si, et seulement si, $\alpha < n - k$. Donc,

la fonction $1/\rho^{\alpha}$ (x, S) *s'avère intégrable sur le domaine* G *si* $\alpha < n - k$, *et dans ce cas seulement.*

3.74. Intégrales impropres dépendant d'un paramètre.

a. La théorie des intégrales impropres dépendant d'un paramètre, que l'on a développée dans *U11.4* pour un domaine d'intégration unidimensionnel, se généralise, sans complication, au cas de domaines multidimensionnels. Considérons une intégrale impropre de première espèce

$$I(\lambda) = \int_G f(x, \lambda)\, dx \tag{1}$$

qui converge pour toutes les valeurs du paramètre λ d'un ensemble Λ. L'intégrale (1) est dite *uniformément convergente sur* Λ si, pour tout $\varepsilon > 0$, il existe $\rho > 0$ tel que, étant donné un domaine borné $Q \subset G$ qui ne contient aucun point de la boule $V_\rho = \{x \in R_n : |x| \leqslant \rho\}$, a lieu la relation

$$\left| \int_Q f(x, \lambda)\, dx \right| < \varepsilon.$$

b. L'intégrale impropre de deuxième espèce dépendant d'un paramètre $\lambda \in \Lambda$

$$I(\lambda) = \int_G f(x, \lambda)\, dx \tag{2}$$

est traitée d'une manière analogue. On suppose que la fonction $f(x, \lambda)$ soit, pour tout $\lambda \in \Lambda$, bornée et continue par morceaux dans un domaine fermé borné $G \subset R_n$, en dehors de chaque voisinage d'un ensemble négligeable fixe (ne dépendant pas de λ) $Z \subset G$. L'intégrale impropre (2) est dite *uniformément convergente sur* Λ si, pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que, étant donné un ensemble Q contenant l'ensemble Z strictement dans son intérieur et contenu lui-même dans le δ-voisinage de l'ensemble Z, a lieu l'inégalité

$$\left| \int_Q f(x, \lambda)\, dx \right| < \varepsilon.$$

c. La définition d'une intégrale de troisième espèce uniformément convergente sur Λ est construite d'une manière analogue. Nous n'insistons pas là-dessus, faute d'applications ultérieures.

Les théorèmes suivants pour l'intégrale de première espèce se démontrent de façon parfaitement analogue au cas de domaine unidimensionnel (*U11.43-45*).

d. *Soient* Λ *un espace métrique et* $f(x, \lambda)$ *une fonction uniformément continue sur tout produit* $\Lambda \times (G \cap V_\rho)$. *Si l'intégrale* (*1*) *converge uniformément sur* Λ, *alors* $I(\lambda)$ *est une fonction continue de* λ.

e. *Si* Λ *est un espace chargé,* $f(x, \lambda)$ *une fonction continue par rapport à l'ensemble* (x, λ) *sur tout produit* $\Lambda \times (G \cap V_\rho)$, *et si l'intégrale* (*1*) *converge uniformément sur* Λ, *alors l'intégrale impropre de la fonction*

$$g(x) = \int_\Lambda f(x, \lambda)\, d\lambda$$

sur le domaine G converge, et l'on a

$$\int_\Lambda \left\{ \int_G f(x, \lambda)\, dx \right\} d\lambda = \int_G \left\{ \int_\Lambda f(x, \lambda)\, d\lambda \right\} dx.$$

f. *Si* Λ *est un espace vectoriel normé, la fonction* $f(x, \lambda)$ *possède une dérivée* $\frac{\partial f(x, \lambda)}{\partial \lambda}$, *continue par rapport à* (x, λ) *sur tout produit* $\Lambda \times (G \cap V_\rho)$, *et l'intégrale*

$$\int_G \frac{\partial f(x, \lambda)}{\partial \lambda}\, dx$$

converge uniformément sur Λ, *alors l'intégrale* (*1*), *dès qu'elle existe pour au moins un* $\lambda = \lambda_0 \in \Lambda$, *existe pour tout* $\lambda \in \Lambda$, *représente une fonction dérivable, et l'on a*

$$\frac{d}{d\lambda} \int_G f(x, \lambda)\, dx = \int_G \frac{\partial f(x, \lambda)}{\partial \lambda}\, dx.$$

Pour démontrer ces propositions, on applique, au lieu des théorèmes *U9.84* et *U9.77* utilisés dans *U11.45*, les théorèmes *1.46* et *3.35c* où il s'agit non pas de dérivées par rapport à R_1, mais de dérivées par rapport à un espace normé.

g. De même que dans *U11.47a*, la plus simple condition suffisante de la convergence uniforme de l'intégrale (1) consiste en l'existence d'*une majorante intégrable*, i.e. d'une fonction admissible $F(x) \geqslant 0$ qui vérifie les conditions:

$$F(x) \geqslant |f(x, \lambda)|, \quad \text{quel que soit} \quad \lambda \in \Lambda;$$

$$\int_G F(x)\, dx < \infty.$$

h. Pour les intégrales de deuxième et de troisième espèces, les théorèmes analogues à *c-g* peuvent être démontrés avec les mêmes procédés.

3.75. Réduction d'une intégrale impropre multiple à des intégrales simples. Pour ne pas encombrer l'exposé, nous étudions ici le cas $n = 2$; le cas général s'épuise d'une manière analogue.

a. Dans l'espace R_2, considérons une fonction non négative et continue par morceaux $f(x, y)$. Supposons qu'elle admette, sur tout intervalle fini $-c \leqslant y \leqslant c$, une *majorante intégrable*, i.e. une fonction continue par morceaux $F_c(x) \geqslant 0$ telle que

$$\int_{-\infty}^{\infty} F_c(x)\,dx < \infty, \quad f(x, y) \leqslant F_c(x) \tag{1}$$

$$(-\infty < x < \infty, \ -c \leqslant y \leqslant c).$$

Supposons encore que la fonction

$$\Phi(y) = \int_{-\infty}^{\infty} f(x, y)\,dx \tag{2}$$

(dont l'existence est garantie par le critère de comparaison *3.72b*) soit continue par morceaux sur tout intervalle $-c \leqslant y \leqslant c$.

Théorème. *Dans les conditions formulées, l'existence de l'une quelconque des intégrales suivantes implique l'existence de l'autre et leur égalité :*

$$I = \iint_{R_2} f(x, y)\,dx\,dy, \quad I_1 = \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} f(x, y)\,dy \right\} dy. \tag{3}$$

Démonstration. L'existence d'une majorante intégrable implique la convergence uniforme sur tout intervalle $-c \leqslant y \leqslant c$ de l'intégrale (2) et l'égalité

$$\int_{-c}^{c} \left\{ \int_{-\infty}^{\infty} f(x, y)\,dx \right\} dy = \int_{-\infty}^{\infty} \left\{ \int_{-c}^{c} f(x, y)\,dy \right\} dx =$$

$$= \lim_{a \to \infty} \int_{-a}^{a} \left\{ \int_{-c}^{c} f(x, y)\,dy \right\} dx = \lim_{a \to \infty} \int_{-a}^{a} \int_{-c}^{c} f(x, y)\,dx\,dy. \tag{4}$$

Si la première des intégrales (3) converge, le second membre de (4) reste majoré par le nombre I pour tous les c ; cela implique l'existence de la deuxième intégrale et l'inégalité $I_1 \leqslant I$. Si converge la deuxième intégrale (3), alors le premier membre de (4) est majoré par le nombre I_1, quel que soit c ; par conséquent, quels que soient a et c, le même nombre I_1 sert de majorant à l'intégrale

$$\int_{-a}^{a} \int_{-c}^{c} f(x, y)\,dx\,dy,$$

d'où l'existence de la première intégrale (3) et la relation $I \leqslant I_1$. Ainsi, l'existence de n'importe quelle des intégrales (3) implique celle de l'autre et deux relations $I \leqslant I_1$, $I_1 \leqslant I$, donc l'égalité $I = I_1$, ce qu'il fallait.

b. Soient $g(x) \geqslant 0$ $(-\infty < x < \infty)$ et $h(y) \geqslant 0$ $(-\infty < y < \infty)$ deux fonctions continues par morceaux et telles que les intégrales

$$\int_{-a}^{a} g(x)\,dx, \qquad \int_{-c}^{c} h(y)\,dy$$

sont positives pour certains a et c. Posons $f(x, y) = g(x)\,h(y)$.

Théorème. *Dans les hypothèses formulées, l'existence de l'intégrale*

$$I = \iint_{R_2} f(x, y)\,dx\,dy \tag{5}$$

est équivalente à l'existence de deux intégrales

$$I_1 = \int_{-\infty}^{\infty} g(x)\,dx, \quad I_2 = \int_{-\infty}^{\infty} h(y)\,dy \tag{6}$$

et a pour conséquence l'égalité

$$\iint_{R_2} f(x, y)\,dx\,dy = \int_{-\infty}^{\infty} g(x)\,dx \int_{-\infty}^{\infty} h(y)\,dy. \tag{7}$$

Démonstration. Pour n'importe quels a et c, on a

$$\int_{-a}^{a}\int_{-c}^{c} f(x, y)\,dx\,dy = \int_{-a}^{a} g(x)\,dx \int_{-c}^{c} h(y)\,dy. \tag{8}$$

Si l'intégrale (5) existe, le premier membre de (8) est majoré par le nombre I quels que soient a et c. En passant à la limite pour $a \to \infty$, $c = c_0$, nous établissons l'existence de la première intégrale (6) ; en passant à la limite pour $c \to \infty$, nous établissons l'existence de la deuxième intégrale et l'inégalité $I \geqslant I_1 I_2$. Si les intégrales (6) existent, alors le second membre de (8) est majoré par le nombre $I_1 I_2$, d'où l'existence de l'intégrale (5) et l'inégalité $I \leqslant I_1 I_2$. Dans tous les cas ont lieu les deux inégalités $I \leqslant I_1 I_2$, $I_1 I_2 \leqslant I$, i.e. l'égalité $I = I_1 I_2$, ce qu'il fallait.

3.76. Changement de variables dans les intégrales impropres. Nous nous bornons toujours au cas $n = 2$ et à une fonction non négative et continue par morceaux $f(x, y)$.

a. Soit

$$I = \iint_{G} f(x, y)\,dx\,dy \tag{1}$$

une intégrale convergente. Supposons que le domaine G (borné ou non) dans le plan x, y soit l'image par une application dérivable

$x = x(u, v)$, $y = y(u, v)$ d'un domaine Q (également borné ou non) dans le plan u, v. Montrons que l'intégrale

$$\iint_G f(x(u, v),\ y(u, v)) \left| \frac{\partial(x, y)}{\partial(u, v)} \right| du\, dv \tag{2}$$

existe et que sa valeur coïncide avec celle de l'intégrale I.

Une suite exhaustive $G_1 \subset G_2 \subset \ldots \subset G$ est, évidemment, transformée par l'application $x = x(u, v)$, $y = y(u, v)$ d'une suite exhaustive $Q_1 \subset Q_2 \subset \ldots \subset Q$. En appliquant la règle *3.57* de changement de variables au domaine G_m nous trouvons

$$\iint_{G_m} f(x, y)\, dx\, dy = \iint_{Q_m} f(x(u, v),\ y(u, v)) \left| \frac{\partial(x, y)}{\partial(u, v)} \right| du\, dv.$$

Faisons tendre m vers $-\infty$. Le premier membre de la dernière égalité tendra, en vertu de la convergence de l'intégrale (1), vers la limite I; par conséquent, le second membre aura la même limite, d'où, conformément à *3.72a*, l'existence de l'intégrale (2) et sa coïncidence avec l'intégrale (1).

b. Exemple. En partant de la définition (*U11.51*) de la fonction gamma

$$\Gamma(a) = \int_0^\infty e^{-x} x^{a-1}\, dx$$

et de *3.75b*, nous voyons que

$$\Gamma(a)\,\Gamma(b) = \int_0^\infty e^{-x} x^{a-1}\, dx \int_0^\infty e^{-y} y^{b-1}\, dy = \iint_{R_2} e^{-x-y} x^{a-1} y^{b-1}\, dx\, dy. \tag{3}$$

Passons aux nouvelles coordonnées

$$x = u(1 - v), \qquad y = uv. \tag{4}$$

L'application (4) transforme le premier quadrant du plan x, y en la demi-bande $0 \leqslant u < \infty$, $0 \leqslant v \leqslant 1$ du plan u, v et vice versa. On a

$$\left| \frac{\partial(x, y)}{\partial(u, v)} \right| = \begin{vmatrix} 1-v & v \\ -u & u \end{vmatrix} = u.$$

En vertu de *a*:

$$\Gamma(a)\,\Gamma(b) = \int_{u=0}^{\infty} \int_{v=0}^{1} e^{-v} u^{a+b-2} (1-v)^{a-1} v^{b-1} u\, du\, dv =$$

$$= \int_0^\infty e^{-u} u^{a+b-1}\, du \int_0^1 v^{b-1} (1-v)^{a-1}\, dv = \Gamma(a+b)\, \mathrm{B}(a, b);$$

nous obtenons, encore une fois, la relation entre les fonctions gamma et bêta. La première déduction *U11.53* qui n'utilise que les fonctions d'une variable s'avère plus compliquée que celle-ci basée sur les fonctions de deux variables.

3.77. Intégrales impropres avec une singularité variable.

a. Dans la suite, nous aurons souvent affaire à des intégrales *avec une singularité variable* qui sont de la forme:

$$F(y)=\int_G \frac{f(x, y)\,dx}{|x-y|^{\alpha}}, \tag{1}$$

où $|x-y|$ est la distance du point $x \in R_n$ au point $y \in R_n$, $G \subset R_n$ étant un domaine borné. La fonction $f(x, y)$ est supposée bornée et continue dans un domaine $G \times H$, $H \subset R_n$.

Pour $y \in G$ et $\alpha > 0$, l'intégrale (1) est une intégrale impropre de deuxième espèce et, d'après *3.73b*, converge (absolument) si $\alpha < n$.

Bien que l'intégrale (1) soit une intégrale impropre fonction d'un paramètre, la théorie générale des intégrales dépendant d'un paramètre qu'on a exposée dans *3.74* n'y est pas applicable, puisque la singularité dans l'intégrale (1) n'est pas fixe, mais dépend de la position du paramètre y. Pour cette raison, nous établirons des propriétés de l'intégrale (1) par une analyse directe, sans nous adresser à la théorie générale *3.74*.

b. Lemme. *Si* $\alpha < n$, *alors, pour tout* $\varepsilon > 0$, *il existe* $\delta_0 > 0$ *tel que quelle que soit une boule* V_δ *de rayon* $\delta \leqslant \delta_0$, *on a la relation*

$$\left| \int_{V_\delta} \frac{f(x, y)\,dx}{|x-y|^{\alpha}} \right| < \varepsilon. \tag{2}$$

Démonstration. La fonction $f(x, y)$ est continue par rapport à l'ensemble x, y, donc bornée, de sorte que, par exemple, $|f(x, y)| \leqslant C$. L'intégrale de $\frac{1}{|x-y|^{\alpha}}$ étant convergente, on peut trouver, d'après $\varepsilon > 0$ donné, un nombre $\tau > 0$ tel que, pour la boule $V_\tau(y)$ de rayon τ et de centre au point y, on a

$$\left| \int_{V_\tau(y)} \frac{f(x, y)}{|x-y|^{\alpha}}\,dx \right| \leqslant C \int_{V_\tau(y)} \frac{dx}{|x-y|^{\alpha}} \leqslant \varepsilon,$$

le nombre τ ne dépendant pas du choix du point y. Il reste à considérer les boules centrées aux autres points. Si une boule $V_\rho(y_1)$ se trouve à l'intérieur de la boule $V_\tau(y)$, alors l'inégalité (2) est, évidemment, valable pour $V_\delta = V_\rho(y_1)$. Ensuite, à l'extérieur de la boule $V_{\tau/3}(y)$, la fonction $\frac{f(x, y)}{|x-y|^{\alpha}}$ est majorée en module par un nombre K. Par conséquent, son intégrale sur n'importe quelle boule V_ρ de centre à l'extérieur de la boule $V_{2\tau/3}(y)$ et de rayon $\rho < \tau/3$ (une telle boule est disjointe avec la boule $V_{\tau/3}(y)$) ne dépasse pas la quantité $K|V_\rho| = K\Omega_n\rho^n$ (Ω_n étant le volume de la boule unité dans R_n) qui est inférieure à ε pour ρ suffisamment

petit, par exemple pour $\rho \leqslant \delta \leqslant \tau/3$. Or, du moment qu'une boule de rayon $\rho < \tau/3$ ne soit pas entièrement incluse dans $V_\tau(\rho)$, elle a son centre à l'extérieur de la boule $V_{2\tau/3}(y)$. Par conséquent, quelle que soit une boule $V_\rho(y_1)$ de rayon $\rho < \delta$, l'inégalité (2) a lieu, ce qu'il fallait démontrer.

c. Ici nous établissons un théorème sur la continuité de la fonction $F(y)$ (1).

Théorème. *Si $f(x, y)$ est continue par rapport à l'ensemble des variables x, y, et si $\alpha < n$, alors la fonction*

$$F(y) = \int_G \frac{f(x, y)}{|x-y|^\alpha}\, dx$$

est continue.

Démonstration. En appliquant le lemme *b* trouvons, pour $\varepsilon > 0$ donné, un $\delta > 0$ de façon à avoir, pour toute boule V_δ de rayon δ, l'inégalité

$$\left| \int_{V_\delta} \frac{f(x, y)}{|x-y|^\alpha}\, dx \right| < \frac{\varepsilon}{3}. \tag{3}$$

Désignons $\varphi(x, y) = \frac{f(x, y)}{|x-y|^\alpha}$. Fixons arbitrairement un point $y_0 \in H$. Soit $|y - y_0| < \delta/2$. Alors on a

$$|F(y) - F(y_0)| = \left| \int_G [\varphi(x, y) - \varphi(x, y_0)]\, dx \right| \leqslant$$

$$\leqslant \left| \int_{V_\delta(y_0)} [\varphi(x, y) - \varphi(x, y_0)]\, dx \right| + \left| \int_{G - V_\delta(y_0)} [\varphi(x, y) - \varphi(x, y_0)]\, dx \right| \leqslant$$

$$\leqslant \int_{V_\delta(y_0)} |\varphi(x, y)|\, dx + \int_{V_\delta(y_0)} |\varphi(x, y_0)|\, dx + \left| \int_{G - V_\delta(y_0)} [\varphi(x, y) - \right.$$

$$\left. - \varphi(x, y_0)]\, dx \right|. \tag{4}$$

La fonction $\varphi(x, y)$ est continue par rapport à l'ensemble des variables x, y dans le domaine $\{x \in G - V_\delta(y_0),\ y \in V_\delta(y_0)\}$; d'après le théorème *3.35a*, la dernière intégrale dans le membre droit de (4) représente une fonction continue de y. En particulier, pour ε donné, on peut indiquer un $\rho < \delta/2$ tel que

$$\left| \int_{G - V_\rho(y_0)} [\varphi(x, y) - \varphi(x, y_0)]\, dx \right| < \frac{\varepsilon}{3}, \tag{5}$$

dès que $|y - y_0| < \rho$.

Il résulte des inégalités (3)-(5):

$$|F(y)-F(y_0)|<\frac{\varepsilon}{3}+\frac{\varepsilon}{3}+\frac{\varepsilon}{3}=\varepsilon$$

pour $|y-y_0|<\rho$.

Ainsi, la fonction $F(y)$ est continue au point y_0. Comme y_0 est un point arbitraire du domaine H, $F(y)$ est continue partout, ce qu'il fallait.

d. Intégration d'une intégrale impropre avec une singularité variable par rapport au paramètre.

Théorème. *Si $\alpha<n$ et le point y parcourt une surface k-dimensionnelle S d'aire finie* ($1\leqslant k<n$), *alors*

$$\int_S\left\{\int_G\frac{f(x,y)}{|x-y|^\alpha}dx\right\}dy=\int_G\left\{\int_S\frac{f(x,y)}{|x-y|^\alpha}dy\right\}dx. \tag{6}$$

Démonstration. Remarquons que, en vertu du théorème *c*, l'intégrale intérieure dans le premier membre de (6) est une fonction continue de y, ce qui garantit l'existence de toute l'intégrale du premier membre. L'intégrale intérieure dans le second membre, en tant que fonction de x, n'est définie qu'aux points x n'appartenant pas à la surface S; lorsque le point x s'approche de la surface S, la valeur de l'intégrale intérieure croît indéfiniment, en général, de sorte que l'intégrale itérée du second membre s'avère impropre, son ensemble des points singuliers coïncidant avec la surface S. Pour cette raison, l'existence de l'intégrale dans le second membre n'est pas évidente à priori; elle découlera, de même que la relation (6), de nos raisonnements.

Notons, comme plus haut, $\varphi(x,y)=\frac{f(x,y)}{|x-y|^\alpha}$. Pour démontrer l'existence de l'intégrale dans le second membre de (6), nous avons à considérer la suite d'intégrales propres

$$I_m=\int_{G-B_m}\left\{\int_S\varphi(x,y)\,dy\right\}dx\quad(m=1,2,\ldots),$$

où les domaines B_m se contractent vers l'ensemble S. Dans chaque intégrale I_m on peut, d'après le théorème *3.35b*, permuter l'ordre d'intégrations:

$$I_m=\int_S\left\{\int_{G-B_m}\varphi(x,y)\,dx\right\}dy.$$

Désignons l'intégrale intérieure par $F_m(y)$. Montrons que la fonction $F_m(y)$ converge uniformément dans H, pour $m\to\infty$, vers

la fonction

$$F(y) = \int_G \varphi(x, y)\, dx.$$

La différence de $F(y)$ et $F_m(y)$ représente une intégrale sur le domaine B_m:

$$F(y) - F_m(y) = \int_{B_m} \varphi(x, y)\, dx. \tag{7}$$

Pour $\varepsilon > 0$ donné, trouvons en appliquant le lemme b un $\delta > 0$ de façon à avoir, pour chaque boule V_ρ de rayon $\rho \leqslant \delta$, l'inégalité

$$\int_{V_\rho} |\varphi(x, y)|\, dx < \frac{\varepsilon}{2}.$$

Décomposons l'intégrale (7) en somme de deux intégrales: la première sur la partie B'_m du domaine B_m qui est dans la boule $V_\rho(y)$, la deuxième sur la partie B''_m qui reste. Vu le choix de δ, nous avons

$$\int_{B'_m} |\varphi(x, y)|\, dx < \frac{\varepsilon}{2}$$

indépendamment de la valeur de m. Dans le domaine B''_m la fonction $\varphi(x, y)$ est majorée (par une constante qui ne dépend pas de m); comme le volume du domaine $B_m \supset B''_m$ tend vers 0 lorsque $m \to \infty$, il existe un N tel que, pour $m > N$, on a:

$$\int_{B''_m} |\varphi(x, y)|\, dx < \frac{\varepsilon}{2}.$$

Ainsi, pour tous les $m > N$:

$$|F(y) - F_m(y)| \leqslant \int_{B'_m} |\varphi(x, y)|\, dx + \int_{B''_m} |\varphi(x, y)|\, dx < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

cette estimation étant indépendante de la position du point $y \in H$. Nous voyons donc que la suite $F_m(y)$ converge uniformément vers $F(y)$. Il en résulte d'après *3.14 f* que I_m tend vers la limite

$$\int_S F(y)\, dy = \int_S \left\{ \int_G \varphi(x, y)\, dx \right\} dy.$$

Nous avons démontré l'existence de l'intégrale dans le second membre de l'égalité (6) ainsi que cette égalité elle-même.

e. L'hypothèse de la continuité de la fonction $f(x, y)$ par rapport aux variables $x \in G$, $y \in H$, dans les théorèmes c, d peut être remplacée par les hypothèses plus faibles:

1) la fonction $f(x, y)$ est bornée pour $x \in G$, $y \in H$;

2) quel que soit $\delta > 0$, la fonction $f(x, y)$ est continue par rapport à l'ensemble x, y dans le domaine $(G \times H) = \{x, y: |x - y| < \delta\}$.

En effet, dans les dernières hypothèses, on peut écrire

$$\frac{f(x, y)}{|x-y|^{\alpha}} = \frac{f(x, y)\,|x-y|^{\gamma}}{|x-y|^{\alpha+\gamma}}.$$

Lorsque $\gamma > 0$ est suffisamment petit, l'exposant dans le dénominateur ne dépasse pas n; or, le numérateur se trouve continu par rapport à l'ensemble des variables x, y partout dans $G \times H$, et les théorèmes c, d restent valables.

f. Il est aisé de prouver que les théorèmes c, d avec le complément e restent en vigueur pour les fonctions $f(x, y)$ à valeurs dans un espace vectoriel normé. On peut les appliquer, par exemple, au produit d'une fonction numérique $f(x, y)$ par le vecteur unitaire $e(x, y)$ dont le sens est celui du point x vers le point y. Ce produit est une fonction vectorielle qui, même pour une fonction $f(x, y)$ continue, n'est en général plus continue pour $x \to y$.

g. Dérivation de l'intégrale impropre avec une singularité variable par rapport aux coordonnées du point paramètre.

Théorème. *Si la fonction $f(x, y)$ est continue par rapport à l'ensemble des variables $x \in G$, $y \in H$ et possède la dérivée $\frac{\partial f(x, y)}{\partial y_j}$ continue par rapport à l'ensemble x, y (ou au moins vérifiant les conditions f), alors pour $\alpha < n - 1$, on a:*

$$\frac{\partial}{\partial y_j} \int_G \frac{f(x, y)}{|x-y|^{\alpha}}\,dx = \int_G \frac{\partial}{\partial y_j} \frac{f(x, y)}{|x-y|^{\alpha}}\,dx. \tag{8}$$

Démonstration. Soit toujours $\varphi(x, y) = \frac{f(x, y)}{|x-y|^{\alpha}}$. Un calcul élémentaire montre que

$$\frac{\partial}{\partial y_j} \frac{1}{|x-y|^{\alpha}} = -\frac{\alpha}{|x-y|^{\alpha+1}} \cdot \frac{y_j - x_j}{|x-y|} = -\frac{\alpha}{|x-y|^{\alpha+1}} \cos \omega_j,$$

où ω_j désigne l'angle que font le vecteur $y - x$ et le j-ème axe de coordonnées. Il en résulte

$$\frac{\partial \varphi(x, y)}{\partial y_j} = -\frac{\alpha \cos \omega_j}{|x-y|^{\alpha+1}} f(x, y) + \frac{1}{|x-y|^{\alpha}} \frac{\partial f(x, y)}{\partial y_j},$$

ce qui, compte tenu du fait que les fonctions $f(x, y)$ et $\frac{\partial f(x,y)}{\partial y_j}$ sont bornées, amène à l'estimation

$$\left|\frac{\partial \varphi(x, y)}{\partial y_j}\right| \leqslant \frac{C_1}{(x-y)^{\alpha+1}},$$

où C_1 est une constante. Selon cette estimation, l'intégrale par rapport à la variable x de la fonction $\frac{\partial \varphi(x, y)}{\partial y_j}$ converge absolument pour $\alpha < n - 1$.

Pour démontrer notre assertion, supposons pour le moment qu'il y ait une fonction $\varphi_\delta(x, y)$ bornée, continue par rapport à l'ensemble des variables x, y et possédant la dérivée $\frac{\partial \varphi_\delta(x, y)}{\partial y_j}$ qui soit continue par rapport à l'ensemble de x, y et qui vérifie les relations

$$\left.\begin{aligned} \lim_{\delta \to 0} \int_G \varphi_\delta(x, y)\, dx &= \int_G \varphi(x, y)\, dx, \\ \lim_{\delta \to 0} \int_G \frac{\partial \varphi_\delta(x, y)}{\partial y_j}\, dx &= \int_G \frac{\partial \varphi(x, y)}{\partial y_j}\, dx \end{aligned}\right\} \tag{9}$$

uniformément par rapport à y.

Soit

$$F(y) = \int_G \varphi(x, y)\, dx, \quad F_\delta(y) = \int_G \varphi_\delta(x, y)\, dx,$$

$$\widetilde{F}(y) = \int_G \frac{\partial \varphi(x, y)}{\partial y_j}\, dx.$$

Vu le théorème *3.35d*, nous avons:

$$\frac{\partial F_\delta(y)}{\partial y_j} = \frac{\partial}{\partial y_j} \int_G \varphi_\delta(x, y)\, dx = \int_G \frac{\partial \varphi_\delta(x, y)}{\partial y_j}\, dx.$$

Les relations (9) prennent la forme

$$\lim_{\delta \to 0} F_\delta(y) = F(y),$$

$$\lim_{\delta \to 0} \frac{\partial F_\delta(y)}{\partial y_j} = \widetilde{F}(y).$$

Il en résulte, selon le théorème sur la dérivation d'une suite de fonctions (1.46), que la fonction $F(y)$ est dérivable en y_j et que

$$\widetilde{F}(y) = \frac{\partial F(y)}{\partial y_j},$$

ce qu'il nous fallait.

Il reste donc à construire une fonction $\varphi_\delta(x, y)$ possédant les propriétés signalées plus haut. Considérons la fonction

$$h_\delta(r) = \begin{cases} 1/r & \text{pour} \quad r \geqslant \delta \\ B - r^2 A & \text{pour} \quad 0 \leqslant r \leqslant \delta, \end{cases}$$

où A et B sont choisis d'une telle manière que le passage de la valeur r par le nombre δ ne viole pas la continuité et la dérivabilité de la fonction $h_\delta(r)$ (fig. 3.7-2). Il est alors évident qu'ont lieu les inégalités

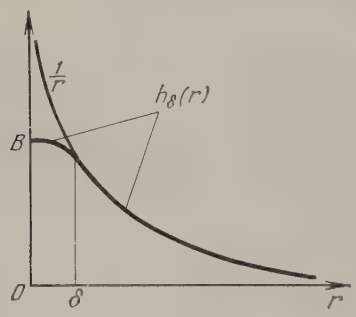


Fig. 3.7-2

$$\left.\begin{aligned} 0 \leqslant h_\delta(r) \leqslant \frac{1}{r} \quad (0 \leqslant r < \infty), \\ |h'_\delta(r)| \leqslant \frac{1}{r^2} \quad (0 \leqslant r < \infty). \end{aligned}\right\} \tag{10}$$

Posons maintenant

$$\varphi_\delta(x, y) = h_\delta^\alpha(|x - y|) f(x, y).$$

Nous avons

$$\frac{\partial \varphi_\delta(x, y)}{\partial y_j} = h_\delta^\alpha(|x-y|) \frac{\partial f(x, y)}{\partial y_j} + \\ + \alpha h_\delta^{\alpha-1}(|x-y|) h'_\delta(|x-y|) \frac{\partial(|x-y|)}{\partial y_j} f(x, y).$$

Il résulte des inégalités (10):

$$|\varphi_\delta(x, y)| \leqslant \frac{C_2}{|x-y|^\alpha}, \quad \left|\frac{\partial \varphi_\delta(x, y)}{\partial y_j}\right| \leqslant \frac{C_3}{|x-y|^{\alpha+1}},$$

C_2 et C_3 étant des constantes. Il est donc possible, pour $\varepsilon > 0$ donné, de choisir δ de façon à ce que l'inégalité

$$|F(y) - F_\delta(y)| = \left|\int_G [\varphi(x, y) - \varphi_\delta(x, y)]\, dx\right| = \\ = \left|\int_{V_\delta(y)} [\varphi(x, y) - \varphi_\delta(x, y)]\, dx\right| \leqslant \\ \leqslant \int_{V_\delta(y)} |\varphi(x, y)|\, dx + \int_{V_\delta(y)} |\varphi_\delta(x, y)|\, dx \leqslant C_4 \int_{V_\delta(y)} \frac{dx}{|x-y|^\alpha} < \varepsilon$$

soit valable indépendamment de la position du point $y \in H$. De même, en diminuant éventuellement $\delta > 0$, on peut obtenir la

relation

$$\left|\widetilde{F}(y)-\frac{\partial F_\delta(y)^{\sim}}{\partial y_j}\right|=\left|\int\limits_G\left[\frac{\partial\varphi(x,y)}{\partial y_j}-\frac{\partial\varphi_\delta(x,y)}{\partial y_j}\right]dx\right|=\left|\int\limits_{V_\rho(y)}\left[\frac{\partial\varphi(x,y)}{\partial y_j}-\right.\right.$$

$$\left.\left.-\frac{\partial\varphi_\delta(x,y)}{\partial y_j}\right]dx\right|\leqslant\int\limits_{V_\rho(y)}\left|\frac{\partial\varphi(x,y)}{\partial y_j}\right|dx+\int\limits_{V_\rho(y)}\left|\frac{\partial\varphi_\delta(x,y)^{\sim}}{\partial y_j}\right|dx\leqslant$$

$$\leqslant C_5\int\limits_{V_\rho(y)}\frac{dx}{|x-y|^{\alpha+1}}<\varepsilon$$

indépendamment de la position du point $y\in H$. Ainsi les relations (9) sont établies et le théorème est démontré.

EXERCICES

1. Réduire à une intégrale unidimensionnelle l'expression

$$I=\int\limits_{x_n=0}^{x}\ldots\int\limits_{x_2=0}^{x_3}\int\limits_{x_1=0}^{x_2}x_1\ldots x_nf(x_1)\,dx_1\ldots dx_n.$$

2. Calculer l'intégrale

$$I=\int\limits_{R_n}\ldots\int e^{-A(x,x)}\,dx,$$

où $A(x, x)$ est une forme quadratique définie positive.

3. Exprimer l'intégrale

$$I=\int\limits_{R_n}\ldots\int f(r)\,dx\qquad\left(r=\sqrt{\sum_{i=1}^{n}x_i^2}\right)$$

par une intégrale unidimensionnelle.

4. Réduire à une intégrale unidimensionnelle l'intégrale de surface

$$I=\int\limits_{S_{n-1}(1)}f[(a,x)]\,dS,$$

où $S_{n-1}(1)$ est une sphère de rayon 1 dans R_n.

5. (Exemple de Schwartz.) Sur un cylindre circulaire droit C (1 pour rayon de base, 1 pour hauteur), on introduit les coordonnées cylindriques naturelles φ et z: $0\leqslant\varphi\leqslant 2\pi$, $0\leqslant z\leqslant 1$. Sur ses sections $z=2kh$, $k=0, 1, 2, \ldots$ on marque les points $\varphi=0, 2\alpha, 4\alpha, \ldots$, et sur les sections $z=(2k+1)h$, $k=0, 1, 2, \ldots$, les points $\varphi=\alpha, 3\alpha, 5\alpha, \ldots$ (h et α sont donnés). Deux points voisins d'une section $z=mh$ et le point de la section $z=(m+1)h$ qui se trouve entre eux déterminent un triangle inscrit dans le cylindre C. Montrer que le polyèdre formé de tous les triangles de ce type s'approche indéfiniment, lorsque h et α tendent vers zéro, de la surface du cylindre, mais que l'aire du polyèdre peut avoir n'importe quelle limite supérieure ou égale à $|C|$.

6. Montrer que, si l'intégrale d'une fonction $f(x)$ (continue par morceaux et localement bornée) étendue sur un domaine non borné $G \subset R_n$ converge, alors elle converge absolument. Comment concilier ce fait et l'existence d'intégrales semi-convergentes sur $[0, \infty)$?

7. Considérer la charge suivante d'un intervalle $[a, b]$: on choisit pour cellules tous les intervalles (ordinaires) possibles, de même que leurs intersections avec l'ensemble P des points rationnels ou bien avec l'ensemble Q des points irrationnels de l'intervalle $[a, b]$; pour mesure d'un intervalle ou de son intersection avec Q on choisit la longueur de l'intervalle, pour celle de l'intersection d'un intervalle avec P on prend 0. Vérifier toutes les propriétés d'une charge et indiquer ses cellules non jordaniennes.

8. Montrer que le centre de gravité d'un cône circulaire droit (dans R_3) de n'importe quel angle au sommet se trouve sur son axe et que sa distance de la base vaut un tiers de la hauteur. Lorsque l'angle au sommet du cône tend vers zéro, le cône devient, à la limite, un segment. Or, le centre de gravité d'un segment se trouve à son milieu. Comment expliquer cette discontinuité ?

9. Désignons par $S_r(k, n, \varepsilon)$ la partie de la sphère $S_r(n) = \{ x \in R_n : |x| = r \}$ déterminée par les inégalités

$$|(b_1, x)| \leqslant \varepsilon, \ldots, |(b_k, x)| \leqslant \varepsilon, \quad \varepsilon > 0$$

où $b_1, \ldots, b_k$ sont des vecteurs orthonormés. Montrer que

$$\lim_{n\to\infty} \frac{|S_r(k, n, \varepsilon)|}{|S_r(n)|} = 1.$$

10. Notons $V_r(k, n, \varepsilon, \rho)$ la partie de la boule $V_r(n) = \{ x \in R_n : |x| \leqslant r \}$ déterminée par les inégalités

$$|(b_1, x)| \leqslant \varepsilon, \ldots, |(b_k, x)| \leqslant \varepsilon, \quad \rho \leqslant |x| \leqslant r$$

où $b_1, \ldots, b_k$ sont des vecteurs orthonormés. Montrer que

$$\lim_{n\to\infty} \frac{|V_r(k, n, \varepsilon, \rho)|}{|V_r(n)|} = 1.$$

11. Les coordonnées sphériques $r, \varphi_1, \varphi_2, \ldots, \varphi_{n-1}$ dans l'espace R_n sont définies par les formules :

$$\begin{aligned}
x_1 &= r \cos \varphi_1, \\
x_2 &= r \sin \varphi_1 \cos \varphi_2, \\
x_3 &= r \sin \varphi_1 \sin \varphi_2 \cos \varphi_3, \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
x_{n-1} &= r \sin \varphi_1 \sin \varphi_2 \ldots \sin \varphi_{n-2} \cos \varphi_{n-1}, \\
x_n &= r \sin \varphi_1 \sin \varphi_2 \ldots \sin \varphi_{n-2} \sin \varphi_{n-1}.
\end{aligned}$$

Montrer que

$$\det \frac{\partial(x_1, \ldots, x_n)}{\partial(r, \varphi_1, \ldots, \varphi_{n-1})} = r^{n-1} \sin^{n-2} \varphi_1 \sin^{n-3} \varphi_2 \ldots \sin \varphi_{n-1}.$$

12. Démontrer que, dans l'espace hilbertien des fonctions $f(x)$ continues sur un intervalle $a \leqslant x \leqslant b$, le carré du volume du parallélépipède construit

sur les vecteurs $f_1(x), \ldots, f_n(x)$ est égal à

$$\frac{1}{n!}\int_a^b \cdots \int_a^b \begin{vmatrix} f_1(x_1) & f_1(x_2) & \ldots & f_1(x_n) \\ f_2(x_1) & f_2(x_2) & \ldots & f_2(x_n) \\ \ldots & \ldots & \ldots & \ldots \\ f_n(x_1) & f_n(x_2) & \ldots & f_n(x_n) \end{vmatrix} dx_1 \ldots dx_n.$$

13. L'expression

$$\int_{R_n} \cdots \int f(x_1, \ldots, x_n)\, e^{-i(x_1\sigma_1 + \ldots + x_n\sigma_n)}\, dx_1 \ldots dx_n = \varphi(\sigma_1, \ldots, \sigma_n)$$

s'appelle *transformation n-uple de Fourier de la fonction* $f(x)$. Montrer que la transformation de Fourier d'une fonction symétrique sphérique est encore une fonction symétrique sphérique.

14. Pour $n = 3$, mettre la transformation de Fourier d'une fonction symétrique sphérique $f(x)$ sous la forme d'une intégrale unidimensionnelle.

15. Pour $n = 3$, montrer que la transformation de Fourier $\varphi(\sigma)$ d'une fonction symétrique sphérique et absolument intégrable $f(x)$ est une fonction dérivable.

16. Montrer que la transformation de Fourier de la fonction

$$e^{-\sum_{j,k=1}^{n} a_{jk}x_j x_k}$$

où la forme $\sum_{j,k=1}^{n} a_{jk}x_j x_k$ est définie positive, vaut $\pi^{n/2} e^{D(\sigma)/(4D)}/\sqrt{D}$, où

$$D = \det \| a_{jk} \|, \quad D(\sigma) = \begin{vmatrix} 0 & \sigma_1 & \ldots & \sigma_n \\ \sigma_1 & a_{11} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots & \ldots \\ \sigma_n & a_{n1} & \ldots & a_{nn} \end{vmatrix}.$$

Historique

L'intégrale double (sur un domaine plan borné) apparaît pour la première fois chez Euler (1770) ; Euler donne la règle de son calcul par réduction à une intégrale itérée. Euler, puis Lagrange, considèrent également les intégrales triples. Chacun d'eux propose certaines règles de changement des variables ; vues de plus près, ces règles se trouvent incomplètes ; la méthode correcte est proposée premièrement par Ostrogradski (1836) pour les intégrales doubles et triples et Jacobi (1841) pour les intégrales de multiplicité arbitraire ; c'est en rapport au problème de changement des variables que Jacobi introduit les déterminants fonctionnels nommés jacobiens par Sylvester.

Une théorie générale de la mesure des volumes (dans l'espace n-dimensionnel) est développée par Jordan dans son « Cours d'Analyse » (1882-1887).

Bien qu'on sache calculer l'aire d'une surface à l'aide des intégrales doubles depuis le XVIII-ème siècle (Euler 1770), sa définition rigoureuse fait défaut encore longtemps ; on croit qu'une telle aire puisse être définie comme limite des aires de polyèdres inscrits, enfin H. Schwartz (1870), par son exemple frappant, fait réfléchir sur la nécessité d'en avoir une définition stricte. Celle-ci est donnée (en termes de recouvrement « en tuiles » et de « couche engendrée par une surface ») par Jordan dans le même Cours.

En 1894, Stieltjes propose une nouvelle notion d'intégrale qui diffère de la notion classique d'intégrale de Riemann par les faits que deux intervalles

égaux sur la droite possèdent des mesures différentes et que même des points isolés sont capables de porter une mesure positive; de même que pour l'intégrale de Riemann, celle de Stieltjes se déduit de sommes intégrales par un passage à la limite. Cette définition est ensuite considérée par nombre d'auteurs. Mais l'apparition, en 1902, de l'intégrale de Lebesgue que Radon étend de la droite à l'espace multidimensionnel en 1913, dirige l'attention des mathématiciens vers des questions liées à une mesure dénombrablement additive, et l'intégrale de Riemann-Stieltjes qui n'exige que l'additivité finie de la mesure (et, dans le cas où la mesure prend des valeurs de deux signes, la condition supplémentaire que la mesure soit de variation bornée) reste à l'ombre. Elle conserve sa valeur dans des questions concernant les fonctions continues; ainsi, F. Riesz obtient en 1909 l'expression générale d'une fonctionnelle linéaire dans l'espace des fonctions continues sur un intervalle fermé borné, puis Radon en fait autant pour les fonctions continues sur un compact n-dimensionnel (1913), sous la forme d'intégrale sur une mesure additive tout court.

Les fonctions additives de domaines se rencontrent bien souvent en physique. Certains savants disent même que seules ces fonctions ont un sens réel, les fonctions d'un point étant une abstraction non adéquate à la réalité (cf., par exemple, l'article de В. И. Смирнов, С. Л. Соболев, *Николай Максимович Гюнтер* dans le livre: Н. М. Гюнтер, Теория потенциала и её применения, ГИТТЛ, 1953). Ainsi, la température et la densité de masse sont des notions abstraites qui correspondent aux notions réelles respectives: quantité de chaleur et de masse dans un volume contenant un point donné.

En 1938-1943, les mesures additives sont étudiées (d'un point de vue plus général) par A. A. Markov et A. D. Alexandrov. Notre construction de l'intégrale est aussi inspirée par les idées de l'intégrale de Riemann-Stieltjes. Quant aux liaisons entre les mesures additive et dénombrablement additive sur un compact n-dimensionnel, voir Г. Е. Шилов и Б. Л. Гуревич. Интеграл, мера и производная, 2-е изд., «Наука», 1967, гл. 2.

Le schéma ici considéré n'est plus valable dans un espace de dimension infinie, ce qui n'est d'ailleurs pas étonnant: une boule de dimension infinie contient une infinité de boules plus petites qui sont deux à deux géométriquement égales et disjointes; par conséquent, si l'on tient à ce que les boules géométriquement égales soient d'un même volume, on se heurte forcément à certaines complications. D'autre part, un volume est l'intégrale de la fonction qui vaut 1 dans le domaine correspondant, donc les difficultés avec les volumes deviennent difficultés dans la théorie d'intégration. En fait, une théorie d'intégration dans les espaces de dimension infinie existe, mais elle diffère si fortement de la théorie ordinaire, que leur réunion en un même exposé n'est pas raisonnable (cf. Г. Е. Шилов и Фан Дык Тинь, Интеграл, мера и производная на линейных пространствах, « Наука », 1967, et la bibliographie qu'on y trouvera).

L'intégration et la dérivation

Les opérations différentielles et intégrales sur les fonctions de plusieurs variables sont en certaines relations qui généralisent la formule de Newton-Leibniz valable pour les fonctions d'une variable. Nous avons déjà vu, dans le chapitre précédent, l'une des relations de ce genre en définissant la densité d'une fonction additive d'un domaine et en construisant cette fonction d'après sa densité. Il en existe d'autres où un rôle essentiel appartient à la frontière du domaine considéré (aux extrémités d'un intervalle dans le cas de fonctions d'une variable). Pour les fonctions de plusieurs variables, ces relations sont exprimées par les formules d'Ostrogradski, de Green et de Stokes. Au début, on les écrivait en forme numérique ou « scalaire ». Plus tard, en rapport avec les problèmes de la physique mathématique, se forme l'analyse dite vectorielle avec ses opérations différentielles spécifiques (gradient, divergence, rotationnel). Il devient clair que les problèmes de l'analyse vectorielle sont juste les problèmes sur la nature des relations entre la dérivation et l'intégration (dans l'espace). Or, le langage de l'analyse vectorielle se montre beaucoup plus commode et expressif que le langage classique « scalaire » des mathématiciens, ce qui provoque la reconstruction « vectorielle » de toute la branche discutée du calcul différentiel et intégral. Quant à l'analyse vectorielle elle-même, elle se dépouille de « vêtements » physiques, et l'assortiment de ses opérations différentielles du premier ordre s'avère naturel et, à un certain sens, complet. Remarquons que l'analyse vectorielle classique, dans sa partie essentielle (rotationnel), s'applique à l'espace tridimensionnel (ou bidimensionnel). Cela suffisait pour ses premières applications à la physique mathématique. Mais les applications suivantes en demandent une généralisation multidimensionnelle, surtout que la théorie mathématique elle-même cherche également les énoncés généraux, valables pour n'importe quelle dimension. *Il s'avère possible de généraliser* l'analyse vectorielle au cas de dimensions supérieures ; comment, on le verra dans le chapitre 7.

§ 4. 1. Formule d'Ostrogradski

4.11. Flux d'un champ vectoriel et son interprétation hydromécanique. Soit $R = R(x)$ un *champ vectoriel* continu, donné dans un domaine G de l'espace R_3 ; cela signifie qu'à tout point $x \in G$ il correspond un vecteur $R(x)$ qui dépend continûment du point x. (La notion de *champ vectoriel* apparue en physique est identique à la notion mathématique de *fonction vectorielle*.) Considérons dans le domaine G une surface S lisse par morceaux, dont le vecteur unitaire normal $m(x)$ varie

continûment sur toute partie lisse de la surface S. L'expression

$$W(R(x), S) = \iint_S (m(x), R(x))\, dS \qquad (1)$$

s'appelle *flux du champ* $R(x)$ *à travers la surface* S.

Cette quantité admet une simple interprétation mécanique. Supposons que le domaine G soit rempli d'un milieu continu en mouvement (d'un « fluide ») et que le vecteur $R(x)$ représente la vitesse du fluide au point x; évaluons la quantité de fluide qui traverse la surface S durant le temps dt.

Soit dS un élément de la surface S; en négligeant la variation du vecteur $R(x)$ sur l'élément dS on peut dire que le fluide qui traverse

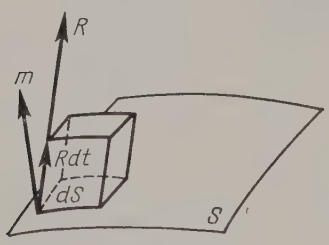


Fig. 4.1-1

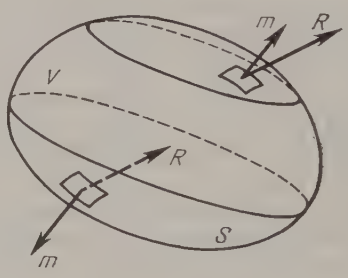


Fig. 4.1-2

dS durant le temps dt remplit le cylindre de base dS et de génératrice $R(x)\,dt$ (fig. 4.1-1). Le volume du cylindre vaut $dS\,|R(x)| \times$ $\times \cos(\widehat{m, R})\,dt = (R, m)\,dS\,dt$.

En sommant les quantités ainsi obtenues sur tous les éléments dS, nous voyons que la quantité de fluide traversant toute la surface S durant le temps dt est

$$dt \iint_S (m, R)\, dS = dt \cdot W(R, S).$$

Définitivement, le flux $W(R, S)$ peut être interprété comme *vitesse* (volumique) *du mouvement de fluide à travers toute la surface* S, ou bien, puisque cette vitesse ne dépend pas du temps, comme *quantité de fluide traversant la surface durant l'unité du temps.*

Ajoutons que, comme $(m(x), R(x)) > 0$ lorsque l'angle formé par $m(x)$ et $R(x)$ est aigu, et $(m(x), R(x)) < 0$, s'il est obtus, l'intégrale (1) représente non pas la quantité absolue du fluide traversant la surface S en l'unité du temps, mais la somme algébrique de la quantité du fluide qui coule dans le sens de la normale (avec le signe +), et de la quantité du fluide coulant dans le sens opposé (avec le signe —).

Considérons le cas où la surface S sert de frontière à un domaine $V \subset G$ et le vecteur de la normale $m(x)$ est dirigé vers l'extérieur de V (fig. 4.1-2). De telles surfaces qui partagent l'espace en deux ensembles, intérieur (domaine V) et extérieur, seront appelées *fermées* (cf. *contour fermé*), ce qui n'est pas à confondre avec le sens métrique du terme « fermé » (contenant ses points d'accumulation). L'intégrale d'une fonction $f(x)$ sur une surface fermée S dans R_3 sera désignée comme suit :

$$\oiint_S f(x)\, dS.$$

En cas de surface fermée, le mouvement du fluide *dans le sens de la normale* signifie que le fluide, dans l'endroit considéré, *s'écoule du domaine* V, et le mouvement dans le sens opposé, que le fluide entre *dans* V. L'intégrale (1) représente alors la différence de la quantité de fluide s'écoulant du domaine en l'unité du temps et de la quantité de fluide qui entre dans le domaine durant le même temps. Si l'intégrale (1) est nulle, ces deux quantités sont égales, sinon elles sont différentes. En particulier, lorsque l'intégrale (1) est positive, la quantité de fluide s'écoulant dépasse celle de fluide entrant. Cela veut dire que le domaine V contient des *sources*, i.e. des endroits où le fluide est créé d'une façon quelconque (comme, par exemple, l'eau apparaît lorsque la neige fond). Dans ce cas, la valeur numérique de l'intégrale (1) peut être interprétée comme débit total des sources dans le domaine V. Si l'intégrale (1) est négative, la quantité de fluide entrant dans V dépasse celle de fluide qui s'en écoule ; il doit y avoir des *puits* dans le domaine, i.e. des endroits où le fluide disparaît (s'évapore, par exemple). Naturellement, on peut dire source au lieu de puits en admettant les quantités négatives de fluide. Pour cette raison, l'intégrale (1) peut être interprétée comme *débit total des sources se trouvant dans le domaine* V, et le rapport de cette intégrale au volume $|V|$ comme *débit moyen des sources dans le domaine* V. Fixons dans le domaine G un point y et considérons une suite $V_1, \ldots, V_k, \ldots$ de domaines limités respectivement par les surfaces $S_1, \ldots, S_k, \ldots$ qui se contracte vers le point y. Le nombre $W(R, V_k)/|V_k|$, on vient de le dire, représente le débit moyen des sources dans le domaine V_k. Supposons que ce débit moyen tende, pour $k \to \infty$, vers une limite $\rho = \rho(y)$ qui ne dépend pas du choix des domaines V_k (pourvu qu'ils se contractent vers le point y). Il est naturel d'appeler le nombre $\rho(y)$ *débit* (ou *densité*) *de source du champ* $R(x)$ *au point* y. Cette limite s'appelle également *divergence du champ* $R(x)$ *au point* y. Ainsi, par définition, on a :

$$\operatorname{div} R(y) = \lim_{V_k \to y} \frac{1}{|V_k|} \oiint_{S_k} (m, R)\, dS \qquad (2)$$

en supposant que la limite existe.

Une fois la divergence, c'est-à-dire le débit de source du champ $R(x)$ à tout point, donnée, on peut en déduire, en intégrant, le débit total des sources de ce champ dans tout le domaine V. Désignons-le par $E(R(x), V)$. Ainsi, par définition, on a :

$$E(R(x), V) = \int\int\int_V \operatorname{div} R(x)\, dx.$$

Le processus étant stationnaire ($R(x)$ ne varie pas avec le temps), tout le fluide créé doit traverser la frontière, ce qui amène à l'égalité

$$\int\int\int_V \operatorname{div} R(x)\, dv = \oiint_S (m, R)\, dS \qquad (3)$$

dont l'interprétation physique est immédiate.

4.12. Passons à la théorie rigoureuse.

a. La définition du flux d'un champ vectoriel donnée dans *4.11* peut être étendue, sans modification essentielle, au cas d'espace R_n. Soit $R = R(x)$ un champ vectoriel continu par morceaux dans un domaine $G \subset R_n$. Quelle que soit une surface $(n-1)$-dimensionnelle lisse par morceaux $S \subset G$ avec le vecteur unitaire continu de la normale $m(x)$ *, on peut définir le flux du champ R à travers la surface S par la formule

$$W(R, S) = \int_S (m(x), R(x))\, dS.$$

De même que pour R_3, une surface $S \subset R_n$ est dite *fermée* si elle partage l'espace R_n en deux ensembles, intérieur et extérieur. L'intégrale sur une surface fermée sera désignée par $\oint$, bien que, conformément à la désignation $\oiint$ dans R_3, il faudrait écrire $\underbrace{\oint\cdots\oint}_{n-1}$ pour R_n.

Le flux du champ R à travers une surface *fermée* découpant un domaine V, la normale étant extérieure, représente une fonction du domaine V :

$$W(R, S) \equiv W_V(R) = \oint_S (m, R)\, dS. \qquad (1)$$

* En général, ce n'est pas pour toute surface lisse S, même dans R_3, qu'il existe un vecteur unitaire continu de la normale (exercice 15). L'existence d'un tel vecteur est une certaine restriction imposée à S (son « orientabilité » ; pour les détails, voir la troisième partie, chap. 8).

b. A présent, portons notre attention au fait que la fonction $W_V(R)$ est une fonction fortement additive du domaine V. Rappelons qu'une fonction $\Phi(V)$ du domaine V est dite fortement additive (*3.41a*) si, pour n'importe quels deux domaines V_1 et V_2 sans point intérieur commun (avec, éventuellement, une partie commune de la frontière) dont la réunion est V, on a l'égalité

$$\Phi(V) = \Phi(V_1) + \Phi(V_2). \tag{2}$$

L'additivité forte du flux (1) est basée sur le fait que le vecteur normal m doit toujours être dirigé vers la partie extérieure de l'espace; par conséquent, en tout point commun aux frontières de V_1 et V_2, s'il en existe (fig. 4.1-3), les quantités (m, R) pour les frontières de V_1 et de V_2 sont de signes opposés, les parties correspondantes des intégrales $W_{V_1}(R)$ et $W_{V_2}(R)$ se réduisent, et le résultat représente une intégrale sur la frontière de tout le domaine V.

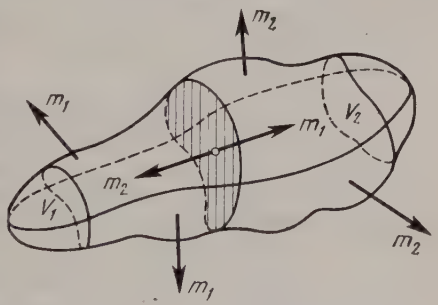


Fig. 4.1-3

Pourtant, le raisonnement ci-dessus va bien pour seuls les domaines d'une forme suffisamment simple. Si les domaines V_1 et V_2 sont tant soit peu compliqués, la démonstration de l'égalité (2) devient délicate. C'est pourquoi, nous n'envisageons ici que certains domaines spéciaux où l'additivité de la fonction (1) est évidente. A savoir, nous considérons un pavé fixe $X \subset R_n$, un domaine fixe $V \subset X$ qui représente un corps jordanien, avec la frontière S lisse par morceaux, et la famille $\mathfrak{A}(V)$ de tous les ensembles jordaniens U dont chacun représente l'intersection d'un sous-pavé du pavé X avec V. La famille $\mathfrak{A}(X)$ est, évidemment, un demi-anneau (*3.12*); les ensembles U sont ses cellules; en choisissant pour mesure d'une cellule U son volume $|U|$ nous obtenons un demi-anneau muni d'une mesure, de sorte que V devient espace chargé, même normalement chargé (*3.51b*). Chaque cellule U, ainsi que V lui-même, a une frontière $S(U)$ lisse par morceaux. La fonction

$$W_U(R) = \oint_{S(U)} (m(x),\ R(x))\, dS$$

est, nous l'avons vu tout à l'heure, fortement additive sur les cellules $U \in \mathfrak{A}(V)$, et nous pouvons y appliquer les résultats *3.41-3.44*. En particulier, conformément à la définition *3.42b*, la densité de la

fonction $W_U(R)$ à un point $y \in V$ est la limite (si elle existe)

$$\lim_{U \to y} \frac{1}{|U|} \oint_{S(U)} (m, R)\, dS.$$

Dans le présent cas, nous désignons cette limite par div $R(y)$. Supposons que div $R(y)$ existe à tout point $y \in V$ et représente une fonction continue de y. Alors a lieu l'égalité *3.44* (1)

$$\oint_{S(U)} (m, R)\, dS \equiv W_U(R) = \oint_U \operatorname{div} R(y)\, dy,$$

donc l'égalité *4.11* (3) se trouve démontrée dans les conditions imposées plus haut au flux $W_U(R)$. Notons que nous l'avons démontrée pour l'espace n-dimensionnel R_n.

4.13. Le résultat *4.12* est basé sur l'hypothèse que la divergence du champ $R(x)$, i.e. la quantité

$$\lim_{U \to y} \frac{1}{|U|} \oint_{S(U)} (m, R)\, dS \qquad (1)$$

existe et représente une fonction continue du point. Pour quels champs vectoriels peut-on garantir l'existence de la limite (1)? Nous répondrons à cette question dans la suite (*4.15*); nous verrons que les champs vectoriels dont les composantes ont leurs dérivées continues possèdent nécessairement une divergence continue. Pour cette démonstration nous utiliserons l'une des plus importantes formules liant les intégrales de volume et de surface, celle d'Ostrogradski.

La formule connue de Newton-Leibniz (voir *U9.32*)

$$\int_a^b F'(x)\, dx = F(b) - F(a)$$

lie la dérivée et l'intégrale de fonction d'une variable. La formule d'Ostrogradski en est une des plus importantes généralisations multidimensionnelles.

Choisissons dans l'espace n-dimensionnel euclidien R_n une base orthonormée quelconque $e_1, \ldots, e_n$, alors le produit scalaire de deux vecteurs $x = (x_1, \ldots, x_n)$ et $y = (y_1, \ldots, y_n)$ s'écrira

$$(x, y) = x_1 y_1 + \ldots + x_n y_n.$$

Considérons dans R_n un domaine G limité par une surface lisse par morceaux S. La surface S peut donc être représentée comme réunion d'un nombre fini de ses parties S_p $(p = 1, \ldots, q)$ deux à deux sans point commun intérieur (par rapport à S); pour tout point intérieur $x \in S_p$ il y a un voisinage V (dans R_n) dans les limites duquel la

surface S_p est représentable par les équations paramétriques de la forme

$$x_i = \varphi_i(u_1, \ldots, u_{n-1}), \quad i = 1, \ldots, n,$$

avec les fonctions $\varphi_i(u)$, $u = (u_1, \ldots, u_{n-1})$ possédant les dérivées premières continues, le rang de la matrice de Jacobi $\left\| \frac{\partial(\varphi_1, \ldots, \varphi_n)}{\partial(u_1, \ldots, u_{n-1})} \right\|$ valant $n - 1$. Le vecteur

$$N = \begin{Vmatrix} e_1 & \ldots & e_n \\ \frac{\partial \varphi_1(u)}{\partial u_1} & \ldots & \frac{\partial \varphi_n(u)}{\partial u_1} \\ \ldots & \ldots & \ldots \\ \frac{\partial \varphi_1(u)}{\partial u_{n-1}} & \ldots & \frac{\partial \varphi_n(u)}{\partial u_{n-1}} \end{Vmatrix},$$

comme nous l'avons vu dans *3.62c*, est normal à la surface S en tout point $u \in S$, sauf en les points de jonction (où ce vecteur est défini d'une façon non univoque, nous ne l'y considérons pas). Etant donnée l'hypothèse sur le rang de la matrice de Jacobi, le vecteur N est partout non nul. En le normalisant, nous obtenons le vecteur

$$m(u) = N/|N| \tag{2}$$

qui peut avoir deux sens opposés possibles. Nous en choisissons le sens extérieur par rapport au domaine G. Maintenant le vecteur $m(u)$ est déterminé d'une façon univoque; nous l'appelons *vecteur unitaire normal à la surface S au point u*. Désignons par ω_j l'angle formé par le vecteur m et l'axe des x_j; alors

$$m = e_1 \cos \omega_1 + \ldots + e_n \cos \omega_n. \tag{3}$$

Soit $P(x) = P(x_1, \ldots, x_n)$ une fonction dans le domaine G, continue et possédant sa dérivée continue $\frac{\partial P}{\partial x_k}$. Alors a lieu la formule (dont la déduction sera donnée dans *4.14*):

$$\int_G \frac{\partial P(x)}{\partial x_k} dx = \oint_S \cos \omega_k \cdot P(x) \, dS. \tag{4}$$

La fonction $\cos \omega_k$ est définie sur toute la surface S, sauf aux points de jonction des surfaces S_j; elle est continue sur chaque partie S_j, donc, continue par morceaux sur S, ce qui garantit l'existence de l'intégrale de surface dans le second membre.

En remplaçant dans (4) $P(x)$ par $P_k(x)$ et sommant les formules qui en résultent sur l'indice k de 1 à n nous arrivons à la *formule*

d'Ostrogradski

$$\int_G \left(\frac{\partial P_1(x)}{\partial x_1} + \ldots + \frac{\partial P_n(x)}{\partial x_n} \right) dx =$$

$$= \oint_S (\cos \omega_1 P_1(x) + \ldots + \cos \omega_n P_n(x))\, dS. \qquad (5)$$

Pour $n = 1$, le domaine G devient intervalle, sa frontière S est formée de deux points, et l'intégrale de surface dans le second membre de (4) se remplace, naturellement, par la différence des valeurs de la fonction $P(x)$ à ces points; nous retrouvons la formule de Newton-Leibniz. Dans le cas général, la formule d'Ostrogradski réduit l'intégrale de volume d'une forme spéciale à une intégrale étendue sur la surface qui sert de frontière au volume en question.

Voyons comment est faite cette formule. Dans l'intégrale de surface du second membre, chaque fonction $P_k(x)$ est multipliée par la composante respective du vecteur normal m; dans l'intégrale de volume du premier membre, chaque fonction $P_k(x)$ est dérivée par rapport à la coordonnée respective. On peut dire que le passage du second membre de l'égalité (5) à son premier membre consiste à remplacer les symboles $\cos \omega_k$ par les symboles $\frac{\partial}{\partial x_k}$, et l'intégrale de surface par l'intégrale de volume.

Pour $P_k(x) \equiv x_k$, la formule amène à un résultat intéressant:

$$\int_G n\, dx = n|G| = \oint_S \sum_{k=1}^{n} \cos \omega_k \cdot x_k\, dS. \qquad (6)$$

Ainsi, *le volume du domaine G s'exprime par une intégrale de surface étendue sur la frontière du domaine.*

Une autre conséquence utile de la formule (4) se déduit en multipliant ses deux membres par chaque vecteur de base e_k, puis sommant les résultats sur k; comme

$$\sum_{k=1}^{n} \frac{\partial P(x)}{\partial x_k} e_k = \operatorname{grad} P(x), \quad \sum_{k=1}^{n} \cos \omega_k \cdot e_k = m = m(x),$$

nous obtenons

$$\int_G \operatorname{grad} P(x)\, dx = \oint_S m(x)\, P(x)\, dS, \qquad (7)$$

ce qui représente l'une des notations vectorielles de la formule d'Ostrogradski.

En y posant $P(x) \equiv 1$, nous sommes amené à la relation

$$\oint_S m(x)\, dS = 0;$$

autrement dit : *la valeur moyenne du vecteur unitaire de la normale sur la surface (lisse par morceaux) qui sert de frontière à un domaine V est nulle.*

4.14. Abordons la déduction de la formule *4.13* (4) qui a pour conséquence la formule *4.13* (5). Désignons par Q la projection du domaine G sur le plan des coordonnées $x_1, \ldots, x_{n-1}$.

a. Supposons pour le moment, comme dans *3.52c*, que le domaine G satisfasse à la condition suivante :

(*) *Toute droite verticale ou bien n'a pas de point commun avec G, ou bien le perce suivant un intervalle (qui peut éventuellement dégénérer en un point).*

L'intervalle mentionné est déterminé par les inégalités

$$\varphi(x_1, \ldots, x_{n-1}) \leqslant x_n \leqslant \psi(x_1, \ldots, x_{n-1})$$

où les fonctions φ et ψ sont continues. Introduisons la désignation $x' = (x_1, \ldots, x_{n-1})$.

Appelons *partie supérieure* de la surface S sa partie S_s déterminée par l'équation $x_n = \psi(x')$ et *partie inférieure* de S sa partie S_i déterminée par l'équation $x_n = \varphi(x')$. Aux points de la partie supérieure S_s où le vecteur m est défini, nous avons l'inégalité $(m, e_n) \geqslant 0$; sur la partie inférieure S_i, a lieu l'inégalité $(m, e_n) \leqslant 0$. Ainsi, il résulte de *3.62d* :

$$(m, e_n) = \left(\frac{N}{|N|}, e_n\right) = \begin{cases} \dfrac{1}{\sqrt{\left(\frac{\partial\psi}{\partial x_1}\right)^2 + \ldots + \left(\frac{\partial\psi}{\partial x_{n-1}}\right)^2 + 1}} & (x \in S_s), \\ -\dfrac{1}{\sqrt{\left(\frac{\partial\varphi}{\partial x_1}\right)^2 + \ldots + \left(\frac{\partial\varphi}{\partial x_{n-1}}\right)^2 + 1}} & (x \in S_i). \end{cases}$$

En appliquant successivement la règle de transformation d'une intégrale multiple en une intégrale itérée (*3.52c*) et la définition d'une intégrale de surface nous obtenons :

$$\int_G \frac{\partial P(x)}{\partial x_n}\, dx = \int_Q \left[\int_{x_n = \varphi(x')}^{\psi(x')} \frac{\partial P(x', x_n)}{\partial x_n}\, dx_n \right] dx' =$$

$$= \int_Q P(x', \psi(x'))\, dx' - \int_Q P(x', \varphi(x'))\, dx' =$$

$$= \int_Q P(x', \psi(x'))(m, e_n) \sqrt{\left(\frac{\partial \psi}{\partial x_1}\right)^2 + \ldots + \left(\frac{\partial \psi}{\partial x_{n-1}}\right)^2 + 1}\, dx' +$$

$$+ \int_Q P(x', \varphi(x'))(m, e_n) \sqrt{\left(\frac{\partial \varphi}{\partial x_1}\right)^2 + \ldots + \left(\frac{\partial \varphi}{\partial x_{n-1}}\right)^2 + 1}\, dx' =$$

$$= \int_{S_s} P(x)(m, e_n)\, dS_s + \int_{S_i} P(x)(m, e_n)\, dS_i = \oint_S P(x) \cos \omega_n\, dS.$$

La formule *4.13* (4) est donc démontrée dans l'hypothèse (*).

b. Comme dans *3.52d*, on peut considérer un cas plus général. Supposons que le domaine G soit réunion des domaines $G_1, \ldots, G_k$ sans point intérieur commun dont chacun satisfasse à la condition (*) ; écrivons pour chaque tel domaine la formule (4) et additionnons les résultats ; la somme des intégrales de volume sur les domaines G_k fournira l'intégrale de volume sur le domaine G, celle des intégrales de surface sur les frontières S_j des domaines G_j, l'intégrale de surface sur la frontière S du domaine G, puisque les termes correspondant aux intégrales sur les parties des frontières de G_j qui sont à l'intérieur de G se réduisent, ce qu'on a vu plus haut.

Nous avons déjà dit qu'avec la formule *4.13* (4) se trouve établie dans le cas général la formule d'Ostrogradski *4.13* (5).

4.15. Montrons à présent que la divergence d'un champ $R(x) = \{P_1(x), \ldots, P_n(x)\}$ existe et est continue dès que les composantes $P_1(x), \ldots, P_n(x)$ du champ $R(x)$ admettent les dérivées continues. A cette fin, écrivons pour le domaine G la formule d'Ostrogradski *4.13* (5) :

$$\oint_S (m, R)\, dS \equiv \oint_S (\cos \omega_1 P_1 + \ldots + \cos \omega_n P_n)\, dS =$$

$$= \int_G \left(\frac{\partial P_1(x)}{\partial x_1} + \ldots + \frac{\partial P_n(x)}{\partial x_n}\right) dx. \qquad (1)$$

La fonction additive du domaine G qui figure dans le second membre a pour densité (*3.42c*) la fonction sous le signe somme :

$$\lim_{U \to y} \frac{1}{|U|} \int_G \left(\frac{\partial P_1(x)}{\partial x_1} + \ldots + \frac{\partial P_n(x)}{\partial x_n}\right) dx = \frac{\partial P_1(y)}{\partial x_1} + \ldots + \frac{\partial P_n(y)}{\partial x_n}.$$

Par conséquent, le flux du champ R qui figure dans le premier membre de (1) a la même densité, donc

$$\lim_{U \to y} \frac{1}{|U|} \oint_{S(U)} (m, R)\, dS \equiv \operatorname{div} R(y) = \frac{\partial P_1(y)}{\partial x_1} + \ldots + \frac{\partial P_n(y)}{\partial x_n}, \qquad (2)$$

ce qui démontre notre assertion. Ainsi, nous obtenons une forme explicite de la divergence du champ R.

4.16. Cas $n = 2$; formule de Green. En cas de $n = 2$, le domaine G est une figure plane limitée par un contour L lisse par morceaux. Si l'on pose $R = R(x, y) = e_1 M(x, y) + e_2 N(x, y)$, où $M(x, y)$ et $N(x, y)$ sont des fonctions avec les dérivées premières continues, la formule d'Ostrogradski prend la forme

$$\iint_G \left(\frac{\partial M}{\partial x} + \frac{\partial N}{\partial y} \right) dx\, dy = \oint_L (m, R)\, dL, \tag{1}$$

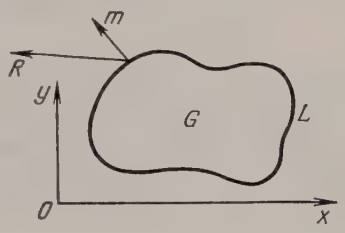


Fig. 4.1-4

où m est le vecteur unitaire normal au contour L et dirigé vers l'extérieur du domaine G (fig. 4.1-4). Pour le moment, l'intégrale curviligne dans le second membre n'est pas liée à un sens de parcours du contour L, car la paramétrisation du contour n'est pas encore fixée. Choisissons pour paramètre la longueur d'arc s comptée dans le sens positif de parcours (i.e. dans le sens inverse des aiguilles d'une montre); alors le vecteur polaire r de la courbe L, de même que les vecteurs R et m, devient fonction de s, et nous aurons la relation

$$\int_L (m, R)\, dL = \oint_L (m(s), R(s))\, ds.$$

Nous affirmons ensuite que

$$\oint_L (m, R)\, ds = \oint_L M\, dy - N\, dx.$$

En effet, le vecteur unitaire $\tau = \frac{dr}{ds}$ tangent au contour L a pour composantes $\frac{dx}{ds} = \cos(\widehat{\tau, e_1})$, $\frac{dy}{ds} = \cos(\widehat{\tau, e_2})$, de sorte que $dx = ds \cos(\widehat{\tau, e_1})$, $dy = ds \cos(\widehat{\tau, e_2})$, donc

$$\oint_L M\, dy - N\, dx = \oint_L [M \cos(\widehat{\tau, e_2}) - N \cos(\widehat{\tau, e_1})]\, ds =$$

$$= \oint_L [M \cos(\widehat{m, e_1}) + N \cos(\widehat{m, e_2})]\, ds = \oint_L (m, R)\, ds.$$

La formule d'Ostrogradski prend la forme:

$$\iint_G \left(\frac{\partial M}{\partial x} + \frac{\partial N}{\partial y} \right) dx\, dy = \oint_L M\, dy - N\, dx, \tag{2}$$

sous cette forme elle s'appelle *formule de Green.*

Pour $M \equiv x$, $N \equiv y$, le premier membre représente le double de l'aire du domaine G. Ainsi l'on a:

$$|G| = \frac{1}{2} \oint_L x\,dy - y\,dx.$$

On a déjà vu cette formule obtenue dans *U9.94* par une autre méthode.

En remplaçant dans (2) M par N et N par $-M$ nous avons:

$$\iint_G \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) dx\,dy = \oint_L M\,dx + N\,dy. \tag{3}$$

Si l'on choisit, comme avant, pour paramètre sur la courbe L la longueur d'arc s, on aura

$$\oint_L M\,dx + N\,dy = \oint_L \left(M\frac{dx}{ds} + N\frac{dy}{ds}\right) ds = \oint_L (\tau, R)\,ds$$

où τ désigne toujours le vecteur unitaire tangent au contour L. Nous arrivons donc à une autre notation de la formule d'Ostrogradski pour $n = 2$:

$$\iint_G \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) dx\,dy = \oint_L (\tau, R)\,ds, \tag{4}$$

où $R = (M, N)$ est une fonction vectorielle avec les composantes $M(x, y)$, $N(x, y)$ continûment dérivables dans le domaine G.

4.17. Fonctions à valeurs vectorielles. Les composantes du champ $R = \{P_1, \ldots, P_n\}$ ne sont pas forcément des fonctions numériques: toutes les conclusions de *4.11-4.16* restent valables pour les fonctions $P_j(x)$ à valeurs, par exemple, dans un espace de Banach B. Bien entendu, il faut toujours exiger que les fonctions $P_j(x)$ soient continues et que, pour tout j fixe, la fonction $P_j(x)$ possède la dérivée continue par rapport à x_j. En particulier, a lieu la formule d'Ostrogradski *4.13* (5) sous la forme

$$\int_G \sum_{k=1}^{n} \frac{\partial P_k}{\partial x_k}\,dx = \oint_S \sum_{k=1}^{n} \cos\alpha_k \cdot P_k(x)\,dS \quad (\in B). \tag{1}$$

La règle signalée dans *4.13* reste toujours valable: en passant du second au premier membre les composantes $\cos\omega_j$ du vecteur m sont remplacées par les symboles de dérivation par rapport aux coordonnées correspondantes, et l'intégrale de surface par l'intégrale de volume.

Enfin, pour $n = 2$, $R = \{M, N\}$, les deux variantes de la formule d'Ostrogradski, *4.16* (1) et *4.16* (4), restent valables:

$$\iint_G \left(\frac{\partial M}{\partial x} + \frac{\partial N}{\partial y}\right) dx\,dy = \oint_L (m, R)\,dL, \tag{2}$$

$$\iint_G \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}\right) dx\,dy = \oint_L (\tau, R)\,dL. \tag{3}$$

4.18. Certaines conséquences intéressantes de *4.17* ont lieu déjà dans le cas de fonctions $P_j(x)$ à valeurs dans un espace de dimension finie. Considérons le produit vectoriel dans R_3 de deux vecteurs $q = (q_1, q_2, q_3)$ et $R = (Q_1, Q_2, Q_3)$:

$$[q, R] = \begin{vmatrix} e_1 & e_2 & e_3 \\ q_1 & q_2 & q_3 \\ Q_1 & Q_2 & Q_3 \end{vmatrix} = q_1 \begin{vmatrix} e_3 & e_2 \\ Q_3 & Q_2 \end{vmatrix} + q_2 \begin{vmatrix} e_1 & e_3 \\ Q_1 & Q_3 \end{vmatrix} + q_3 \begin{vmatrix} e_2 & e_1 \\ Q_2 & Q_1 \end{vmatrix}. \tag{1}$$

Posons

$$P_1 = \begin{vmatrix} e_3 & e_2 \\ Q_3 & Q_2 \end{vmatrix}, \quad P_2 = \begin{vmatrix} e_1 & e_3 \\ Q_1 & Q_3 \end{vmatrix}, \quad P_3 = \begin{vmatrix} e_2 & e_1 \\ Q_2 & Q_1 \end{vmatrix}.$$

Alors la formule *4.17* (1) amène au résultat:

$$\oint_S [m, R]\,dS = \oint_S \sum_{k=1}^{3} \cos\alpha_k \cdot P_k(x)\,dS \equiv \oint_S \begin{vmatrix} e_1 & e_2 & e_3 \\ \cos\alpha_1 & \cos\alpha_2 & \cos\alpha_3 \\ Q_1(x) & Q_2(x) & Q_3(x) \end{vmatrix} dS =$$

$$= \int_G \int \int \begin{vmatrix} e_1 & e_2 & e_3 \\ \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \\ Q_1(x) & Q_2(x) & Q_3(x) \end{vmatrix} dx \equiv \int_G \int \int \left[\left(\frac{\partial Q_3}{\partial x_2} - \frac{\partial Q_2}{\partial x_3}\right) e_1 + \right.$$

$$\left. + \left(\frac{\partial Q_1}{\partial x_3} - \frac{\partial Q_3}{\partial x_1}\right) e_2 + \left(\frac{\partial Q_2}{\partial x_1} - \frac{\partial Q_1}{\partial x_2}\right) e_3\right] dx.$$

L'expression $\oint_S [m, R]\,dS$ s'appelle *rotation du champ vectoriel* $R(x)$ *sur la frontière* S *du domaine* G. Dans *4.24*, nous mettrons en évidence le sens géométrique et mécanique de cette notion.

§ 4. 2. Rotationnel du champ vectoriel

4.21. Gradient et potentiel.

a. Soit $L = \{x \in R_n : x = x(t),\ a \leqslant t \leqslant b\}$ une courbe lisse par morceaux, donnée dans un domaine $G \subset R_m$ et sans points singuliers (i.e. $x'(t)$ ne s'y annule pas). Alors on peut paramétriser la courbe L avec un paramètre naturel, i.e. avec la longueur d'arc s;

le vecteur $\tau = x'(s)$ sera le vecteur unitaire tangent à la courbe L. Soit, de plus, $R = \{P_1(x), \ldots, P_n(x)\}$ un champ vectoriel continu donné dans le domaine G. Le nombre

$$\int_L (\tau(s), R(s))\, ds = \int_L P_1\, dx_1 + \ldots + P_n\, dx_n \qquad (1)$$

s'appelle *intégrale du champ R sur la courbe L.*

Si l'on interprète le champ R comme champ d'une force (par exemple, de gravitation), l'expression (1) sera interprétée comme travail de la force R sur le chemin L. Dans certains problèmes de mécanique, il est essentiel que ce travail dépende non pas du chemin L, mais seulement des points de départ et d'arrivée. Or, cette situation n'est pas habituelle, et nous posons la question suivante, purement mathématique: *quelles sont les conditions à imposer au champ R (i.e. aux $P_1, \ldots, P_n$), pour que l'intégrale du champ R sur n'importe quel chemin $L \subset G$ dépende des seules les extrémités de L, et ne dépende pas de la courbe L elle-même*?

Supposons que le champ R soit le champ du gradient d'une fonction scalaire $\varphi(x)$, c'est-à-dire que

$$P_1(x) = \frac{\partial \varphi(x)}{\partial x_1}, \ \ldots, \quad P_n(x) = \frac{\partial \varphi(x)}{\partial x_n}. \qquad (2)$$

Dans ce cas, l'intégrale du champ R sur n'importe quel chemin L n'est déterminée que par le point de départ A et le point d'arrivée B du chemin L, puisque

$$\int_A^B (\tau, R)\, ds = \int_A^B P_1\, dx_1 + \ldots + P_n\, dx_n = \int_A^B \frac{\partial \varphi}{\partial x_1}\, dx_1 + \ldots +$$

$$+ \frac{\partial \varphi}{\partial x_n}\, dx_n = \int_A^B d\varphi = \varphi(B) - \varphi(A). \qquad (3)$$

Etant données les relations (2), le champ R est dit *potentiel*, et la fonction $\varphi(x)$ est appelée son *potentiel*. En termes de mécanique, l'égalité (3) se lit: le travail d'un champ de force potentiel $R = \{P_1, \ldots, P_n\}$ sur un chemin L liant des points A et B vaut la différence des valeurs du potentiel aux points B et A.

c. La réciproque est également valable:

Théorème. *Si l'intégrale d'un champ R sur tout chemin L dans le domaine G ne dépend que des extrémités de L, alors R est un champ potentiel.*

Démonstration. Fixons dans le domaine G un point A et posons, pour tout autre point, $B = B(x_1, \ldots, x_n)$:

$$\varphi(B) \equiv \varphi(x_1, \ldots, x_n) = \int_A^B (\tau, R)\, ds \qquad (4)$$

où l'intégration s'effectue sur n'importe quel chemin allant de A à B dans le domaine G ; comme, par l'hypothèse, l'intégrale du champ ne dépend pas du chemin, la fonction $\varphi(B)$ est définie d'une façon univoque. Montrons que

$$\operatorname{grad} \varphi(B) = R(B).$$

Pour le faire, composons la dérivée de la fonction φ suivant une direction quelconque λ déterminée par un vecteur unitaire e. Choisissons un point B_1 sur la demi-droite issue du point B et allant dans le sens de λ (fig. 4.2-1) et faisons le rapport

$$\frac{\varphi(B_1) - \varphi(B)}{|B_1 - B|}.$$

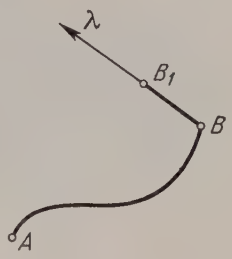


Fig. 4.2-1

On peut dire que la quantité $\varphi(B_1)$ est l'intégrale du champ R sur le chemin ABB_1, où BB_1 est le segment de droite. Alors

$$\frac{\varphi(B_1) - \varphi(B)}{|B_1 - B|} = \frac{1}{|B_1 - B|} \int_B^{B_1} (\tau, R)\, ds$$

est la valeur moyenne de la fonction $(\tau, R) = (e, R)$ sur le segment BB_1. Lorsque B_1 s'approche de B, cette valeur moyenne tend vers la limite $(e, R(B))$. Par conséquent, la dérivée de la fonction φ suivant la direction λ au point B existe, et l'on a :

$$\frac{\partial \varphi(B)}{\partial \lambda} = (e, R(B)).$$

En particulier,

$$\frac{\partial \varphi(B)}{\partial x_1} = P_1(B), \ \ldots, \quad \frac{\partial \varphi(B)}{\partial x_n} = P_n(B).$$

Or, ceci veut dire que le vecteur $R(B)$ coïncide avec le gradient de la fonction φ (4) au point B. Le théorème est démontré.

Remarquons encore qu'un champ vectoriel donné R ne peut avoir, à un terme constant près, qu'une seule fonction potentielle φ. En effet, si ψ est n'importe quelle autre fonction potentielle du champ R, nous avons d'après la formule (3) :

$$\psi(B) - \psi(A) = \int_A^B (\tau, R)\, ds = \varphi(B),$$

d'où

$$\psi(B) = \varphi(B) + \psi(A) = \varphi(B) + \text{const},$$

ce qui achève la démonstration.

d. La question posée dans *a* est donc réduite à la suivante : *quelles sont les conditions à imposer aux fonctions* $P_1(x), \ldots, P_n(x)$ *pour*

qu'il existe une fonction $\varphi(x)$ *telle que*

$$P_1(x) = \frac{\partial \varphi(x)}{\partial x_1}, \ \ldots, \quad P_n(x) = \frac{\partial \varphi(x)}{\partial x_n}? \tag{5}$$

Nous connaissons déjà la réponse, au moins pour des domaines suffisamment petits ou bien simplement connexes de l'espace R_n (*2.61b*) : une condition nécessaire et suffisante de l'existence d'une fonction $\varphi(x)$ qui satisfasse aux conditions (5) est donnée par les identités

$$\frac{\partial P_i(x)}{\partial x_j} = \frac{\partial P_j(x)}{\partial x_i} \quad (i, j = 1, \ldots, n). \tag{6}$$

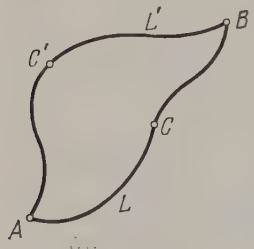


Fig. 4.2-2

Remarquons que, pour les domaines non simplement connexes, cette condition n'est plus suffisante (exercice 6).

4.22. Circulation d'un champ vectoriel. Soient $L = ACB$ et $L' = AC'B$ deux courbes (lisses par morceaux) qui vont, dans un domaine $G \subset R_n$, d'un point donné A à un point donné B (fig. 4.2-2). En suivant la courbe L de A vers B et revenant à A le long de la courbe L', on parcourt un contour fermé $\Gamma = ACBC'A$. Dans ce cas, pour un champ vectoriel continu R, choisi arbitrairement, nous avons :

$$\int_L (\tau, R)\, ds - \int_{L'} (\tau, R)\, ds = \int_{ACB} (\tau, R)\, ds + \int_{BC'A} (\tau, R)\, ds = \oint_\Gamma (\tau, R)\, ds. \tag{1}$$

L'intégrale d'un champ R le long d'un contour fermé L s'appelle *circulation du champ* R sur ce contour.

L'égalité (1) montre que les questions sur l'indépendance de l'intégrale d'un champ du chemin d'intégration et sur les conditions sous lesquelles la circulation du champ sur tout contour fermé soit nulle sont équivalentes. Donc, la réponse à la dernière question est la même qu'à la précédente : *la circulation d'un champ* R *sur tout contour fermé dans un domaine* G *est nulle si, et seulement si, le champ* R *est potentiel dans le domaine* G. Dans les domaines suffisamment petits ou simplement connexes, la réponse consiste à satisfaire aux relations *4.21* (6).

4.23. Encore un problème de mécanique. Soit $R(x)$ un champ vectoriel donné dans un domaine G de l'espace R_3. Nous l'interprétons ici, de même que dans *4.11*, comme champ de vitesse du fluide en mouvement. Voici une expérience de mécanique idéalisée. Portons dans le domaine G une petite rouette cylindrique Q munie d'un grand nombre de palettes qui peut tourner librement autour d'un axe fixé arbitrairement ; nous déterminons le sens positif

de l'axe par un vecteur unitaire l (fig. 4.2-3). En agissant sur les palettes les particules du fluide font tourner la rouette Q à une vitesse angulaire $\omega\ (Q,\ l,\ A)$ qui dépend (en un point donné A du domaine G) du vecteur l. Convenons sur la règle des signes (en supposant que le système de coordonnées soit à droite) : la vitesse angulaire de la rouette est positive, si elle tourne dans le sens opposé à celui des aiguilles d'une montre en regardant de l'extrémité du vecteur l (dont l'origine est au centre A de la rouette Q), et négative dans le cas contraire. En faisant varier de toutes les façons possibles

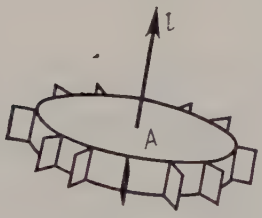


Fig. 4.2-3

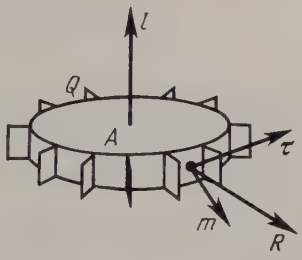


Fig. 4.2-4

la direction du vecteur l (la position de l'origine A étant la même), nous obtenons des valeurs différentes, en module et en signe, de la vitesse angulaire ; pour une position l_0 du vecteur l, la vitesse angulaire prend sa valeur maximale ω_0. On appelle *rotationnel du champ* R au point A le vecteur $2\omega_0 l_0$.

La définition du rotationnel d'un champ vectoriel donnée sous cette forme mécanique ne peut être considérée comme correcte du point de vue mathématique. (Elle est basée sur certaines hypothèses sur l'interaction des particules de fluide avec les palettes ; elle ne dit pas comment faire si la vitesse angulaire maximale correspond à deux positions différentes de l'axe de la rouette Q ; enfin, le rôle que jouent les dimensions de la rouette Q reste un peu vague.) Essayons de formuler une définition purement mathématique.

L'action rotatrice des particules de fluide sur une palette de la rouette Q est déterminée par la composante du vecteur R le long de la ligne d'action de la palette, autrement dit, par la quantité $(\tau\ (M),\ R\ (M))$, où $\tau\ (M)$ est le vecteur unitaire tangent à la circonférence de rotation de la palette et dirigé dans le sens positif (fig. 4.2-4). C'est l'action totale de toutes les particules sur les palettes qui fait tourner la rouette ; la plus naturelle supposition consiste en ce que la vitesse linéaire v de chaque point sur la circonférence de la rouette vaut la moyenne arithmétique de toutes les quantités $(\tau,\ R)$, en d'autres

termes, l'intégrale sur la surface latérale Σ de la rouette Q

$$v = \frac{1}{2\pi rh} \iint_{\Sigma} (\tau, R)\, dS,$$

où r est le rayon de la circonférence, h l'épaisseur de la rouette Q.

Pour obtenir la valeur de la vitesse angulaire, on doit diviser la vitesse linéaire par le rayon de la circonférence ; comme le volume $|Q|$ de la rouette Q vaut $\pi r^2 h$, nous avons l'expression suivante pour la vitesse angulaire cherchée :

$$\omega(Q, l) = \frac{1}{2|Q|} \iint_{\Sigma} (\tau, R)\, dS.$$

Si nous tenons à avoir une caractéristique de la rotation qui se rapporte non pas à la circonférence de la rouette, mais au point A même, il faut faire tendre vers zéro, dans l'expression obtenue, le rayon de la rouette et son épaisseur, puis trouver la limite de la quantité $\omega(Q, l)$.

Soit m (M) le vecteur unitaire normal au point M de la circonférence. Comme les vecteurs m, τ, l sont orientés de même que les vecteurs de base e_1, e_2, e_3 (le système de coordonnées est à droite!), nous avons $\tau = [l, m]$; ensuite, d'après la propriété cyclique du produit mixte, nous trouvons $(\tau, R) = ([l, m], R) = (l, [m, R])$, d'où

$$\omega(Q, l) = \frac{1}{2|Q|} \iint_{\Sigma} (l, [m, R])\, dS = \frac{1}{2} \left(l, \frac{1}{|Q|} \iint_{\Sigma} [m, R]\, dS \right).$$

Sur la base du cylindre Q, les vecteurs m et l sont colinéaires, i.e. $(l, [m, R]) = 0$; par conséquent, l'expression dans le second membre ne change pas, si nous remplaçons l'intégrale sur la surface latérale du cylindre par l'intégrale sur sa surface totale S :

$$\omega(Q, l) = \frac{1}{2} \left(l, \frac{1}{|Q|} \oiint_{S} [m, R]\, dS \right).$$

4.24. Il est temps de recourir au langage mathématique rigoureux.

a. Soit $R = \{P_1(x), P_2(x), P_3(x)\}$ un champ vectoriel continu donné dans un domaine $G \subset R_3$. Considérons un domaine $Q \subset G$ limité par une surface lisse par morceaux S, avec la normale extérieure m. L'expression

$$\oiint_{S} [m, R]\, dS \tag{1}$$

nous est déjà familière depuis *4.18* ; rappelons qu'elle s'appelle *rotation du champ R sur la frontière S du domaine Q*. Pour la même

raison que dans *4.12b* où l'on considérait le flux, la rotation est une fonction fortement additive du domaine G.

b. Appelons *rotation moyenne du champ R dans le domaine Q* le rapport de la rotation du champ R sur la surface fermée S au volume $|Q|$ du domaine Q qu'elle enferme. Supposons que le domaine Q se contracte vers un point $y \in G$. Si la rotation moyenne du champ R dans le domaine Q possède une limite qui ne dépend pas de la forme du domaine Q se contractant vers le point y, cette limite s'appelle *rotationnel* du champ R au point y et est désignée par $\operatorname{rot} R(y)$. Ainsi,

$$\operatorname{rot} R(y) = \lim_{Q \to y} \frac{1}{|Q|} \iint_S [m, R]\, dS.$$

Nous voyons que $\operatorname{rot} R(y)$ *représente la densité de la fonction fortement additive* (1) qui est la rotation du champ R sur la frontière du domaine Q. Contrairement au flux et à la divergence, la rotation et le rotationnel ne sont plus des fonctions scalaires, mais vectorielles (d'un domaine et d'un point respectivement).

c. En remplaçant dans le raisonnement *4.12* le produit scalaire (τ, R) par le produit vectoriel $[\tau, R]$ et compte tenu de *4.18* nous trouvons : *si* $\operatorname{rot} R(y)$ *existe et représente une fonction continue du point $y \in G$, alors on a*

$$\iiint_Q \operatorname{rot} R(y)\, dy = \iint_S [m, R]\, dS,$$

quel que soit le domaine $Q \subset G$ avec la frontière S lisse par morceaux.

d. En appliquant le théorème d'Ostrogradski on peut démontrer l'existence et la continuité du rotationnel en supposant que les composantes P_1, P_2, P_3 du champ R admettent les dérivées continues. En effet, sous cette hypothèse nous avons, d'après la formule *4.18* (2) :

$$\iint_S [m, R]\, dS = \iint_S \begin{vmatrix} e_1 & e_2 & e_3 \\ \cos\omega_1 & \cos\omega_2 & \cos\omega_3 \\ P_1 & P_2 & P_3 \end{vmatrix} dS =$$

$$= \iiint_Q \begin{vmatrix} e_1 & e_2 & e_3 \\ \dfrac{\partial}{\partial x_1} & \dfrac{\partial}{\partial x_2} & \dfrac{\partial}{\partial x_3} \\ P_1 & P_2 & P_3 \end{vmatrix} dx\, dy\, dz,$$

donc,

$$\operatorname{rot} R(y) = \lim_{Q \to y} \frac{1}{|Q|} \iint_S [m, R]\, dS = \begin{vmatrix} e_1 & e_2 & e_3 \\ \dfrac{\partial}{\partial x_1} & \dfrac{\partial}{\partial x_2} & \dfrac{\partial}{\partial x_3} \\ P_1(y) & P_2(y) & P_3(y) \end{vmatrix} =$$

$$= e_1\left(\frac{\partial P_3(y)}{\partial x_2} - \frac{\partial P_2(y)}{\partial x_3}\right) + e_2\left(\frac{\partial P_1(y)}{\partial x_3} - \frac{\partial P_3(y)}{\partial x_1}\right) + e_3\left(\frac{\partial P_2(y)}{\partial x_1} - \frac{\partial P_1(y)}{\partial x_2}\right).$$

4.25. Revenons à l'expérience idéalisée décrite dans *4.23*. Si rot R (A) existe, alors nous avons pour la vitesse angulaire de la rouette Q d'axe l l'expression limite suivante:

$$\omega\ (Q,\ l) = \frac{1}{2}(l,\ \text{rot}\ R\ (A)).$$

Soit θ l'angle des vecteurs l et rot R (A). Alors, d'après la définition du produit scalaire:

$$\omega\ (Q,\ l) = \frac{1}{2}\ |\ \text{rot}\ R\ (A)\ |\ \cos\theta,$$

d'où les conclusions suivantes:

a. *La quantité* $\omega\ (Q,\ l)$ *prend, à un point* A, *sa valeur maximale* $\frac{1}{2}\ |\ \text{rot}\ R\ (A)\ |$ *lorsque le vecteur* l *est dirigé suivant le vecteur* rot R (A). (A cette position de l'axe, la rouette Q a la vitesse angulaire maximale; ceci approuve la définition initiale « mécanique » du rotationnel comme vecteur qui est dirigé dans le sens de l'axe de la rotation à la vitesse angulaire maximale et dont la longueur vaut le double de la vitesse angulaire maximale.)

b. *Lorsque* l *est orthogonal au vecteur* rot R (A), *la valeur de* $\omega\ (Q,\ l)$ *est nulle* (pour de telles directions de l'axe, la rouette ne tournera pas).

c. *Pour les autres directions de l'axe, la quantité* $\omega\ (Q,\ l)$ *prend ses valeurs entre* $-\frac{1}{2}\ |\ \text{rot}\ R\ (A)\ |$ *et* $\frac{1}{2}\ |\ \text{rot}\ R\ (A)\ |$.

4.26. Formule de Stokes. Comme nous l'avons vu dans *4.24*, le rotationnel d'un champ vectoriel R de composantes dérivables P_1, P_2, P_3 s'exprime en coordonnées comme suit:

$$\text{rot}\ R = \begin{vmatrix} e_1 & e_2 & e_3 \\ \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \\ P_1 & P_2 & P_3 \end{vmatrix} = e_1\left(\frac{\partial P_3}{\partial x_2} - \frac{\partial P_2}{\partial x_3}\right) + e_2\left(\frac{\partial P_1}{\partial x_3} - \frac{\partial P_3}{\partial x_1}\right) + e_3\left(\frac{\partial P_2}{\partial x_1} - \frac{\partial P_1}{\partial x_2}\right). \quad (1)$$

Si, dans le domaine de définition G du champ R, son rotationnel est nul, alors, dans ce domaine, on a:

$$\frac{\partial P_2}{\partial x_3} = \frac{\partial P_3}{\partial x_2}, \quad \frac{\partial P_3}{\partial x_1} = \frac{\partial P_1}{\partial x_3}, \quad \frac{\partial P_2}{\partial x_1} = \frac{\partial P_1}{\partial x_2}. \quad (2)$$

Si le domaine G est assez simple, par exemple simplement connexe, alors, comme on l'a vu dans *4.22*, la circulation du champ R

$$\oint_L (\tau, R)\, ds$$

est nulle sur tout contour L. Inversement, si la circulation du champ R sur chaque contour fermé dans le domaine G est nulle, les relations (2) ont lieu en vertu de *4.22*, donc le rotationnel du champ R est nul.

Si rot $R \not\equiv 0$, alors, en général, la circulation du champ R n'est pas nulle non plus. Les deux grandeurs caractérisant des propriétés rotationnelles du champ R, il est naturel de s'attendre à ce qu'elles soient liées d'une façon ou d'une autre. Une telle dépendance est décrite par la formule de Stokes.

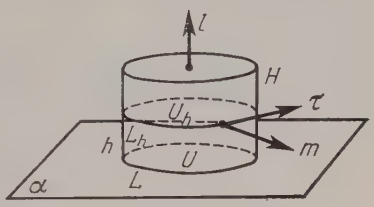


Fig. 4.2-5

Dans *4.24d*, pour un champ vectoriel $R = \{P_1, P_2, P_3\}$ avec le rotationnel continu dans un domaine Q limité par une surface lisse par morceaux S, on a établi la formule

$$\iiint_Q \operatorname{rot} R\, dQ = \oiint_S [m, R]\, dS, \qquad (3)$$

où m est le vecteur unitaire de normale extérieure.

Choisissons pour domaine Q le cylindre droit d'une hauteur H (fig. 4.2-5) dont la base représente un domaine U se trouvant dans un plan α et limité par un contour L. Désignons par l le vecteur unitaire normal au plan α et par τ le vecteur unitaire orthogonal à l, tangent à la surface latérale du cylindre Q et de sens positif (par rapport à l). Soient S et Σ les surfaces totale et latérale respectivement du cylindre Q. On a

$$\left(l, \oiint_S [m, R]\, dS\right) = \oiint_S (l, [m, R])\, dS =$$

$$= \iint_\Sigma (l, [m, R])\, dS = \iint_\Sigma ([l, m], R)\, dS = \iint_\Sigma (\tau, R)\, dS, \qquad (4)$$

car les vecteurs l et m sont colinéaires sur les bases du cylindre Q. Désignons par U_h la section du cylindre Q par le plan parallèle aux bases et passant à une distance h de la base inférieure, et par L_h le contour de la section U_h. Alors, pour n'importe quelles fonctions

$F(x)$ et $g(x)$, on a

$$\iiint_Q F(x)\,dx = \int_0^H \left\{ \iint_{U_h} F(x)\,dS \right\} dh,$$

$$\iint_\Sigma g(x)\,dS = \int_0^H \left\{ \oint_{L_h} g(x)\,ds \right\} dh;$$

en particulier,

$$\iiint_Q (l,\ \mathrm{rot}\, R)\,dx = \int_0^H \left\{ \iint_{U_h} (l,\ \mathrm{rot}\, R)\,dS \right\} dh,$$

$$\iint_\Sigma (\tau,\ R)\,dS = \int_0^H \left\{ \oint_{L_h} (\tau,\ R)\,dS \right\} dh.$$

Divisons ces relations par H, puis passons à la limite pour $H \to 0$. Comme, pour toute fonction continue $\Phi(h)$, le théorème de la moyenne fournit

$$\lim_{H\to 0} \frac{1}{H} \int_0^H \Phi(h)\,dh = \Phi(0),$$

nous avons

$$\iint_U (l,\ \mathrm{rot}\, R)\,dS = \oint_L (\tau,\ R)\,ds, \tag{5}$$

vu la relation (4) et l'égalité (3) multipliée scalairement par l.

La formule obtenue s'appelle *formule de Stokes pour un domaine plan U*. Elle relie le rotationnel du champ R et la circulation de ce champ sur le contour L limitant le domaine U. Si le champ R est dérivable, la formule (5) peut être déduite d'une manière beaucoup plus simple. Choisissons les coordonnées de façon que le contour L se trouve dans le plan x, y. Alors

$$l = (0,\ 0,\ 1), \quad (l,\ \mathrm{rot}\, R) = \frac{\partial P_2}{\partial x} - \frac{\partial P_1}{\partial y},$$

et la formule (5) prend la forme

$$\iint_U \left(\frac{\partial P_2}{\partial x} - \frac{\partial P_1}{\partial y} \right) dx\,dy = \oint_L (\tau,\ R)\,ds,$$

i.e. coïncide avec la formule connue *4.16* (4).

On peut écrire une formule plus générale de Stokes, non pas pour une partie du plan, mais pour une surface *orientable* (i.e. possédant e vecteur normal unitaire continu) U qui est limitée par un contour L

(lisse par morceaux) (fig. 4.2-6) :

$$\iint_U (l, \operatorname{rot} R)\, dS = \oint_L (\tau, R)\, ds. \qquad (6)$$

Ici $l = l(M)$ désigne le vecteur unitaire orthogonal à la surface U au point M ; cette fois il n'est pas constant, mais varie avec le point M. Le vecteur τ est toujours le vecteur unitaire, tangent au contour L et orienté positivement (par rapport à l). Si le vecteur τ est de sens opposé, le second membre de (6) change de signe, et l'égalité n'aura plus lieu.

Lorsque la surface U est sensiblement courbée, la règle de concordance des sens de la normale et du parcours du contour limitant la surface n'est plus correcte, et on doit la remplacer par celle-ci

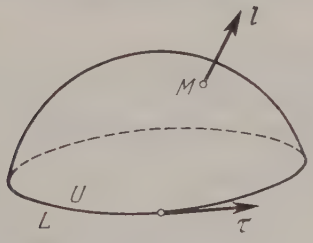


Fig. 4.2-6

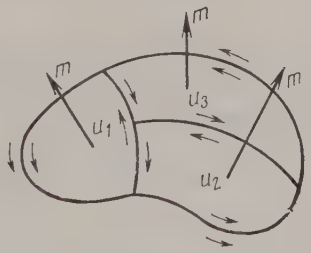


Fig. 4.2-7

qui est plus exacte : une surface orientable U peut toujours être partagée en morceaux suffisamment petits $U_1, \ldots, U_p$ limités par les contours respectifs fermés $L_1, \ldots, L_p$ dont les sens de parcours sont concordés comme suit : sur chaque morceau du contour L_i qui coïncide avec un morceau de L le sens de parcours est déterminé par un même vecteur tangent, et sur chaque morceau commun de deux contours L_i et L_j les sens de parcours sont opposés (fig. 4.2-7) ; alors le vecteur normal m est orienté sur la surface U_i d'une telle façon que le sens de parcours du contour L_i, vu depuis son extrémité, soit le sens inverse des aiguilles d'une montre. La possibilité d'une telle concordance pour toutes les surfaces orientables est démontrée en topologie ; dans notre cours nous n'aurons affaire qu'à des surfaces élémentaires où cette possibilité sera évidente.

La démonstration de la formule (6) peut être construite de la façon suivante. Etant valable pour les domaines plans, elle le reste pour chaque surface orientable *polyèdre* (composée d'un nombre fini de parties planes, par exemple de triangles). Ensuite, étant donnée une surface courbée (qui possède le plan tangent continu ou continu par morceaux), on la représente comme limite des surfaces polyèdres inscrites U_n composées de triangles (*3.64d*). En écrivant la formule

(6) pour chaque domaine U_n et passant ensuite à la limite pour $n \to \infty$, on aboutit à la formule (6) dans le cas général (*3.64e*).

Si l'on divise l'égalité (5) par l'aire du domaine U, puis passe à la limite lorsque le contour L se contracte dans le plan α vers un point A, on aura

$$(l,\ \mathrm{rot}\, R(A)) = \lim_{U \to A} \frac{1}{|U|} \oint_L (\tau,\ R)\, ds. \tag{7}$$

C'est encore une formule importante qui exprime la projection du rotationnel d'un champ R sur une direction donnée l par la circulation du champ R.

§ 4. 3. Opérateur hamiltonien

4.31. Les définitions de la divergence d'un champ vectoriel $R(x)$ (*4.15*)

$$\mathrm{div}\, R(y) = \lim_{V \to y} \frac{1}{|V|} \oint_S (m,\ R)\, dS \quad (\in R_1)$$

et du rotationnel du champ vectoriel $R(x)$ (*4.23*)

$$\mathrm{rot}\, R(y) = \lim_{V \to y} \frac{1}{|V|} \oint_S [m,\ R]\, dS \quad (\in R_3)$$

suggèrent l'idée de l'existence d'une définition générale qui contienne les définitions ci-dessus comme cas spéciaux. Une telle définition générale existe et a la forme suivante.

Soit $T(x)$ un champ opératoriel donné dans un domaine G de l'espace $X = R_n$; plus exactement, soit $T(x)$ une fonction dont la valeur pour tout $x \in G$ représente un opérateur linéaire de l'espace X dans un espace de Banach fixe Y. Pour le moment, il nous est plus commode de placer le symbole de la variable vectorielle indépendante à gauche du symbole de l'opérateur ; ainsi, pour tout $q \in X$, la fonction $qT(x)$ à valeurs dans Y est définie, quel que soit $x \in G$. Supposons ensuite que la fonction $T(x)$ soit continue (sur la norme opératorielle) dans le domaine G ; la fonction $qT(x)$ sera alors elle aussi continue dans G pour tout $q \in X$, et même plus, continue par rapport à l'ensemble des variables $q \in X$ et $x \in G$. Considérons dans le domaine G une surface fermée lisse par morceaux Γ qui découpe un sous-domaine V de volume $|V|$; désignons par $m(\xi)$ le vecteur unitaire de la normale extérieure au point $\xi \in \Gamma$ et par $d\Gamma(\xi)$ l'élément de la surface Γ.

La quantité

$$W_\Gamma(T) = \oint_\Gamma m(\xi)\, T(\xi)\, d\Gamma(\xi) \quad (\in Y) \tag{1}$$

s'appelle flux du champ opératoriel $T(x)$ à travers la surface Γ. Comme plus haut, le flux du champ $T(x)$ à travers la surface Γ est

une fonction (fortement) additive (à valeurs dans Y) du domaine V. La densité de flux du champ $T(x)$ à un point $A \subset G$

$$\nabla T(A) = \lim_{V \to A} \frac{1}{|V|} \oint_{\Gamma} m(\xi)\, T(\xi)\, d\Gamma(\xi) \quad (\in Y), \tag{2}$$

si elle existe, s'appelle *hamiltonien du champ $T(x)$ au point A*. L'opérateur ∇ (nabla *) de passage du champ $T(x)$ à la densité de son flux s'appelle *opérateur hamiltonien*. Le passage à la limite dans le second membre de (2) se fait, de façon naturelle, sur la métrique de l'espace Y. Si $\nabla T(A)$ existe partout dans le domaine G et représente une fonction continue (par morceaux) du point A (à valeurs dans Y), alors, de même que dans *4.12*, la fonction d'un domaine $W_\Gamma(T)$ est restituée d'après sa densité $\nabla T(x)$ avec la formule

$$W_\Gamma(T) \equiv \oint_{\Gamma} m(\xi)\, T(\xi)\, d\Gamma(\xi) = \int_V \nabla T(x)\, dv(x). \tag{3}$$

A présent, montrons que *si le champ opératoriel $T(x)$ est non seulement continu dans le domaine G, mais aussi possède dans G la dérivée continue, la quantité $\nabla T(A)$ existe en tout point $A \in G$; nous en obtiendrons une forme explicite.* Soient $e_1, \ldots, e_n$ une base orthonormale de l'espace X et $\alpha_1, \ldots, \alpha_n$ les angles que le vecteur $m(\xi)$ fait avec les vecteurs de base respectifs. Nous avons $m(\xi) = \cos\alpha_1 \cdot e_1 + \ldots + \cos\alpha_n \cdot e_n$, donc,

$$m(\xi)\, T(\xi) = \cos\alpha_1 \cdot e_1 T(\xi) + \ldots + \cos\alpha_n \cdot e_n T(\xi).$$

D'après la formule généralisée d'Ostrogradski *4.17* (1), nous avons

$$\oint_{\Gamma} m(\xi)\, T(\xi)\, d\Gamma(\xi) \equiv \oint_{\Gamma} \sum_{k=1}^{n} \cos\alpha_k \cdot e_k T(\xi)\, d\Gamma(\xi) =$$

$$= \int_V \sum_{k=1}^{n} \frac{\partial}{\partial x_k} (e_k T(x))\, dv.$$

Il en résulte l'existence de la limite (2) et la formule

$$\nabla T(A) \equiv \lim_{V \to A} \frac{1}{|V|} \int_V \sum_{k=1}^{n} \frac{\partial}{\partial x_k} (e_k T(x))\, dv = \sum_{k=1}^{n} \frac{\partial}{\partial x_k} (e_k T(x)) \Big|_{x=A}. \tag{4}$$

Ainsi, l'expression $\nabla T(A)$ se déduit en remplaçant formellement, dans l'expression de $qT(x)$, les coordonnées du vecteur q par les symboles de dérivation par rapport aux variables correspondantes. Conformément à la notation du vecteur q sous la forme $q = \sum e_k q_k$,

* Nabla est un ancien instrument de musique, en forme d'un triangle renversé.

l'opérateur ∇ peut être noté $\nabla = \sum e_k \frac{\partial}{\partial x_k}$, ce qui coïncide avec la notation introduite par Hamilton.

On aurait pu considérer la formule (4) comme définition de l'opérateur hamiltonien. Mais alors on aurait à démontrer que l'expression y figurant ne dépend pas du choix de la base orthonormale $e_1, \ldots, e_n$. Notre définition directe (2) n'est pas liée à une base.

4.32. Exemples.

a. Soient $P(x)$ un champ vectoriel dans un domaine $G \subset X = R_n$, $T(x): X \to R_1$ l'opérateur qui agit selon la formule $qT(x) = (q, P(x))$. En posant $P(x) = \{P_1(x), \ldots, P_n(x)\}$, nous trouvons pour la quantité correspondante $\nabla T(x) \equiv (\nabla, P(x))$:

$$(\nabla, P(x)) = \sum_{k=1}^{n} \frac{\partial}{\partial x_k}(e_k, P) = \sum_{k=1}^{n} \frac{\partial P_k}{\partial x_k} = \operatorname{div} P;$$

ainsi, l'opération div est un cas spécial de l'opération ∇.

b. Soient $P(x)$ un champ vectoriel dans un domaine $G \subset X = R_3$ et $T(x): X \to R_3$ l'opérateur qui agit selon la formule

$$qT(x) = [q, P(x)] = \begin{vmatrix} e_1 & e_2 & e_3 \\ q_1 & q_2 & q_3 \\ P_1(x) & P_2(x) & P_3(x) \end{vmatrix}.$$

En remplaçant les symboles q_1, q_2, q_3 par $\frac{\partial}{\partial x_1}$, $\frac{\partial}{\partial x_2}$, $\frac{\partial}{\partial x_3}$, nous aboutissons à l'expression

$$\nabla T(x) = [\nabla, P(x)] = \begin{vmatrix} e_1 & e_2 & e_3 \\ \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \\ P_1(x) & P_2(x) & P_3(x) \end{vmatrix} = \operatorname{rot} P;$$

ainsi, l'opération rot s'avère un autre cas spécial de l'opération ∇.

c. Soient $P(x)$ un champ scalaire dans un domaine $G \subset X = R_n$ et $T(x): X \to X$ l'opérateur agissant selon la formule $qT(x) = q \cdot P(x)$. Ici, pour la quantité correspondante $\nabla T(x) = \nabla P(x)$, nous trouvons

$$\nabla P(x) = \sum_{k=1}^{n} \frac{\partial}{\partial x_k} e_k P(x) = \sum_{k=1}^{n} \frac{\partial P(x)}{\partial x_k} e_k = \operatorname{grad} P.$$

Par conséquent, le gradient est lui aussi un cas spécial de l'opération hamiltonienne.

4.33. Application de l'opérateur ∇ aux produits. L'application de l'opérateur linéaire ∇ à un produit $T_1 T_2$ d'opérateurs, dans l'analyse vectorielle classique, se réalise d'après la règle suivante. On convient que, s'il y a une notation concrète de

l'expression $\nabla T(x)$ où certaines variables (champs vectoriels ou scalaires, etc.) sont placées *à gauche* du symbole ∇, et d'autres *à sa droite*, alors l'opérateur ∇ agit sur seules les variables qui sont *à sa droite*. Par exemple, dans la notation $(P, \nabla) Q$, l'opérateur ∇ agit sur Q et n'agit pas sur P. Parfois on utilise une autre convention : indépendamment de l'ordre des variables, on munit les variables, sur lesquelles le symbole ∇ n'agit pas de l'indice supplémentaire c qui montre que ces variables jouent le rôle de constantes. Par exemple, la notation $(\nabla, P_c) Q$ veut dire que ∇ agit sur Q et n'agit pas sur P. Dans de tels cas, on tâche de transformer l'expression $\nabla T(x)$, si c'est possible, de façon à ce que les variables munies de l'indice c soient à gauche du symbole ∇ ; alors, conformément à la première convention, on pourra omettre l'indice c, car la position même de la variable correspondante dira qu'elle est considérée comme constante. Pour l'exemple ci-dessus, il suffit de permuter ∇ et P_c ; alors, en considérant P comme vecteur et Q comme scalaire, nous obtenons

$$(\nabla, P_c) Q = (P_c, \nabla) Q = (P_c, \nabla Q) = (P, \operatorname{grad} Q).$$

Ceci dit, la règle dont il s'agit est : *le résultat de l'action de l'opérateur hamiltonien sur un produit de deux facteurs est égal à la somme de deux termes dans chacun desquels l'opérateur hamiltonien agit sur un seul des facteurs ; sur l'un dans le premier terme et sur l'autre dans le deuxième.*

Afin de préciser cet énoncé qui est un peu vague et d'en tirer une proposition rigoureuse, considérons les espaces de dimension finie $X = R_n$, Y, Z, U, $V \subset L(X, Z)$, $W \subset L(X, U)$ avec les applications bilinéaires données :

$$\begin{aligned} &\times : V \text{ et } W \text{ dans } L(X, Y), \\ &\bigvee : V \text{ et } U \text{ dans } Y, \\ &\bigwedge : W \text{ et } Z \text{ dans } Y, \end{aligned}$$

et supposons que, pour tout vecteur $q \in X$ et pour n'importe quels opérateurs $R \in V$ et $S \in W$, a lieu la relation fondamentale

$$q(R \times S) = S \bigwedge qR = R \bigvee qS. \tag{1}$$

Supposons ensuite que les opérateurs $R = R(x)$ et $S = S(x)$ dépendent d'un paramètre x parcourant un domaine $G \subset X$ et possèdent dans G les dérivées continues du premier ordre. D'après *4.32*, les opérateurs $R(x)$ et $S(x)$ admettent l'application de l'opérateur ∇, et l'on a

$$\nabla R(x) \subset Z, \quad \nabla S(x) \in U.$$

Théorème. *Dans les conditions ci-formulées, l'opérateur* $T(x) = R(x) \times S(x)$ *admet l'application de l'opérateur* ∇, *et l'on a*

$$\nabla(R(x) \times S(x)) = S(x) \bigwedge \nabla R(x) + R(x) \bigvee \nabla S(x). \tag{2}$$

Démonstration. L'opérateur $T(x)$ est dérivable par rapport à x_j ($j = 1, \ldots, n$) en vertu de *1.34b*, et a lieu la relation

$$\frac{\partial}{\partial x_j}(R(x) \times S(x)) = \frac{\partial R(x)}{\partial x_j} \times S(x) + R(x) \times \frac{\partial S(x)}{\partial x_j}.$$

En utilisant (1), nous avons

$$q \frac{\partial}{\partial x_j}(R(x) \times S(x)) = S(x) \wedge q \frac{\partial R(x)}{\partial x_j} + R(x) \vee q \frac{\partial S(x)}{\partial x_j}.$$

Il en résulte

$$\nabla(R(x) \times S(x)) = \sum_{k=1}^{n} \frac{\partial}{\partial x_k} e_k (R(x) \times S(x)) =$$

$$= \sum_{k=1}^{n} S(x) \wedge e_k \frac{\partial R(x)}{\partial x_k} + \sum_{k=1}^{n} R(x) \vee e_k \frac{\partial S(x)}{\partial x_k} =$$

$$= S(x) \wedge \sum_{k=1}^{n} \frac{\partial}{\partial x_k}(e_k R(x)) + R(x) \vee \sum_{k=1}^{n} \frac{\partial}{\partial x_k}(e_k S(x)) =$$

$$= S \wedge \nabla R + R \vee \nabla S,$$

ce qu'il fallait démontrer.

4.34. Exemples.

a. Soient $R(x)$ et $S(x)$ des champs scalaires dans un domaine $G \subset X = R_n$; trouvons grad $(R(x) \cdot S(x))$.

Quel que soit un vecteur $q \in X$, nous avons évidemment:

$$q(R(x)\, S(x)) = S(x) \cdot qR(x) = R(x) \cdot qS(x).$$

On peut considérer cette relation comme cas spécial de la relation *4.33* (1), pour la concrétisation convenable des espaces et des opérations bilinéaires qui y figurent (il faut poser: $X = Y = Z = U = R_n$, $V = W \subset L(R_n)$ (ensemble des opérateurs de multiplication par un nombre, $\times$ est la multiplication des nombres, $\vee$ et $\wedge$ représentent chacun la multiplication d'un nombre par un vecteur). En appliquant la formule *4.33* (2), nous trouvons:

$$\text{grad}(R(x)\, S(x)) = \nabla(RS) = S(x) \cdot \nabla R(x) + R(x) \cdot \nabla S(x) =$$

$$= S(x) \cdot \text{grad}\, R(x) + R(x) \cdot \text{grad}\, S(x). \qquad (1)$$

Bref, en se servant de l'énoncé initial de la régle, on peut procéder comme suit:

$$\nabla(RS) = \nabla(R_c S) + \nabla(RS_c) = R_c \cdot \nabla S + S_c \nabla R =$$

$$= R \text{ grad } S + S \text{ grad } R.$$

b. Soient $R(x)$ un champ scalaire, $S(x)$ un champ vectoriel dans R_n; trouvons div $(R(x)\, S(x))$. En procédant par analogie avec

ce qui précède, nous avons:

$$\operatorname{div}(RS) = (\nabla, (RS)) = R(\nabla, S) + S, (\nabla, R) = R \operatorname{div} S + (S, \operatorname{grad} R). \tag{2}$$

c. Pour $n = 3$, nous obtenons aussi:

$$\operatorname{rot}(RS) = [\nabla, RS] = R[\nabla, S] - [S, \nabla R] = \\ = R \operatorname{rot} S - [S, \operatorname{grad} R].$$

d. Soient $R(x)$ et $S(x)$ des champs vectoriels dans G. Pour tout vecteur $q \in X$, nous avons

$$(q, [R, S]) = (S, [q, R]) = -(R, [q, S]), \tag{3}$$

$$[q, [R, S]] = R(q, S) - (R, q) S = -S(q, R) + (S, q) R, \tag{4}$$

d'où

$$\operatorname{div}[R, S] = (\nabla, [R, S]) = (S, [\nabla, R]) - (R, [\nabla, S]) = \\ = (S, \operatorname{rot} R) - (R, \operatorname{rot} S), \tag{5}$$

$$\operatorname{rot}[R, S] = [\nabla, [R, S]] = R(\nabla, S) - (R, \nabla) S - S(\nabla, R) + \\ + (S, \nabla) R = R \operatorname{div} S - S \operatorname{div} R - (R, \nabla) S + (S, \nabla) R \tag{6}$$

En coordonnées, l'expression $(R, \nabla) S$ s'écrit:

$$(R, \nabla) S = \sum_{k=1}^{3} R_k \frac{\partial}{\partial x_k} \sum_{j=1}^{3} S_j e_j = \sum_{j=1}^{3} \left(\sum_{k=1}^{3} R_k \frac{\partial S_j}{\partial x_k} \right) e_k; \tag{7}$$

l'expression $(S, \nabla) R$ est notée d'une manière analogue.

4.35. Opérateurs du deuxième ordre. Supposons que dans un domaine G de l'espace $X = R_n$ soit donnée une fonction $T(x)$ dont la valeur pour tout $x \in G$ représente un opérateur bilinéaire qui agit de $X \times X$ dans un espace de Banach Y, de sorte que soit définie la forme bilinéaire $pqT(x)$ $(p \in X, q \in X)$. Supposons que $T(x)$ soit une fonction continue en x, donc que $pqT(x)$ soit continue par rapport à l'ensemble des variables p, q, x.

Conformément à *4.31*, on peut définir l'expression $\nabla qT(x)$ linéaire par rapport à q. En supposant la fonction $T(x)$ dérivable et choisissant arbitrairement une base orthonormale $e_1, \ldots, e_n$ dans X on peut écrire:

$$\nabla qT(x) = \sum_{k=1}^{n} \frac{\partial}{\partial x_k} (e_k qT(x)).$$

Si $T(x)$ est deux fois dérivable dans le domaine G, on peut également remplacer la variable q par ∇, ce qui définit l'expression

$\nabla\nabla T(x)$; dans la même base $e_1, \ldots, e_n$ cette expression a la forme:

$$\nabla\nabla T(x) = \sum_{k=1}^{n}\sum_{j=1}^{n} \frac{\partial^2}{\partial x_k\, \partial x_j}(e_k e_j T(x)). \qquad (1)$$

On peut procéder autrement: remplacer d'abord q par ∇, puis p par ∇; on aura

$$p\nabla T(x) = \sum_{j=1}^{n} \frac{\partial}{\partial x_j}(pe_j T(x)),$$

$$\nabla\nabla T(x) = \sum_{j=1}^{n}\sum_{k=1}^{n} \frac{\partial^2}{\partial x_j\, \partial x_k}(e_k e_j T(x)), \qquad (2)$$

ce qui coïncide avec (1), vu l'indépendance des dérivées mixtes de l'ordre de dérivations. On peut considérer la forme adjointe $qpT(x)$; en y remplaçant p par ∇ et q par ∇ (dans n'importe quel ordre) nous retrouvons l'expression (1) ou (2), ce qui est le même. En réalité, le résultat ne dépend que de la forme quadratique $qqT(x)$, ce qui résulte du lemme suivant:

Lemme. *Si, pour deux opérateurs bilinéaires* $T_1(x)$ et $T_2(x)$ *satisfaisant aux conditions formulées plus haut, on a* $qqT_1(x) \equiv qqT_2(x)$, *alors*

$$\nabla\nabla T_1(x) = \nabla\nabla T_2(x).$$

Démonstration. Considérons l'opérateur $T(x) = T_1(x) - T_2(x)$. Pour les formes bilinéaires nous avons $(p+q)(p+q)T = ppT + pqT + qpT + qqT$. Or, par hypothèse, $(p+q)(p+q)T = 0$, $ppT = 0$, $qqT = 0$, donc

$$pqT + qpT = 0.$$

En remplaçant ici p et q par ∇, et compte tenu de ce qu'on a dit plus haut, nous trouvons

$$\nabla\nabla T + \nabla\nabla T = 2\nabla\nabla T = 0,$$

d'où $\nabla\nabla T_1(x) = \nabla\nabla T_2(x)$, ce qu'il fallait.

4.36. Si $P(x)$ est une fonction scalaire, on peut composer les expressions suivantes:

(a) $(\nabla, \nabla)P \equiv \nabla^2 P = (\nabla, \nabla P) = \text{div grad } P$;

(b) $[\nabla, \nabla]P \equiv [\nabla, \nabla P] = \text{rot grad } P$ (dans R_3).

Si $P(x)$ est un champ vectoriel, on peut construire les expressions:

(c) $(\nabla, \nabla)P = \nabla^2 P$;

(d) $\nabla(\nabla, P) = \text{grad div } P$;

(e) $(\nabla, [\nabla, P]) = \text{div rot } P$;

(f) $[\nabla, [\nabla, P]] = \text{rot rot } P$;

(dans R_3).

Si l'on remplace ∇ par un vecteur ordinaire q, les résultats des opérations (b) et (e) seront nuls d'après les règles de l'algèbre vectorielle.

Nous concluons: *quel que soit un champ scalaire* $P(x)$ *dans* R_3,

$$\text{rot grad } P(x) \equiv 0, \tag{1}$$

et quel que soit un champ vectoriel $P(x)$ *dans* R_3,

$$\text{div rot } P(x) \equiv 0. \tag{2}$$

On peut verifier les égalités (1) et (2) obtenues des raisonnements généraux par un calcul direct:

$$\text{rot grad } P(x) = \begin{vmatrix} e_1 & e_2 & e_3 \\ \dfrac{\partial}{\partial x_1} & \dfrac{\partial}{\partial x_2} & \dfrac{\partial}{\partial x_3} \\ \dfrac{\partial P}{\partial x_1} & \dfrac{\partial P}{\partial x_2} & \dfrac{\partial P}{\partial x_3} \end{vmatrix} = 0,$$

$$\text{div rot } P(x) = \begin{vmatrix} \dfrac{\partial}{\partial x_1} & \dfrac{\partial}{\partial x_2} & \dfrac{\partial}{\partial x_3} \\ \dfrac{\partial}{\partial x_1} & \dfrac{\partial}{\partial x_2} & \dfrac{\partial}{\partial x_3} \\ P_1 & P_2 & P_3 \end{vmatrix} = 0,$$

où $P(x) = e_1P_1(x) + e_2P_2(x) + e_3P_3(x)$.

Considérons l'opération (a). Vu que, pour un vecteur q, nous avons

$$q = \sum_{j=1}^{n} q_j e_j, \quad (q, q) = \sum_{1}^{n} q_j^2, \qquad (q, q) P(x) = \sum_{1}^{n} q_j^2 P(x),$$

en remplaçant les composantes du vecteur par les dérivées par rapport aux coordonnées respectives, nous obtenons:

$$(\nabla, \nabla) P(x) = \sum_{j=1}^{n} \frac{\partial^2 P(x)}{\partial x_j^2}. \tag{3}$$

Cet opérateur différentiel du deuxième ordre est nommé *opérateur de Laplace*; c'est un des plus importants opérateurs différentiels de la physique mathématique.

L'opération (f) s'exprime par (c) et (d). La formule *4.35* (4) donne:

$$[\nabla, [\nabla, P]] = \nabla(\nabla, P) - (\nabla, \nabla) P = \text{grad div } P - \nabla^2 P. \tag{4}$$

§ 4. 4. Certains types de champs vectoriels

4.41. Champs symétriques sphériques.

a. Considérons un champ scalaire dans R_n qui, pour y variant et x fixe, est donné par la formule:

$$u(y) = f(r), \quad r = |y - x| = \sqrt{\sum_{i=1}^{n} (y_i - x_i)^2}, \tag{1}$$

où la fonction $f(r)$ est définie pour $r > 0$, continue, et admet une dérivée continue. Comme la fonction $u(y)$ ne dépend que de la distance de y au point fixe x, le champ u s'appelle *champ symétrique sphérique de centre de symétrie en x.*

b. Calculons le gradient du champ symétrique sphérique $u(y)$. D'après *1.37b* on a

$$\frac{\partial f(r)}{\partial y_i} = f'(r) \frac{\partial r}{\partial y_i} = f'(r) \frac{y_i - x_i}{r},$$

d'où

$$\operatorname{grad}_y f(r) = \sum_{i=1}^{n} \frac{\partial f}{\partial y_i} e_i = \frac{f'(r)}{r} \sum_{i=1}^{n} (y_i - x_i) e_i.$$

Nous munissons le symbole de gradient de l'indice y pour montrer les coordonnées par rapport auxquelles on dérive. Si $e(x, y)$ est le vecteur unitaire allant de x vers y, de sorte que $\sum_{i=1}^{n} (x_i - y_i) e_i = re(x, y)$, on peut mettre le gradient cherché sous la forme:

$$\operatorname{grad}_y f(r) = f'(r) e(x, y). \tag{2}$$

c. D'une façon générale, n'importe quel champ vectoriel $P(x)$ de la forme $\varphi(r) \cdot e(x, y)$, où $r = |y - x|$ et $\varphi(r)$ est une fonction numérique, s'appelle *champ vectoriel symétrique sphérique avec x pour centre de symétrie.* Quelle que soit la fonction donnée $\varphi(r)$ (continue), on peut restituer une fonction $f(r)$ en résolvant l'équation différentielle $f'(r) = \varphi(r)$. Par conséquent, *tout champ vectoriel symétrique sphérique et continu est potentiel, c'est-à-dire qu'il est* (pour $x \neq y$) *gradient d'un champ symétrique sphérique scalaire avec le même centre de symétrie.*

d. Trouvons la divergence d'un champ symétrique sphérique vectoriel $P(y) = \psi(r) e(x, y)$. Il vaut mieux noter $P(y) = \varphi(r) \times re(x, y)$ où $\varphi(r) = \psi(r)/r$. D'après *4.34b* nous avons:

$$\operatorname{div}_y P(y) = (\nabla, \varphi(r) re(x, y)) = \varphi(r) (\nabla_y, re(x, y)) +$$
$$+ (re(x, y), \nabla_y \varphi(r)) = \varphi(r) \operatorname{div}_y re(x, y) + (re(x, y), \operatorname{grad} \varphi(r)).$$

Or,

$$\operatorname{div}_y re(x, y) = \operatorname{div}_y \sum_{i=1}^{n} (y_i - x_i) e_i = \sum_{i=1}^{n} \frac{\partial (y_i - x_i)}{\partial y_i} = n,$$

$$\operatorname{grad}_y \varphi(r) = \varphi'(r) e(x, y),$$

donc,

$$\operatorname{div} P(y) = n\varphi(r) + r\varphi'(r).$$

Est-ce qu'il est possible que $\operatorname{div}_y P(y) = 0$? En résolvant l'équation différentielle

$$n\varphi(r) + r\varphi'(r) = 0,$$

nous trouvons

$$\varphi(r) = \frac{C}{r^n},$$

où C est une constante arbitraire.

Ainsi, *parmi les champs vectoriels symétriques sphériques, les champs de la forme*

$$P(y) = \frac{C}{r^n} re(x, y) = \frac{Ce(x, y)}{r^{n-1}} \tag{3}$$

sont seuls à posséder la divergence nulle (pour $y \neq 0$).

Le flux du champ (3) à travers la sphère S_ρ d'un rayon ρ et de centre au point x vaut

$$\oint_{S_\rho} (m, P)\, dS = C \oint_{S_\rho} \frac{dS}{r^{n-1}} = \frac{C}{\rho^{n-1}} |S_\rho| = C |S_1|,$$

où $|S_1|$ est l'aire de la sphère unitaire. Nous voyons donc qu'à l'intérieur de la sphère S_ρ il existe tout de même une source du champ $P(y)$; il est évident qu'elle ne peut se trouver qu'au point x.

e. Disons entre autres que, pour $n = 3$ et $C = -m$, $m > 0$, le champ (3)

$$P(y) = \frac{Ce(x, y)}{r^2} = \frac{me(y, x)}{r^2}$$

est identique au champ de gravitation d'une masse ponctuelle m placée au point x (loi de Newton!). Nous voyons que la loi de Newton peut *être démontrée* à partir des suppositions naturelles que le champ cherché doive être symétrique sphérique et avoir pour seule source une masse ponctuelle. (Dans un univers bidimensionnel, la force de gravitation serait inversement proportionnelle à la distance, dans un univers n-dimensionnel à la puissance $(n - 1)$-ème de la distance.)

f. Lorsque $n = 3$, le rotationnel d'un champ symétrique sphérique vectoriel $P(y)$, pour $y \neq x$, est nul à cause de l'existence de potentiel. La rotation du champ $P(y)$ sur toute sphère S_ρ de centre au point x est nulle (puisque $[m, P] = 0$), donc le rotationnel s'annule même au point $y = x$.

4.42. Champ newtonien.

a. Selon la loi de Newton dans l'espace tridimensionnel, la force attractive avec laquelle une masse m placée à un point x agit sur la masse unitaire se trouvant en un point y est de la forme

$$F(y) = \frac{m}{|x-y|^2} e(y, x), \tag{1}$$

où $|x-y|$ est la distance de x à y, $e(y, x)$ le vecteur unitaire allant de y vers x; on suppose, naturellement, que le point y soit autre que x.

Le champ $F(y)$ est symétrique sphérique (avec x pour centre de symétrie) et, comme nous l'avons vu dans *4.41c*, potentiel. Evidemment, dans notre cas le potentiel du champ $F(y)$ vaut

$$f(y) = \frac{m}{|x-y|}. \tag{2}$$

b. Si le champ de gravitation est généré non pas par un seul mais par plusieurs masses ponctuelles $m_1, \ldots, m_k$ placées aux points respectifs $x_1, \ldots, x_k$, alors chacune d'elles agit sur la masse unitaire se trouvant en un point y d'après une formule analogue à (1). Par la loi d'addition des forces, l'action totale de toutes les masses est représentée par la somme vectorielle

$$F(y) = \sum_{i=1}^{k} \frac{m_i}{|x_i-y|^2} e(y, x_i), \tag{3}$$

sous la condition que le point y soit différent de chaque point x_i ($i = 1, \ldots, k$).

Le champ (3) est toujours potentiel, et sa fonction potentielle vaut

$$f(y) = \sum_{i=1}^{k} \frac{m_i}{|x_i-y|}. \tag{4}$$

c. Figurons-nous une distribution continue de masse: dans un domaine borné $G \subset R_3$, chaque domaine (jordanien) $V \subset G$ est pourvu d'une masse $m(V)$ qui représente une fonction fortement additive du domaine V. Rappelons la notion de densité de masse. Pour un domaine donné ΔV contenant Δm unités de masse, nous définissons la densité moyenne de masse comme rapport $\frac{\Delta m}{|\Delta V|} = \mu(\Delta V)$. Ensuite, fixons un point x et considérons une suite $\Delta V_1, \ldots, \Delta V_s, \ldots$ de domaines se contractant vers le point x lorspue $s \longrightarrow \infty$. Pour chacun de ces domaines, il y a une densité moyenne de masse $\mu(\Delta V_s)$. Admettons que, quelle que soit une telle suite de domaines, les nombres $\mu(\Delta V_s)$ tendent, pour $s \longrightarrow \infty$, à une limite fixe $\mu = \mu(x) = \lim \mu(\Delta V_s)$ qui ne dépend pas du choix de la suite ΔV_s. La quantité $\mu(x)$ est alors appelée *densité de masse au point x.*

Supposons que, dans un domaine V, la masse admette une densité continue (par morceaux) $\mu(x)$. Construisons l'expression du champ gravitationnel créé par cette masse.

Partageons le domaine V en un nombre de petits domaines ΔV_i ($i = 1, \ldots, k$) et désignons par Δm_i la masse portée par le domaine ΔV_i. Il est naturel d'accepter que la force $\Delta F_i(y)$ avec laquelle la masse élémentaire Δm_i agit sur la masse unitaire au point y soit la même que si cette masse élémentaire était concentrée à un seul point du domaine ΔV_i, par exemple au point x_i. Cela

permet d'appliquer la formule (3) pour la force $\Delta F_i(y)$. En additionnant sur l'indice i et passant à la limite nous aboutissons à l'expression cherchée de la force totale de gravitation :

$$F(y) = \int_V \frac{\mu(x)\, e(y, x)}{|x-y|^2}\, dx. \tag{5}$$

L'intégrale (5) est définie aussi bien pour les points y se trouvant en dehors du domaine V qu'à son intérieur. Dans le dernier cas, l'intégrale devient impropre, mais converge absolument, puisque le dénominateur représente le carré de la quantité $|x-y|$ (*3.73b*).

Le champ $F(y)$ (5) s'appelle *champ newtonien dans* R_3.

d. Dans l'espace n-dimensionnel, on appelle *champ newtonien* un champ vectoriel qui est défini pour $y \in R_n$ par la formule

$$F(y) = \int_V \frac{\mu(x)\, e(y, x)}{|x-y|^{n-1}}\, dx \tag{6}$$

où $V \subset R_n$ est un domaine borné, $\mu(x)$ une fonction continue par morceaux.

Montrons que *le champ* (*6*) *possède un potentiel qui vaut*

$$f(y) = \frac{1}{n-2} \int_V \frac{\mu(x)\, dx}{|x-y|^{n-2}} \quad (n > 2). \tag{7}$$

Il suffit de prouver que grad $f(y) = F(y)$.

En appliquant à l'intégrale (7) le théorème *3.77g* et compte tenu de l'égalité *4.41* (2), on a

$$\operatorname{grad}_y \frac{1}{n-2} \frac{1}{|x-y|^{n-2}} = \sum_{i=1}^{n} \frac{\partial}{\partial y_i} \frac{1}{n-2} \frac{1}{|x-y|^{n-2}} \cdot e_i = \frac{e(y, x)}{|x-y|^{n-1}}$$

et

$$\operatorname{grad} f(y) = \frac{1}{n-2} \sum_{i=1}^{n} \frac{\partial}{\partial y_i} \int_V \frac{\mu(x)\, dx}{|x-y|^{n-2}}\, e_i =$$

$$= \int_V \mu(x) \sum_{i=1}^{n} \frac{\partial}{\partial y_i} \frac{1}{n-2} \frac{1}{|x-y|^{n-2}}\, dx \cdot e_i =$$

$$= \int_V \mu(x) \frac{e(y, x)}{|x-y|^{n-1}}\, dx = F(y).$$

Pour $n = 2$, la fonction $1/((n-2)\,|x-y|^{n-2})$ est à remplacer par $\ln(1/|x-y|)$.

e. Trouvons la divergence du champ newtonien. On peut s'attendre à ce qu'elle soit liée à la fonction $\mu(x)$, parce qu'on a déjà vu que les sources d'un champ de gravitation sont les masses qui le créent.

Si le point y est à une distance positive du domaine V, la dérivation de l'intégrale (6) peut être effectuée sous le signe somme (*3.35d*), car elle a, de même que la fonction $1/|x-y|^{n-1}$, les dérivées de n'importe quel ordre. Il en résulte, compte tenu de *4.41d*, le fait:

$$\operatorname{div} F(y) = \int_V \mu(x) \operatorname{div} \frac{e(y, x)}{|x-y|^{n-1}}\, dx = 0.$$

Si le point y appartient au domaine V, l'intégrale (6) représente une intégrale impropre de deuxième espèce, avec une « singularité variable » (*3.77a*). Pour la dérivation de telles intégrales on utilise le théorème *3.77g*.

Tout de même, ce théorème n'est pas applicable à notre cas, car il demande que l'exposant dans le dénominateur soit inférieur à $n-1$. Pour nous en tirer, procédons autrement. Trouvons d'abord le flux du champ $F(y)$ (6) à travers la surface fermée S limitant un domaine Q, puis divisons-le par le volume $|Q|$ de ce domaine et faisons se contracter le domaine Q vers un point fixe. En calculant le flux on peut appliquer le théorème *3.77d*:

$$\oint_S (m(y), F(y))\, dS(y) = \oint_S \left(m(y), \int_V \frac{\mu(x)\, e(y, x)\, dx}{|x-y|^{n-1}}\right) dS(y) =$$

$$= \int_V \mu(x) \left\{ \oint_S \frac{m(y), e(y, x))}{|x-y|^{n-1}}\, dS(y) \right\} dx.$$

L'intégrale intérieure est le flux à travers la surface S du champ de la masse unitaire placée au point x. Si le point x est en dehors du domaine Q, ce flux est nul, car le champ d'une masse ponctuelle n'a pas de sources en dehors de la masse même. Si le point x est à l'intérieur du domaine Q, le flux ne dépend pas de la forme du domaine Q et se trouve égal au flux à travers la sphère S_1 de rayon 1 centrée au point x, ce qui est facile à calculer:

$$\oint_{S_1} \frac{m(y), e(y, x)}{|x-y|^{n-1}}\, dS(y) = -\oint_{S_1} \frac{dS}{1} = -|S_1|$$

où $|S_1|$ est l'aire de la sphère unitaire dans R_n.

Ainsi, dans le cas en question, on a:

$$\oint_S (m(y), F(y))\, dS(y) = -|S_1| \int_Q \mu(x)\, dx.$$

La fonction $\mu(x)$ est supposée continue (par morceaux); en ses points de continuité nous avons:

$$\operatorname{div} F(y_0) = \lim_{Q \to y_0} \frac{1}{|Q|} \oint_S (m(y), F(y))\, dS(y) =$$

$$= -|S_1| \lim_{Q \to |y_0|} \frac{1}{|Q|} \int_Q \mu(x)\, dx = -|S_1|\, \mu(y). \qquad (8)$$

L'égalité (8) exprime la liaison cherchée entre la divergence du champ de gravitation et la densité de masse générant ce champ.

f. Bien que le champ $F(y)$ (6) possède une divergence, cela ne signifie pas encore qu'il existe les dérivées $\frac{\partial F(y)}{\partial y_i}$. Montrons que, si la fonction $\mu(x)$, au voisinage d'un point $x = y_0$, est continue et admet les dérivées partielles continues, les fonctions $\frac{\partial F(y)}{\partial y_i}$ existent et sont continues pour $y = y_0$.

Lemme. *Pour toute fonction dérivable $f(r)$ où $r = |x - y|$, $x \in R_n$, $y \in R_n$ a lieu l'égalité*

$$\nabla_y f(|x - y|) = -\nabla_x f(|x - y|).$$

Démonstration. D'après *4.41* (2), on a

$$\nabla_y f(r) = \operatorname{grad}_y f(r) = f'(r)\, e(x, y);$$

en y remplaçant y par x on trouve

$$\nabla_x f(r) = \operatorname{grad}_x f(r) = f'(r)\, e(y, x) = -\nabla_y f(r),$$

ce qu'on cherchait.

g. Afin de prouver la dérivabilité du champ $F(y)$ (6), supposons d'abord que la fonction $\mu(x)$ soit continue et possède les dérivées continues non seulement au voisinage du point $x = y$, mais partout dans le domaine V et qu'elle s'annule sur la frontière de V. Mettons le champ $F(y)$ sous forme d'un gradient (d), puis, à l'aide du théorème *3.77g*, portons le gradient sous le signe somme et appliquons le lemme précédent; ensuite transformons la fonction sous le signe somme avec *4.34* (1):

$$F(y) = \nabla_y \frac{1}{n-2} \int_V \frac{\mu(x)\, dx}{|x-y|^{n-2}} =$$

$$= \frac{1}{n-2} \int_V \mu(x)\, \nabla_y \frac{1}{|x-y|^{n-2}}\, dx = \frac{-1}{n-2} \int_V \mu(x)\, \nabla_x \frac{1}{|x-y|^{n-2}}\, dx =$$

$$= -\frac{1}{n-2} \int_V \nabla_x \left(\frac{\mu(x)}{|x-y|^{n-2}} \right) dx + \frac{1}{n-2} \int_V \frac{\nabla_x \mu(x)}{|x-y|^{n-2}}\, dx.$$

Appliquons le théorème d'Ostrogradski *4.13* (7) à la première des intégrales obtenues:

$$\int_V \nabla_x \frac{\mu(x)}{|x-y|^{n-2}}\, dx = \oint_S m(x) \frac{\mu(x)}{|x-y|^{n-2}}\, dS_x = 0,$$

parce que la fonction $\mu(x)$ s'annule sur la frontière S du domaine V. Ainsi,

$$F(y) = \frac{1}{n-2} \int_V \frac{\nabla_x \mu(x)}{|x-y|^{n-2}}\, dx,$$

après quoi l'existence des dérivées premières continues du champ $F(y)$ résulte du théorème *3.77g* et des conditions qu'on a imposées à la fonction $\mu(x)$.

h. Supposons que l'existence des dérivées continues de la fonction $\mu(x)$ ne soit garantie que dans un voisinage du point y_0, par exemple pour $|x - y_0| < 2\delta$. Considérons une fonction dérivable $f(r)$ définie pour $r \geqslant 0$ qui vaut 1 pour $0 \leqslant r \leqslant \delta$ et 0 pour $r \geqslant 2\delta$. Alors la fonction $\mu_1(x) = \mu(x) f(|x - y_0|)$ satisfait aux hypothèses de g, et le champ newtonien correspondant

$$F^*(y) = \int_V \frac{\mu_1(x)\, e(y, x)\, dx}{|x - y|^{n-1}}$$

a les dérivées continues partout dans R_n. Nous avons

$$F(y) - F^*(y) = \int_V \frac{[\mu(x) - \mu_1(x)]\, e(y, x)\, dx}{|x - y|^{n-1}}.$$

Comme la fonction $\mu(x) - \mu_1(x)$ s'annule pour $|x - y_0| < \delta$, le champ $F(y) - F^*(y)$, d'après e, possède au voisinage du point y_0 les dérivées de n'importe quel ordre. Il en résulte que le champ $F(y) = F^*(y) + (F(y) - F^*(y))$ est continûment dérivable au point y_0, ce qu'on affirmait.

i. Ceci étant, si $F = \{F_1, \ldots, F_n\} = \operatorname{grad} f$, alors

$$\operatorname{div} F(y) = \sum_{i=1}^{n} \frac{\partial F_i(y)}{\partial y_i} = (\nabla, F) = (\nabla, \nabla f) = \nabla^2 f(y),$$

ce qu'on a montré dans *4.15*.

En vertu de (8) nous arrivons à la *formule de Poisson* classique pour le champ newtonien $F(y)$:

$$\sum_{i=1}^{n} \frac{\partial F_i(y)}{\partial y_i} = \sum_{i=1}^{n} \frac{\partial^2 f(y)}{\partial y_i^2} = -|S_1|\, \mu(y). \tag{9}$$

Nous avons montré qu'elle est valable en tout point y au voisinage duquel la fonction $\mu(x)$ a les dérivées continues.

j. Cette formule montre, en particulier, comment trouver une solution particulière de l'*équation de Poisson*

$$\Delta f(y) \equiv \sum_{i=1}^{n} \frac{\partial^2 f(y)}{\partial y_i^2} = g(y), \tag{10}$$

où $g(y)$ est une fonction donnée dans un domaine borné $V \subset R_n$, $f(y)$ une fonction inconnue définie dans le même domaine. D'après ce qu'on a démontré, le potentiel newtonien avec la densité de masse

$-g(y)/|S_1|$ satisfait à l'équation (10) en tout point y au voisinage duquel la fonction $g(y)$ possède les dérivées continues.

k. En ce qui concerne le rotationnel du champ newtonien $F(y)$ (6), il est nul puisque le champ $F(y)$ admet un potentiel, et on peut appliquer *4.36* (1).

4.43. Champ de Biot et Savart.

a. Soit $\nu(x)$ un champ vectoriel dans un domaine $V \subset R_3$ limité par une surface lisse par morceaux S ; le *champ de Biot et Savart* qui lui correspond est le champ vectoriel

$$G(y) = \int_V \frac{[\nu(x), e(x, y)]}{|x-y|^2}\, dx. \tag{1}$$

Si le domaine V est rempli de charges électriques en mouvement, avec $\nu(x) = u(x)\frac{dq(x)}{dx}$ pour densité de courant électrique ($dq(x)$ étant la quantité de charge dans un volume dx, $u(x)$ sa vitesse), alors, selon la loi de Biot et Savart, le champ magnétique généré par ce courant est donné par la formule (1), ce qui explique notre terminologie.

b. Le champ $G(y)$ (1) est, en général, rotationnel, donc n'a pas de fonction potentielle. On peut tout de même construire un *potentiel vectoriel* de ce champ, c'est-à-dire un champ vectoriel $J(y)$ tel que $G(y) = \operatorname{rot} J(y)$. Posons

$$J(y) = \int_V \frac{\nu(x)\, dx}{|x-y|}. \tag{2}$$

D'après le théorème *3.77g* et la règle *4.34c*, on a :

$$\operatorname{rot} J(y) = [\nabla, J(y)] = \left[\nabla_y, \int_V \frac{\nu(x)\, dx}{|x-y|}\right] =$$

$$= \int_V \left[\nabla_y, \frac{\nu(x)}{|x-y|}\right] dx = -\int_V \left[\nu(x), \nabla_y \frac{1}{|x-y|}\right] dx =$$

$$= -\int_V \left[\nu(x), \operatorname{grad}_y \frac{1}{|x-y|}\right] dx = \int_V \frac{[\nu(x), e(x, y)]}{|x-y|^2}\, dx = G(y),$$

ce qu'il fallait.

c. Trouvons la quantité $\operatorname{div} J(y)$. En procédant par analogie et appliquant le lemme *4.42f* et le théorème d'Ostrogradski, nous obtenons, une fois la dérivabilité de $\nu(x)$ supposée :

$$\operatorname{div} J(y) = (\nabla, J) = \int_V \left(\nabla_y, \frac{\nu(x)}{|x-y|}\right) dx = \int_V \left((\nu(x), \nabla_y \frac{1}{|x-y|}\right) dx =$$

$$= -\int_V \left(\nu(x), \nabla_x \frac{1}{|x-y|}\right) dx = -\int_V \left(\nabla_x, \frac{\nu(x)}{|x-y|}\right) dx +$$

$$+ \int_V \frac{(\nabla_x, \nu(x))}{|x-y|}\, dx = -\oint_S \left(m(x), \frac{\nu(x)}{|x-y|}\right) dS + \int_V \frac{\operatorname{div} \nu(x)}{|x-y|}\, dx. \tag{3}$$

d. Imposons maintenant au champ $\nu(x)$ les conditions suivantes :

(*) $\operatorname{div} \nu(x) \equiv 0$ (« les charges électriques n'apparaissent pas et ne disparaissent pas ») ;

(**) sur la frontière S du domaine V on a la relation $(m(x), \nu(x)) = 0$ (« la frontière est formée de lignes de courant ; les charges ne le quittent pas et n'y entrent pas »).

L'égalité (3) montre que les conditions (*) et (**) ont pour conséquence

$$\operatorname{div} J(y) = 0. \tag{4}$$

e. Etant données les conditions (*) et (**), calculons le rotationnel et la divergence du champ de Biot et Savart $G(y)$ (1).

Compte tenu de (4) et de *4.36* (4), on a

$$\operatorname{rot} G(y) = [\nabla, [\nabla, J]] = \nabla(\nabla, J) - (\nabla, \nabla) J = \\ = \nabla \operatorname{div} J(y) - \nabla^2 J(y) = -\nabla^2 J(y).$$

Or, les composantes du champ $J(y)$ ont la forme de potentiels newtoniens : si $J = \{J_1, J_2, J_3\}$ et $\nu = \{\nu_1, \nu_2, \nu_3\}$, alors

$$J_k(y) = \int_V \frac{\nu_k(x)\, dx}{|x - y|} \quad (k = 1, 2, 3).$$

Nous avons vu dans *4.42f* que, du moment que la fonction $\mu(x)$ soit dérivable, on a

$$\nabla_y^2 \int_V \frac{\mu(x)\, dx}{|x - y|} = -|S_1| \mu(y) = -4\pi\mu(y) \quad (\text{dans } R_3).$$

Il en résulte $\nabla^2 J(y) = -4\pi\nu(y)$, donc

$$\operatorname{rot} G(y) = 4\pi\nu(y). \tag{5}$$

On peut dire que le rotationnel du champ magnétique d'un courant indique la direction du courant générant ce champ magnétique.

Quant à la divergence du champ de Biot et Savart (1), elle est nulle puisque $G(y) = \operatorname{rot} J(y)$, et la divergence d'un rotationnel est nulle (*4.36* (2)).

Soulignons pour conclure les propriétés opposées des champs de Newton et de Biot et Savart : pour le premier, le rotationnel est nul et la divergence se détermine d'après les fonctions données (densité de masse), pour le second, la divergence est nulle et le rotationnel se détermine d'après les fonctions données (densité de courant) ; le premier possède un potentiel scalaire sans avoir un potentiel vectoriel (la divergence est non nulle) ; le second possède un potentiel vectoriel et n'a pas de potentiel scalaire.

§ 4. 5. Champs et fonctions harmoniques

4.51. Champs harmoniques.

a. Un champ vectoriel $H(x)$ donné dans un domaine $V \subset R_3$ et y possédant les dérivées continues $\frac{\partial H}{\partial x_i}$ $(i = 1, 2, 3)$ est dit *harmonique* si, partout dans le domaine V, on a

$$\operatorname{div} H(x) = 0, \quad \operatorname{rot} H(x) = 0. \tag{1}$$

Un champ newtonien (*4.42*) dans un domaine dépourvu de masse, ainsi que le champ de Biot et Savart (*4.43*) dans un domaine dépourvu de courant, sont des champs harmoniques.

b. La notion de champ harmonique peut être formulée pour les champs vectoriels dans l'espace n-dimensionnel. Un champ vectoriel dérivable $H(x) = \{H_1(x), \ldots, H_n(x)\}$ sera dit *harmonique dans un domaine* $V \subset R_n$ si, partout dans V, on a

$$\operatorname{div} H(x) \equiv \frac{\partial H_1}{\partial x_1} + \ldots + \frac{\partial H_n}{\partial x_n} = 0, \tag{2}$$

$$\frac{\partial H_i}{\partial x_j} - \frac{\partial H_j}{\partial x_i} = 0 \quad (i, j = 1, \ldots, n). \tag{3}$$

Dans ce qui suit, on aura affaire à un champ harmonique dans un domaine simplement connexe $V \subset R_n$. Avant tout, il résulte (*4.21d*) de (3) que le champ H admet un potentiel, i.e. qu'il existe une fonction (deux fois dérivable) $h(x)$ telle que

$$H(x) = \operatorname{grad} h(x) = \left\{ \frac{\partial h}{\partial x_1}, \ldots, \frac{\partial h}{\partial x_n} \right\}. \tag{4}$$

La relation (2) peut être représentée comme condition imposée à la fonction $h(x)$:

$$\sum_{k=1}^{n} \frac{\partial H_k}{\partial x_k} \equiv \sum_{k=1}^{n} \frac{\partial^2 h}{\partial x_k^2} = 0. \tag{5}$$

Toute fonction (deux fois dérivable) $h(x)$ qui satisfait à l'équation (5) dans un domaine V s'appelle *fonction harmonique dans le domaine* V. Ainsi, *la fonction potentielle d'un champ vectoriel harmonique est une fonction harmonique.* Comme, en vertu de *4.21d*, un vecteur $H(x) = \operatorname{grad} h(x)$ satisfait aux égalités (3), la réciproque est également valable: *le gradient de toute fonction harmonique représente un champ vectoriel harmonique.*

4.52. Formules de Green. Soient, dans un domaine $V \subset R_n$ avec la frontière S lisse par morceaux, $\psi(x)$ un champ scalaire dérivable et $R(x) = \operatorname{grad} \varphi(x)$ un champ vectoriel dérivable possédant un potentiel. D'après *4.34b*, nous avons

$$\operatorname{div} \psi R = (\nabla, \psi R) = \psi(\nabla, R) + (R, \nabla\psi) = \psi\Delta\varphi + (\nabla\varphi, \nabla\psi).$$

En intégrant les deux membres sur le domaine V, puis appliquant au premier la formule d'Ostrogradski *4.13* (5), nous obtenons

$$\int_V \operatorname{div} \psi R \, dx = \oint_S (m, \psi R) \, dS = \oint_S \psi (m, R) \, dS =$$

$$= \int_S \psi \frac{\partial \varphi}{\partial m} \, dS = \int_V \psi \Delta \varphi \, dx + \int_V (\nabla \varphi, \nabla \psi) \, dx,$$

de sorte que

$$\int_V (\nabla \varphi, \nabla \psi) \, dx + \int_V \psi \Delta \varphi \, dx = \oint_S \psi \frac{\partial \varphi}{\partial m} \, dS \qquad (1)$$

(*première formule de Green*). En permutant ici φ et ψ puis faisant la soustraction, nous arrivons à la *deuxième formule de Green*:

$$\int_V (\psi \Delta \varphi - \varphi \Delta \psi) \, dx = \oint_S \left(\psi \frac{\partial \varphi}{\partial m} - \varphi \frac{\partial \psi}{\partial m} \right) dS. \qquad (2)$$

4.53. Théorème.

a. *Si une fonction harmonique $h(x)$ s'annule sur la frontière du domaine V, alors $h(x) \equiv 0$ partout à l'intérieur du domaine V.*

b. *Si deux fonctions harmoniques coïncident sur la frontière du domaine V,* elles le font à son intérieur.

c. *Si un champ harmonique $H(x)$ a la composante (m, H) nulle partout sur la frontière du domaine V, alors le champ $H(x)$ est identiquement nul à l'intérieur de V.*

d. *Si deux champs harmoniques $H_1(x)$ et $H_2(x)$ ont, partout sur la frontière du domaine V, leurs composantes normales (m, H_1) et (m, H_2) égales, ces champs coïncident à l'intérieur du domaine V.*

Démonstration. Soit $h(x)$ une fonction harmonique dans le domaine V. Posons $\varphi = \psi = h$ dans la formule de Green *4.52* (1); alors $\Delta h \equiv 0$, et, dans l'hypothèse de *a*, l'intégrale sur la frontière disparaît. Ainsi la formule *4.52* (1) fournit

$$\int_V |\nabla h|^2 \, dx = 0.$$

Il en résulte que $\nabla h = \operatorname{grad} h(x) = 0$ dans le domaine V, donc la fonction $h(x)$ est constante; comme elle est nulle, par hypothèse, sur la frontière, elle l'est partout dans le domaine V. La proposition *a* est donc démontrée.

Soit $H(x)$ un champ harmonique dans le domaine V. Posons $\psi = \varphi = h(x)$ dans la formule de Green *4.52* (1), $h(x)$ étant le potentiel du champ $H(x)$. Alors $\Delta \varphi = 0$ cette fois aussi, et, dans l'hypothèse de *c*, la quantité $\frac{\partial \varphi}{\partial m} = (m, H)$ s'annule sur la frontière de V, de sorte que toute l'intégrale sur S soit nulle. La formule

4.52 (1) fait conclure que

$$\int_V |\nabla h|^2\, dx = \int_V |H(x)|^2\, dx = 0,$$

d'où $H(x) \equiv 0$ dans le domaine V. La proposition c est donc démontrée.

Les assertions b et d résultent de a et c, car la différence de deux fonctions (champs) harmoniques est encore une fonction (un champ) harmonique.

4.54. Autres conséquences des formules de Green.

a. *Si $h(x)$ est une fonction harmonique dans un domaine V, alors*

$$\oint_S \frac{\partial h(x)}{\partial m}\, dS = 0. \tag{1}$$

En effet, la formule (1) est immédiate en posant dans la formule de Green *4.52* (2) $\varphi = h$, $\psi = 1$.

b. Théorème. *La valeur d'une fonction harmonique $h(x)$ au centre y d'une boule $W \subset V$ est égale à la moyenne arithmétique de ses valeurs sur la frontière Σ de la boule W:*

$$h(y) = \frac{1}{|\Sigma|} \oint_\Sigma h(x)\, dS. \tag{2}$$

Démonstration. Soient $W \subset V$ et $Q \subset W$ des boules de centre au point y, de rayons r et ρ, $\rho < r$, de frontières Σ_r et Σ_ρ respectivement. Soit $V_{r\rho} = W - Q$. La frontière S du domaine $V_{r\rho}$ est la réunion de Σ_r et Σ_ρ ; l'opérateur dérivée $\frac{\partial}{\partial m}$ suivant la normale extérieure sur la frontière du domaine W coïncide avec $\frac{\partial}{\partial r}$ sur Σ_r et avec $-\frac{\partial}{\partial r}$ sur Σ_ρ.

Dans la formule de Green *4.52* (2), posons $V = V_{r\rho}$, $\varphi = h$, $\psi = \frac{1}{|x-y|^{n-2}}$; nous savons (*4.51a*) que cette dernière fonction est harmonique par rapport à x pour $x \neq y$. Pour cette raison, l'intégrale de volume dans la formule *4.52* (2) disparaît, et nous obtenons :

$$\int_{\Sigma_r} \left(\frac{1}{r^{n-2}} \frac{\partial h}{\partial r} - h \frac{\partial}{\partial r} \frac{1}{r^{n-2}} \right) dS = \oint_{\Sigma_\rho} \frac{1}{\rho^{n-2}} \frac{\partial h}{\partial r} - h \frac{\partial}{\partial r} \frac{1}{r^{n-2}} \Big) dS,$$

ou bien

$$\frac{1}{r^{n-2}} \oint_{\Sigma_r} \frac{\partial h}{\partial r}\, dS + \frac{n-2}{r^{n-1}} \oint_{\Sigma_r} h\, dS = \frac{1}{\rho^{n-2}} \oint_{\Sigma_\rho} \frac{\partial h}{\partial r}\, dS + \frac{n-2}{\rho^{n-1}} \oint_{\Sigma_\rho} h\, dS.$$

D'après a, les premiers termes dans les deux membres s'annulent. Pour les deuxièmes termes, nous trouvons en introduisant le facteur $1/|S_1|$ ($|S_1|$ étant l'aire de la sphère unitaire dans R_n) :

$$\frac{1}{r^{n-1}|S_1|}\oint\limits_{\Sigma_r} h\,dS = \frac{1}{\rho^{n-1}|S_1|}\oint\limits_{\Sigma_\rho} h\,dS.$$

Notons que $r^{n-1}|S_1| = |\Sigma_r|$, $\rho^{n-1}|S_1| = |\Sigma_\rho|$. En faisant ρ tendre vers 0, et compte tenu de la continuité de la fonction $h(x)$ pour $x = 0$, nous aboutissons à l'égalité cherchée (2).

c. *Une fonction harmonique $h(x)$ qui est définie dans tout l'espace R_n et qui tend vers 0 uniformément pour $|x| \to \infty$ est identiquement nulle.*

En effet, quel que soit un point $y \in R_n$, la valeur $h(y)$ représente, d'après b, la moyenne arithmétique des valeurs de la fonction $h(x)$ sur la sphère de rayon r centrée en y. Lorsque $r \to \infty$ les valeurs de la fonction $h(x)$ sur la sphère tendent uniformément vers 0, par hypothèse. Il en est donc de même de la moyenne arithmétique. Comme cette dernière ne dépend pas de r (elle vaut $h(y)$), elle est nulle, ce qu'on affirmait.

4.55. Représentation d'une fonction harmonique à l'intérieur d'une boule par ses valeurs sur la frontière. Soit $W \subset R_n$ une boule de rayon r et de centre z, qui se trouve à l'intérieur d'un domaine où une fonction harmonique $h(x)$ est définie. En vertu de *4.53b*, les valeurs de la fonction harmonique $h(x)$ à l'intérieur de la boule W sont déterminées d'une façon univoque par ses valeurs sur la frontière Σ de cette boule ; nous tenons à établir une formule explicite exprimant la valeur de la fonction $h(x)$ en un point arbitraire y à l'intérieur de la boule par ses valeurs sur la frontière. Pour $y = z$, on l'a déjà fait dans *4.54b* ; on y considérait une boule $Q \subset W$ centrée en z, et l'on a appliqué la formule de Green *4.52* (2) au domaine $W - Q$ en posant $\varphi = h$, $\psi = \frac{1}{|x-z|^{n-2}}$. L'intégrale de volume a disparu (les deux fonctions étant harmoniques dans $W - Q$), et les intégrales de surface sont devenues beaucoup plus simples, la fonction $\psi(x)$ étant constante sur les frontières de W et Q. Il est impossible d'utiliser cette idée telle quelle pour $y \neq z$, car la fonction $1/|x - z|^{n-2}$ n'est pas constante sur Σ ; mais il s'avère possible de soustraire de $\psi(z)$ une telle fonction harmonique $\psi_0(z)$ que la différence $\Psi(x) = \psi(x) - \psi_0(x)$ soit constante sur Σ et la construction réussira.

Sans nuire à la généralité, on peut supposer que le centre z de la boule W coïncide avec l'origine des coordonnées de l'espace R_n ; alors l'équation de la sphère Σ_r a la forme $|x| = r$. Désignons $y^* = \frac{r^2}{|y|^2}\,y$. Nous affirmons que, pour $x \in \Sigma_r$ le rapport des distances

$|x-y^*|$ et $|x-y|$ est constant. En effet, si $x \in \Sigma_r$, alors $(x, x) = r^2$ et $(x, y^*) = (x, y)\frac{r^2}{|y|^2}$, de sorte que

$$\frac{|x-y|^2}{|x-y^*|^2} = \frac{r^2-2(x, y)+|y|^2}{r^2-2(x, y)\frac{r^2}{|y|^2}+\frac{r^4}{|y|^2}} = \frac{|y|^2}{r^2} = \text{const.} \qquad (1)$$

Pour cette raison, la fonction

$$\Psi(x) = \frac{1}{|x-y|^{n-2}} - \frac{r^{n-2}}{|y|^{n-2}}\frac{1}{|x-y^*|^{n-2}}$$

est constante (même égale à 0) sur la sphère Σ_r.

Notons ensuite que la fonction

$$\psi_0(x) = \frac{r^{n-2}}{|y|^{n-2}}\frac{1}{|x-y^*|^{n-2}}$$

est harmonique partout dans W. Posons $V = W - Q$ dans la formule de Green *4.52* (2) où $Q \subset V$ est une boule de centre y et de rayon ρ, $\varphi = h_1^y(x)$, $\psi = \Psi(x)$. Alors, tout comme dans *4.54b*, l'intégrale de volume disparaît, et nous obtenons

$$-\oint_{\Sigma_r} h(x)\frac{\partial\Psi}{\partial r}\,dS = \oint_{\Sigma_\rho}\left(\Psi\frac{\partial h(x)}{\partial\rho} - h\frac{\partial\Psi(x)}{\partial\rho}\right)dS_\rho,$$

où $\frac{\partial}{\partial\rho}$ désigne la dérivation suivant le rayon menant du point y au point x sur la sphère Σ_ρ de rayon ρ et de centre y. Ensuite,

$$\left|\oint_{\Sigma_\rho}\Psi\frac{\partial h(x)}{\partial\rho}\,dS_\rho\right| \leqslant C_1\frac{1}{\rho^{n-2}}C_2\rho^{n-1} \to 0 \quad \text{pour } \rho\to 0,$$

$$\left|\oint_{\Sigma_\rho} h\frac{\partial\psi_0(x)}{\partial\rho}\,dS\right| \leqslant C\rho^{n-1} \to 0 \qquad \text{pour } \rho\to 0,$$

$$-\oint_{\Sigma_\rho} h\frac{\partial}{\partial\rho}\frac{1}{\rho^{n-2}}\,dS = (n-2)\oint_{\Sigma_\rho} h\frac{1}{\rho^{n-1}}\,dS =$$

$$= \frac{(n-2)|S_1|}{|\Sigma_\rho|}\oint_{\Sigma_\rho} h\,dS \to (n-2)|S_1|h(y) \quad \text{pour } \rho\to 0.$$

En passant à la limite pour $\rho \to 0$ on trouve

$$h(y) = \frac{-1}{(n-2)|S_1|}\oint_{\Sigma_r} h(x)\frac{\partial\Psi(x)}{\partial r}\,dS.$$

Il reste à calculer $\frac{\partial\Psi(x)}{\partial r}$ sur la sphère Σ_r. Comme cette sphère est une surface de niveau de la fonction $\Psi(x)$ et que $\Psi(x)$ croît lorsque

x se déplace vers le centre de la sphère ($\Psi(y) = \infty$), on a $\frac{\partial \Psi(x)}{\partial r} = -|\operatorname{grad} \Psi|$. Pour $|x| = r$, en utilisant (1), on a

$$|\operatorname{grad} \Psi| = \left|\operatorname{grad}_x \frac{1}{|x-y|^{n-2}} - \frac{r^{n-2}}{|y|^{n-2}} \operatorname{grad}_x \frac{1}{|x-y^*|^{n-2}}\right| =$$

$$= (n-2)\left|\frac{x-y}{|x-y|^n} - \frac{r^{n-2}}{|y|^{n-2}} \frac{x-y^*}{|x-y^*|^n}\right| =$$

$$= (n-2)\left|\frac{x-y}{|x-y|^n} - \frac{|y|^2}{r^2} \frac{x-y^*}{|x-y|^n}\right| =$$

$$= \frac{n-2}{|x-y|^n}\left|x-y-\frac{|y|^2}{r^2}\left(x-\frac{r^2}{|y|^2}y\right)\right| = \frac{(n-2)(r^2-|y|^2)}{r|x-y|^n}. \quad (2)$$

Définitivement, la formule cherchée (pour $z = 0$) a la forme

$$h(y) = \frac{r^2-|y|^2}{|S_1|\,r} \oint_\Sigma \frac{h(x)}{|x-y|^n}\, dS\,; \quad (3)$$

elle s'appelle *formule de Poisson.*

4.56. Conséquences de la formule de Poisson.

a. Supposons qu'une fonction harmonique $h(x)$ satisfasse, sur la frontière Σ de la boule W, aux inégalités

$$A \leqslant h(x) \leqslant B.$$

Nous affirmons que *la fonction $h(x)$ satisfait aux mêmes inégalités à l'intérieur de la boule W.* En effet, « le noyau de Poisson »

$$P(x, y) = \frac{r^2-|y|^2}{|S_1|\,r} \cdot \frac{1}{|x-y|^n} \quad (1)$$

est positif dans la boule W; par conséquent, en écrivant la formule de Poisson pour les fonctions harmoniques et non négatives sur la frontière de la boule $B - h(x)$ et $h(x) - A$, nous trouvons

$$B - h(y) = \oint_\Sigma P(x, y)\,(B - h(x))\, dS \geqslant 0,$$

$$h(y) - A = \oint_\Sigma P(x, y)\,(h(x) - A)\, dS \geqslant 0,$$

ce qu'il fallait.

Ainsi, *le maximum et le minimum de n'importe quelle fonction harmonique sur la boule W sont réalisés sur la frontière de la boule.*

b. La formule de Poisson *4.55* (3) a la forme d'intégrale fonction de paramètre y. Comme la fonction $P(x, y)$ (1) peut être indéfiniment dérivable par rapport aux coordonnées du point y, on peut appliquer le théorème *3.35d*; la fonction $h(y)$ qui figure dans le premier membre de la formule de Poisson se trouve donc elle aussi indéfiniment déri-

vable par rapport à y. Ainsi, *toute fonction harmonique est indéfiniment dérivable en tout point de son domaine de définition.*

c. En vertu de l'égalité

$$\Delta \frac{\partial}{\partial x_i} h(x) = \frac{\partial}{\partial x_i} \Delta h(x) = 0,$$

chaque dérivée d'une fonction harmonique est encore une fonction harmonique.

En particulier, *les composantes d'un champ harmonique* $H(x) = \{H_1(x), \ldots, H_n(x)\}$, *en tant que dérivées d'une fonction harmonique* $h(x)$, *i.e. du potentiel du champ* $H(x)$ (*4.51b*), *sont elles-mêmes des fonctions harmoniques.*

d. *Si un champ harmonique* $H(x)$ *est défini dans tout l'espace* R_n *et si* $\lim_{|x|\to\infty} H(x) = 0$, *alors* $H(x) \equiv 0$. En effet, chaque composante du champ $H(x)$ tend vers 0 pour $|x| \to \infty$; en vertu de *c*, ces composantes sont des fonctions harmoniques, et il reste à appliquer *4.54c.*

e. Supposons qu'une fonction arbitraire (continue) $\lambda(x)$ soit donnée sur la sphère Σ qui sert de frontière à la boule W ; *définissons* une fonction $h(y)$, à l'intérieur de la boule W, par la formule :

$$h(y) = \oint_{\Sigma} P(x, y)\, \lambda(x)\, dS, \qquad P(x, y) = \frac{r^2 - |y|^2}{|S_1| r} \cdot \frac{1}{|x-y|^n}.$$

Nous affirmons que $h(y)$ est une *fonction harmonique.* Pour le prouver, il suffit de voir que le noyau de Poisson $P(x, y)$ représente une fonction harmonique (par rapport à y) à l'intérieur de la boule W, puisque les dérivations de la fonction $h(y)$ par rapport à y peuvent être effectuées sous le signe somme. Nous savons que la fonction $\frac{1}{|x-y|^{n-2}}$ est harmonique (*4.51a*) ; en vertu de *c*, ses dérivées sont elles aussi harmoniques ; par conséquent, est harmonique la fonction

$$\frac{1}{|x-y|^{n-2}} - \frac{2}{n-2} \sum_{i=1}^{n} x_i \frac{\partial}{\partial y_i} \frac{1}{|x-y|^{n-2}} = \frac{1}{|x-y|^{n-2}} +$$

$$+ 2 \sum_{i=1}^{n} \frac{x_i (y_i - x_i)}{|x-y|^n} = \frac{1}{|x-y|^n} \left(\sum_{i=1}^{n} (x_i - y_i)^2 + 2 \sum_{i=1}^{n} x_i y_i - 2 \sum_{i=1}^{n} x_i^2 \right) =$$

$$= \frac{1}{|x-y|^n} \sum_{i=1}^{n} (y_i^2 - x_i^2) = \frac{|y|^2 - r^2}{|x-y|^n} = -|S_1| r P(x, y),$$

ce qu'il nous faut.

f. Montrons que la fonction $h(y)$ que nous venons de construire dans *e*, à l'intérieur de la boule W, a pour valeur limite, lorsque le point intérieur y s'approche d'un point x_0 sur la frontière, la quantité $\lambda(x_0)$.

Pour le faire, prouvons que le noyau de Poisson $P(x, y)$ possède les propriétés d'une suite en forme de delta (*3.36*) lorsque $y = y_m \to x_0$, à savoir, les propriétés $P(x, y_m) > 0$ (ce qui est immédiat) et

1) $\oint\limits_{\Sigma} P(x, y_m)\, dS(x) = 1,$

2) $\oint\limits_{\Sigma'} P(x, y_m)\, dS(x) \to 0 \quad$ pour $\quad m \to \infty,$

où Σ' est la surface Σ dépourvue d'un voisinage arbitrairement petit U du point x_0.

La propriété 1) résulte de la formule de Poisson en substituant à $h(x)$ la fonction harmonique qui est identiquement égale à 1. La propriété 2) découle de la relation

$$\lim_{m\to\infty} \frac{r^2 - |y_m|^2}{|x - y_m|^n} = 0$$

qui est uniforme par rapport à $x \in \Sigma - U$, puisque $|y_m| \to r$ et $|x - y_m| \geqslant \text{const} > 0$ en dehors de U.

En appliquant *3.36*, on obtient

$$\lim_{m\to\infty} \oint\limits_{\Sigma} P(x, y_m)\, \lambda(x)\, dS(x) = \lambda(x_0),$$

ce qui achève la démonstration.

g. Théorème. *Pour toute fonction $\lambda(x)$ donnée et continue sur une sphère $\Sigma \subset R_n$ il existe une unique fonction $h(y)$ définie sur la boule W limitée par la sphère Σ, continue sur la boule (fermée) W, harmonique à l'intérieur de W et coïncidant sur la sphère Σ avec la fonction $\lambda(x)$.*

Démonstration. Une fonction vérifiant les propriétés exigées est déjà construite dans *e* et *f*. C'est l'unicité de cette fonction $h(y)$ qu'il reste à établir. Il est impossible d'appliquer à ce cas le théorème *4.53b*, car sa démonstration exige que les fonctions qui y figurent soient deux fois dérivables partout dans W (y compris la frontière de la boule), tandis que les dérivées premières des fonctions que nous considérons puissent s'avérer non continues sur la boule fermée W. Voici la démonstration que nous proposons. Tout d'abord, il suffit de prouver que l'unique fonction $h(y)$ qui remplie l'hypothèse du théorème pour $\lambda(x) \equiv 0$, est identiquement nulle. Désignons par r le rayon de la boule W. Pour $\varepsilon > 0$ donné, il existe, en vertu de la continuité uniforme de $h(y)$, une sphère concentrique $\Sigma_{r-\delta}$ de rayon $r - \delta$, $\delta > 0$ en tout point de laquelle a lieu l'inégalité $|h(y)| \leqslant \varepsilon$. Alors, en appliquant la conséquence *a* à la boule $W_{r-\delta}$ de frontière $\Sigma_{r-\delta}$, l'inégalité $|h(y)| \leqslant \varepsilon$ est également valable pour tous les $y \in W_{r-\delta}$. En faisant tendre ε et δ vers 0 nous voyons que $h(y) \equiv 0$ à l'intérieur de W, ce qu'il fallait.

b. Théorème. *Supposons que, dans une boule $W \subset R_n$, il y ait une suite $h_1(y), \ldots, h_m(y), \ldots$ de fonctions continues dans W et harmoniques à l'intérieur de W, cette suite étant uniformément convergente dans W vers une fonction $h(y)$. Nous affirmons que la fonction $h(y)$ est elle aussi continue dans W et harmonique à l'intérieur de W.*

Démonstration. D'après *e*, *f* et *g*, les fonctions $h_m(y)$ peuvent être mises sous la forme

$$h_m(y) = \oint_{\Sigma} P(x, y)\, h_m(x)\, dS.$$

Dans cette formule, passons à la limite lorsque $m \to \infty$; nous aurons

$$h(y) = \oint_{\Sigma} P(x, y)\, h(x)\, dS,$$

la fonction $h(y)$ étant continue dans la boule fermée W. Alors, d'après *e*, la fonction $h(y)$ sera harmonique à l'intérieur de la boule W, ce qu'il fallait.

4.57. Représentation d'un champ harmonique à l'intérieur d'une boule par les valeurs de sa composante normale sur la frontière de la boule.

a. Soit $W \subset R_n$ une boule de rayon r et de centre à un point z contenue dans l'intérieur d'un domaine V dans lequel un champ harmonique $H(y)$ soit défini. En vertu de *4.53d*, le champ $H(y)$ est déterminée d'une façon univoque à l'intérieur de la boule W d'après les valeurs de sa composante normale sur la frontière de la boule W ; nous tenons à donner une règle explicite pour la construction du champ $H(y)$ d'après sa composante normale sur la frontière.

Soit $h(y)$ un potentiel du champ $H(y)$, i.e. une fonction harmonique telle que $H(y) = \operatorname{grad} h(y)$; évidemment, la composante normale du champ $H(y)$ sur Σ coïncide avec la dérivée normale de la fonction $h(y)$ sur Σ. Si l'on sait trouver la fonction $h(y)$ à l'intérieur de la boule W d'après les valeurs de sa dérivée normale sur Σ, on trouvera le champ même $H(y)$ d'après la formule $H(y) = \operatorname{grad} h(y)$. Pour cette raison, *le problème de la restitution d'un champ harmonique dans W d'après sa composante normale sur Σ se réduit au problème de la recherche d'une fonction harmonique dans W d'après les valeurs de sa dérivée normale sur Σ.*

b. Soit $h(y) \equiv h(y_1, \ldots, y_n)$ une fonction harmonique dans une boule $W \subset R_n$. Nous affirmons que *la fonction* $\varphi(y) = \sum_{i=1}^{n} y_i \frac{\partial h(y)}{\partial y_i}$ *est elle aussi harmonique dans la boule W.*

En effet, on peut dériver indéfiniment la fonction $\varphi(y)$, de même que la fonction $h(y)$ (*4.56b*), et les dérivées $\frac{\partial h(y)}{\partial y_i}$ sont harmoniques de même que $h(y)$ (*4.56c*). Par conséquent,

$$\sum_{j=1}^{n} \frac{\partial^2}{\partial y_j^2}\left(\sum_{i=1}^{n} y_i \frac{\partial h}{\partial y_i}\right) = \sum_{j=1}^{n}\left(\sum_{i=1}^{n} y_i \frac{\partial^3 h}{\partial y_j^2\, \partial y_i} + 2\sum_{i=1}^{n} \frac{\partial^2 h}{\partial y_i^2}\right) = 0.$$

c. La fonction $\varphi(y)$ (*b*) peut être mise sous la forme $\varphi(y) = (y, \text{grad } h)$. Dans ce qui suit, nous supposons (sans restreindre la généralité) que $z = 0$, $r = 1$. Posons $y = \rho x, |x| = 1, 0 \leqslant \rho \leqslant 1$; alors la fonction $\varphi(y)$ prend la forme $\varphi(y) = \rho \frac{\partial h(\rho x)}{\partial \rho}$. Il en découle que la fonction $h(y)$ peut être déterminée d'après $\varphi(y)$ avec la formule

$$h(y) \equiv h(\rho x) = \int_0^{\rho} \frac{\varphi(\tau x)}{\tau}\, d\tau + h(0), \tag{1}$$

la valeur $h(0)$ étant arbitraire.

Or, en posant $\rho = 1$ nous avons

$$\varphi(x) = \frac{\partial h(x)}{\partial \rho},$$

c'est-à-dire que la fonction harmonique $\varphi(y)$ coïncide, sur la frontière de la boule W, avec la dérivée normale de la fonction $h(y)$. Il en résulte une méthode de restitution de la fonction $h(y)$ d'après sa dérivée normale sur Σ : étant donné $\frac{\partial h(x)}{\partial \rho}$, on trouve la fonction $\varphi(y)$ avec la formule de Poisson *4.55* (3) ; ensuite, sachant $\varphi(y)$, on obtient la fonction $h(y)$ d'après la formule (1).

d. Montrons que la méthode signalée permet non seulement de retrouver une fonction harmonique $h(y)$ d'après les valeurs de sa dérivée normale sur Σ, mais aussi de construire une fonction harmonique ayant les valeurs arbitrairement données $\lambda(x)$ de la dérivée normale sur Σ, pourvu que ces valeurs satisfassent à une seule condition

$$\oint_{\Sigma} \lambda(x)\, dS = 0, \tag{2}$$

dont la nécessité découle de *4.54a*.

Soit $\lambda(x)$ une fonction donnée, continue sur Σ et satisfaisant à la condition (2). Posons

$$\varphi(y) = \oint_{\Sigma} P(x, y)\, \lambda(x)\, dS,$$

où $P(x, y)$ est le noyau de Poisson. En vertu de *4.56e*, *f*, la fonction $\varphi(y)$ est continue dans la boule (fermée) W, harmonique à l'intérieur de W, et coïncide avec la fonction $\lambda(x)$ sur la frontière. De plus, il résulte de (2) que

$$\varphi(0) = \oint_{\Sigma} \lambda(x)\, dS = 0.$$

Posons encore

$$h(y) \equiv h(\rho x) = \int_0^{\rho} \frac{\varphi(\tau x)}{\tau}\, d\tau + h(0), \tag{3}$$

avec une valeur arbitraire $h(0)$; comme la fonction $\varphi(y)$ est dérivable pour $y = 0$ et $\varphi(0) = 0$, l'intégrale (3) existe. Montrons que la fonction $h(y)$ (3) est harmonique à l'intérieur de W, continue dans la boule fermée W. Le dernier découle, évidemment, de la continuité de la fonction $\varphi(\tau x)/\tau$; les valeurs limites de la fonction $h(y)$ sont données par la formule

$$h(x) = \int_0^1 \frac{\varphi(\tau x)}{\tau}\, d\tau. \tag{4}$$

Pour démontrer que la fonction $h(y)$ est harmonique, il suffit, vu *4.56e*, de prouver qu'elle peut être restituée, avec la formule de Poisson, à partir de ses valeurs limites (4). En posant, pour simplifier, $h(0) = 0$, nous avons

$$\oint_{\Sigma} P(s, y)\, h(s)\, dS = \oint_{\Sigma} P(s, y) \left\{ \int_0^1 \frac{\varphi(\tau s)}{\tau}\, d\tau \right\} dS =$$

$$= \int_0^1 \left\{ \oint_{\Sigma} P(s, y)\, \frac{\varphi(\tau s)}{\tau}\, dS \right\} d\tau.$$

Or, la fonction $\varphi(\tau y)/\tau$ est harmonique avec $\varphi(y)$ (τ étant fixe) ; donc, l'intégrale intérieure vaut $\varphi(\tau y)/\tau$, et, après la substitution $\xi = \rho\tau$, on a

$$\oint_{\Sigma} P(s, y)\, h(s)\, dS = \int_0^1 \frac{\varphi(\tau y)}{\tau}\, d\tau = \int_0^{\rho} \frac{\varphi(\xi x)}{\xi}\, d\xi = h(y),$$

ce qu'il fallait.

Enfin, nous avons $\frac{\partial h(\rho x)}{\partial \rho} = \frac{\varphi(\rho x)}{\rho}$, d'où, pour $\rho = 1$, $\frac{\partial h(x)}{\partial \rho} = \varphi(x) = \lambda(x)$, ce qui achève la démonstration.

e. Résumons les résultats *b-d* en le théorème suivant :

Théorème. *Pour toute fonction $\lambda(x)$ donnée et continue sur une sphère $\Sigma \subset R_n$ et vérifiant la condition*

$$\oint_{\Sigma} \lambda(x)\, dS = 0,$$

il existe une fonction $h(y)$ continue dans la boule W limitée par la sphère Σ, harmonique à l'intérieur de W et possédant la dérivée continue $\frac{\partial h(\rho x)}{\partial \rho}$ $(x \in \Sigma,\ 0 < \rho \leqslant 1)$ qui devient égale à $\lambda(x)$ pour $\rho = 1$. Toute autre fonction ayant les mêmes propriétés diffère de $h(y)$ d'un terme constant.

L'existence de la fonction cherchée $h(y)$ est établie dans *d.* Son unicité, à une constante additive près, résulte de *c.*

§ 4. 6. Construction d'un champ vectoriel dans R_3 d'après ses rotationnel et divergence

4.61. Soient $R(y)$ un champ vectoriel et $b(y)$ un champ scalaire dans un domaine borné $V \subset R_3$. On se demande *s'il est possible de construire, dans le domaine V, un champ vectoriel $Q(y)$ possédant les propriétés*

$$\operatorname{div} Q(y) = b(y), \qquad \operatorname{rot} Q(y) = R(y)? \tag{1}$$

Si oui, comment décrire tous les champs vectoriels vérifiant les équations (1)?

Le problème formulé s'appelle *problème de construction d'un champ vectoriel d'après ses rotationnel et divergence*, ou bien *problème inverse de l'analyse vectorielle.*

Nous supposons que les champs donnés $R(y)$ et $b(y)$ soient dérivables dans le domaine V. De plus, nous exigeons que la divergence du champ $R(y)$ soit nulle partout dans V, et que, sur la frontière de V, la composante normale de ce champ s'annule. La première de ces conditions est nécessaire pour la résolubilité de notre problème, puisque div rot $Q(y) = 0$, quel que soit un champ $Q(y)$. Quant à la deuxième condition, elle est due à seule la méthode de résolution (cf. plus bas, *4.66*).

4.62. Commençons par construire une solution particulière du système (1).

Considérons le cas où le champ donné $R(y)$ est nul, de sorte que le système (1) se réduit à un système plus simple:

$$\operatorname{div} Q(y) = b(y), \qquad \operatorname{rot} Q(y) = 0. \tag{2}$$

Nous connaissons une solution de ce système. En effet, construisons le champ de gravitation avec $-\frac{1}{4\pi}\, b(x)$ pour densité (*4.42* (5)):

$$F(y) = -\frac{1}{4\pi} \int_V \frac{b(x)\, e(y, x)}{|x - y|^2}\, dx. \tag{3}$$

On a vu dans *4.42e* et *k* que

$$\operatorname{div} F(y) = b(y), \qquad \operatorname{rot} F(y) = 0,$$

de sorte que le champ $F(y)$ (3) représente une solution particulière du système (2).

4.63. Considérons l'autre cas où le champ scalaire donné $b(y)$ est nul, de sorte que le système (1) prend la forme

$$\operatorname{div} Q(y) = 0, \qquad \operatorname{rot} Q(y) = R(y). \tag{4}$$

Là aussi, on connaît une solution particulière. Construisons le champ de Biot et Savart pour la densité de courant $\nu(x) = \frac{1}{4\pi} R(x)$ (*4.43a*) ;

$$G(y) = \frac{1}{4\pi} \int_V \frac{[R(x), e(y, x)]}{|x-y|^2}\, dx. \tag{5}$$

Vu *4.43e* et compte tenu des conditions imposées au champ $R(y)$, on a

$$\operatorname{div} G(y) = 0, \qquad \operatorname{rot} G(y) = R(y),$$

de sorte que le champ $G(y)$ fournit une solution particulière du système (4).

4.64. Etant donnés les champs construits $F(y)$ et $G(y)$, posons

$$Q_0(y) = F(y) + G(y). \tag{6}$$

Pour le champ $Q_0(y)$ on a :

$$\operatorname{div} Q_0(y) = \operatorname{div} F(y) + \operatorname{div} G(y) = b(y),$$
$$\operatorname{rot} Q_0(y) = \operatorname{rot} F(y) + \operatorname{rot} G(y) = R(y).$$

Ainsi, le champ $Q_0(y)$ est une solution particulière du système (1). Soit $Q_1(y)$ une autre solution du système (1). Pour le champ $H(y) = Q_1(y) - Q_0(y)$ nous trouvons :

$$\operatorname{div} H(y) = \operatorname{div} Q_1(y) - \operatorname{div} Q_0(y) = 0,$$
$$\operatorname{rot} H(y) = \operatorname{rot} Q_1(y) - \operatorname{rot} Q_0(y) = 0,$$

donc, $H(y)$ est un champ harmonique (*4.51*). Inversement, si $H(y)$ est un champ harmonique et $Q_1(y) = Q_0(y) + H(y)$, alors

$$\operatorname{div} Q_1(y) = \operatorname{div} Q_0(y) + \operatorname{div} H(y) = b(y),$$
$$\operatorname{rot} Q_1(y) = \operatorname{rot} Q_0(y) + \operatorname{rot} H(y) = R(y),$$

de sorte que le champ $Q_1(y)$ représente, avec le champ $Q_0(y)$, une solution du système (1). Nous aboutissons au théorème suivant:

Théorème. *Dans les conditions formulées, le système* (1) est *compatible. L'une de ses solutions est donnée par la formule* (6), *les autres s'obtiennent en y ajoutant un champ harmonique quelconque.*

4.65. On peut imposer à la solution cherchée $Q(y)$ certaines conditions supplémentaires qui la détermineront d'une façon univoque.

Par exemple, supposons qu'à tout point x sur la frontière S du domaine V soit donnée la composante normale $(Q(x), m(x))$ du champ cherché $Q(y)$.

En écrivant la solution $Q(y)$ comme $Q_0(y) + H(y)$ où $Q_0(y)$ est la solution (6) et $H(y)$ un champ harmonique inconnu, nous voyons que la composante normale sur S de ce dernier champ est :

$$(H(x), m(x)) = (Q(x), m(x)) - (Q_0(x), m(x)). \qquad (7)$$

Si le domaine V est une boule W, alors, d'après *4.57a* et *e*, le champ harmonique $H(x)$ satisfaisant à la condition (7) existe et est unique ; par conséquent, existe et est unique la solution du système (1).

D'une façon analogue, conformément à *4.56g*, la solution de ce système existe et est unique lorsqu'on donne, sur la frontière de la boule W, les valeurs du potentiel du champ $H(x)$ (ou bien du champ $Q(x)$, s'il devient harmonique à la proximité de la frontière).

Mais dans le cas général où le domaine V n'est plus une boule, l'existence et l'unicité de solutions pour l'une des conditions formulées plus haut représente un problème assez compliqué ; c'est un des plus importants problèmes de la théorie des équations à dérivées partielles (premier ou deuxième problèmes aux limites pour l'équation de Laplace), et nous n'y touchons pas dans le présent livre *.

Une condition d'un autre type consiste en ce que *le champ cherché* $Q(y)$ *que l'on considère dans tout l'espace* R_n *tend uniformément vers* 0 *pour* $|y| \to \infty$. Il est évident que le champ $Q_0(y)$ que nous avons construit satisfait à cette condition car les expressions (3) et (5) contenant dans le dénominateur la quantité $|x - y|^2$ montrent directement que les intégrales correspondantes tendent vers 0 pour $|y| \to \infty$. Nous affirmons qu'*il n'existe aucune autre solution du système* (1) *possédant ladite propriété.* En effet, la différence d'une telle solution et de la solution $Q_0(y)$ serait un champ harmonique ; or, d'après *4.56d*, le seul champ harmonique qui tend vers 0 pour $|y| \to \infty$ est identiquement nul, d'où l'unicité de solution du système (1) dans la classe considérée des champs.

4.66. Le problème de construction d'un champ d'après la divergence et le rotationnel peut être formulé localement : il faut construire un champ $Q(y)$ qui satisfasse aux équations *4.61* (1) au voisinage d'un point donné $y_0 \in V$; on sait que div $R(y) \equiv 0$, mais il n'y a aucune hypothèse sur la composante normale du champ $R(y)$ sur la frontière du domaine V.

Ce problème se réduit à la construction au voisinage du point y_0 d'un champ $G(y)$ au moins, tel que rot $G(y) = R(y)$. En effet, ayant

* Cf., par exemple, И. Г. Петровский, Лекции об уравнениях с частными производными, изд. 5, « Наука », 1970.

un tel champ nous désignons $\varphi(y) = \operatorname{div} G(y)$, alors le champ cherché $Q(y)$ aura la forme du la somme de champ $G(y)$ et du champ newtonien de densité $-\frac{1}{4\pi}(b(y) - \varphi(y))$.

On peut déterminer les composantes $G_1(y)$, $G_2(y)$, $G_3(y)$ du champ cherché $G(y)$ d'après les composantes $R_1(y)$, $R_2(y)$, $R_3(y)$ de $R(y)$, au voisinage du point $y_0 = (y_1^0, y_2^0, y_3^0)$, à l'aide des formules, par exemple,

$$G_1(y) = \int_{y_3^0}^{y_3} R_2(y_1, y_2, \tau_3)\, d\tau_3,$$

$$G_2(y) = -\int_{y_3^0}^{y_3} R_1(y_1, y_2, \tau_3)\, d\tau_3 + \int_{y_1^0}^{y_1} R_3(\tau_1, y_2, y_3^0)\, d\tau_1,$$

$$G_3(y) \equiv 0,$$

En effet, il en découle:

$$\frac{\partial G_3}{\partial y_2} - \frac{\partial G_2}{\partial y_3} = R_1, \quad \frac{\partial G_1}{\partial y_3} - \frac{\partial G_3}{\partial y_1} = R_2;$$

en utilisant la condition $\operatorname{div} R = \frac{\partial R_1}{\partial y_1} + \frac{\partial R_2}{\partial y_2} + \frac{\partial R_3}{\partial y_3} = 0$, nous trouvons:

$$\frac{\partial G_2}{\partial y_1} - \frac{\partial G_1}{\partial y_2} = -\int_{y_3^0}^{y_3} \frac{\partial R_1(y_1, y_2, \tau_3)}{\partial y_1}\, d\tau_3 + R_3(y_1, y_2, y_3^0) -$$

$$-\int_{y_3^0}^{y_3} \frac{\partial R_2(y_1, y_2, \tau_3)}{\partial y_3}\, d\tau_3 = R_3(y_1, y_2, y_3^0) +$$

$$+\int_{y_3^0}^{y_3} \frac{\partial R_3(y_1, y_2, \tau_3)}{\partial y_3}\, d\tau_3 = R_3(y_1, y_2, y_3),$$

ce qu'il nous fallait.

EXERCICES

1. Trouver le rotationnel du champ $P(A) = f(\rho)\, \tau(A)$, où ρ est la distance du point A à une droite fixe λ, $\tau(A)$ le vecteur unitaire orthogonal à λ ainsi qu'à la perpendiculaire abaissée du point A sur la droite λ. Dans quel cas $\operatorname{rot} P(A) \equiv 0$?

2. Pour une famille uniparamétrique de courbes suffisamment lisses sur le plan, démontrer la relation

$$\operatorname{div} m = -k,$$

où m est le vecteur normal unitaire, k la courbure d'une courbe de la famille au point correspondant.

3. Montrer qu'un champ vectoriel R dans R_3 qui possède les surfaces orthogonales vérifie la relation

$$(R,\ \text{rot}\ R) \equiv 0.$$

4. Soit $(R,\ \text{rot}\ R) \equiv 0$. Traçons par un point donné A une ligne L orthogonale aux lignes du champ R, et par chaque point M de la ligne L une courbe $\gamma\ (M)$ tangente en tout son point au vecteur rot R. Démontrer que la surface obtenue S est orthogonale au champ R.

5. Soit $(R,\ \text{rot}\ R) = 0$. Construire une famille uniparamétrique de surfaces orthogonales au champ R.

6. Le champ vectoriel $R = \{X,\ Y\}$ dans le plan $R_2 - \{0\}$, avec

$$X = \frac{y}{x^2+y^2}, \quad Y = -\frac{x}{x^2+y^2},$$

satisfait à la condition $\dfrac{\partial X}{\partial y} = \dfrac{\partial Y}{\partial x}$, mais n'a pas de potentiel défini dans tout le domaine de définition du champ. Comment expliquer la contradiction apparente avec le théorème *4.21d*?

7. Le champ $R\ (M) = \dfrac{e\ (0, M)}{r^2\ (0, M)}$ dans $R_3 - \{0\}$ a une divergence nulle. En considérant la sphère de rayon 1 centrée à l'origine des coordonnées et dépourvue d'un petit voisinage d'un point fixe, montrer que le champ $R\ (M)$ n'a pas de potentiel vectoriel.

8. Calculer le champ newtonien créé par une boule homogène de masse 1 et de rayon r, centrée à l'origine des coordonnées.

9. Montrer qu'une fonction harmonique non négative $h\ (x)$ satisfait, dans une boule de rayon r et de centre à l'origine des coordonnées de l'espace R_n, à l'*inégalité de Harnack*

$$\frac{(r-|\,y\,|)\,r^{n-2}}{(r+|\,y\,|)^{n-2}}\,h\,(0) \leqslant h\,(y) \leqslant \frac{(r+|\,y\,|)\,r^{n-2}}{(r-|\,y\,|)^{n-2}}\,h\,(0).$$

10. Montrer qu'une fonction harmonique et non négative dans tout l'espace R_n est constante.

11. Supposons qu'une fonction harmonique $h\ (y)$, dans une boule $W \subset R_n$ de rayon r, satisfasse à l'inégalité $|\,h\ (y)\,| \leqslant M$. Montrer qu'au centre z de la boule W a lieu l'inégalité

$$|\,\text{grad}\ h\ (z)\,| \leqslant \frac{n}{r}\,M.$$

12. Supposons qu'une fonction harmonique $h\ (y)$ satisfasse à l'inégalité $|\,h\ (y)\,| \leqslant M$ dans un domaine $V \subset R_n$. Démontrer que, dans chaque domaine fermé $\overline{W} \subset V$, a lieu l'inégalité

$$|\,\text{grad}\ h\ (y)\,| \leqslant CM,$$

où C ne dépend pas du choix de la fonction $h\ (y)$.

13. Si un ensemble infini $\{h\ (y)\} = B$ de fonctions harmoniques est uniformément borné dans un domaine $V \subset R_n$, on peut en extraire une suite $h_1\ (y)$, $h_2\ (y)$, ... de fonctions qui converge uniformément dans chaque domaine fermé $\overline{W} \subset V$.

14. Montrer qu'une suite monotone $h_1\ (y) \geqslant h_2\ (y) \geqslant \ldots$ de fonctions harmoniques dans une boule $|\,y\,| < r$, qui converge au centre de la boule, converge dans toute la boule vers une fonction harmonique.

15. Peut-on appliquer la formule de Stokes au ruban de Möbius?

Historique

Les formules qui lient une intégrale sur la frontière d'un domaine à une intégrale sur ce domaine lui-même sont établies:

la formule *4.16* (3) (« formule de Green ») par Euler en 1771-1772, par Green en 1828;

la formule *4.13* (5) (« formule d'Ostrogradski ») par Gauss pour un cas très spécial en 1813, par Ostrogradski en 1828 ($n = 3$) et 1834 (n arbitraire);

la formule *4.26* (6) (« formule de Stokes ») par Thomson en 1849; Stokes l'inclut dans l'examen annuel qu'il organise à Cambridge de 1849 à 1882;

les formules *4.52* (1) et (2) (« formules de Green ») par Green en 1828.

Il va de soi que, sous leur forme initiale, toutes ces formules s'écrivent dans les désignations habituelles, sans vecteurs. Le calcul vectoriel apparaît chez Hamilton (son ouvrage principal « Lectures on quaternions » est publié en 1853) en tant que partie de sa théorie des quaternions. Dans le langage moderne, les quaternions forment une algèbre quadridimensionnelle sur le corps R_1, avec une base $1, i, j, k$ et les règles suivantes de multiplication: $i^2 = j^2 = k^2 = -1$; $ij = -ji = k$, $jk = -kj = i$, $ki = -ik = j$. Pour un quaternion $a + bi + cj + dk$, le nombre a représente, selon Hamilton, un « scalaire », et $bi + cj + dk$ un « vecteur » (c'est ici que Hamilton introduit ces termes pour la première fois). La multiplication des vecteurs, en tant que quaternions, conduit à l'égalité

$$(b_1 i + c_1 j + d_1 k)(b_2 i + c_2 j + d_2 k) = (b_1 b_2 + c_1 c_2 + d_1 d_2) + \\ + i(c_1 d_2 - d_1 c_2) + j(d_1 b_2 - d_2 b_1) + k(b_1 c_2 - b_2 c_1),$$

où nous voyons, en même temps, les produits « scalaire » et « vectoriel » des vecteurs.

L'opérateur quaternionique de Hamilton

$$\nabla = i\frac{\partial}{\partial x} + j\frac{\partial}{\partial y} + k\frac{\partial}{\partial z}$$

agit sur une fonction vectorielle $f = iu + jv + kw$ d'après la règle formelle de multiplication des quaternions et amène au résultat:

$$\nabla f = \left(i\frac{\partial}{\partial x} + j\frac{\partial}{\partial y} + k\frac{\partial}{\partial z}\right)(iu + jv + kw) = -\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z}\right) + \\ + i\left(\frac{\partial w}{\partial y} - \frac{\partial v}{\partial z}\right) + j\left(\frac{\partial u}{\partial z} - \frac{\partial w}{\partial x}\right) + k\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right).$$

Dans la partie scalaire, nous voyons la divergence (avec le signe moins) connue du temps d'Euler; la partie vectorielle, i.e. le rotationnel, représente, à l'époque de Hamilton, une nouveauté, son interprétation physique n'est pas encore trouvée. Hamilton rêve d'une théorie des fonctions d'une variable quaternionique qui généralise le calcul des fonctions d'une variable complexe. Mais le sort ne veut pas que les espérances qu'il place en quaternions se réalisent. La partie algébrique de la théorie de Hamilton se trouve entièrement incluse dans la théorie des matrices se développant très vite. De la partie analytique on dégage l'analyse vectorielle « non quaternionique » (J. Gibbs, 1881) qui commence à jouer un rôle important dans la physique mathématique. Et si, avant Gibbs, plusieurs savants de renom considèrent avec prudence les vecteurs (ainsi, J. Maxwell, le créateur de la théorie d'électromagnétisme, n'utilise pas les vecteurs et écrit les équations fondamentales de l'électrodynamique, pour la première fois, en forme scalaire), alors H. Hertz, le successeur de Maxwell, met

déjà ces équations sous la forme vectorielle transparente (1890):

$$\frac{\partial E}{\partial t} = -c \operatorname{rot} H, \qquad \frac{\partial H}{\partial t} = +c \operatorname{rot} E$$

(pour le vide). Ce système d'équations admet, entre autres, une solution de la forme $H = F(x - ct)\, j$, $E = F(x - ct)\, k$, avec la fonction F arbitraire; elle représente une onde transversale de profil F qui se propage le long de l'axe des x, à la vitesse c. La réalisation expérimentale de telles ondes par Hertz, puis la découverte de la radio par A. Popov (1895) fournissent une confirmation brillante de la théorie de Maxwell et une manifestation de l'importance primordiale des mathématiques pour le progrès des sciences.

La fonction potentielle du champ newtonien est indiquée par Lagrange en 1773. L'équation $\Delta u = 0$ peut être rencontrée dans les travaux d'Euler, mais c'est Laplace qui l'étudie systématiquement (depuis 1782). L'équation $\Delta u = -4\pi\rho$ pour le champ newtonien est déduite par Poisson en 1813; c'est toujours lui qui construit une fonction harmonique (terme de Laplace) dans une boule d'après ses valeurs sur la frontière de la boule. La construction d'une fonction harmonique d'après les valeurs sur la frontière de sa dérivée normale est réalisée par Neumann en 1877. L'état actuel des problèmes aux limites pour les fonctions harmoniques est décrit, par exemple, dans le livre И. Г. Петровский, Лекции об уравнениях с частными производными, изд. 5, « Наука », 1970.

DEUXIÈME PARTIE

Des espaces vectoriels aux variétés différentiables

Géométrie différentielle classique

La géométrie — plus exactement, la géométrie différentielle — déjà sur la surface bidimensionnelle dans l'espace tridimensionnel, est riche en idées, faits, positions de problèmes admettant une vaste généralisation; en même temps, elle sert de champ d'application naturel aux méthodes de l'analyse mathématique. Certains faits d'un caractère général sont plus naturels quand on les expose pour la surface multidimensionnelle, ce que nous ferons chaque fois qu'une telle possibilité se présente. La première forme quadratique (§ 5.1) introduit une métrique sur la surface m-dimensionnelle dans l'espace euclidien n-dimensionnel en l'empruntant, à l'echelle infiniment petite, à cet espace. Au début, une telle métrisation de la surface ne laisse rien à désirer. La deuxième forme quadratique (§ 5.2) qu'on utilise pour calculer la courbure de lignes se trouvant sur la surface demande déjà que la dimension de la surface et celle de tout l'espace ne diffèrent que d'une unité. C'est moyennant la deuxième forme quadratique qu'on définit la courbure de la surface même qui est l'une des plus importantes caractéristiques de la surface; les surfaces dont les courbures sont de signes contraires possèdent les propriétés géométriques essentiellement différentes. Ensuite, une métrique sur la surface permet de mettre en évidence sa connexion, i.e. la dépendance des propriétés locales de la position du point sur la surface. A l'aide de la connexion on définit les lignes géodésiques (§ 5.4), la translation sur la surface (§ 5.6) et, en fin des comptes, la courbure en tant que résultat de la rotation d'un vecteur translaté le long d'un contour fermé. La courbure s'avère une propriété intrinsèque de la surface, c'est-à-dire ne dépend que de la métrique (i.e. de la première forme quadratique) et ne dépend pas de la façon dont la surface est immergée dans l'espace qui l'entoure.

§ 5. 1. Première forme quadratique

5.11. Dans l'espace n-dimensionnel R_n, une surface m-dimensionnelle P est déterminée par n fonctions exprimant les coordonnées $x_1, \ldots, x_n$ du point de la surface par les paramètres $u_1, \ldots, u_m$, le point paramétrique $u = (u_1, \ldots, u_m)$ parcourant un domaine $U \subset R_m$:

$$\left.\begin{aligned} x_1 &= x_1(u_1, \ldots, u_m), \\ &\cdots\cdots\cdots\cdots \\ x_n &= x_n(u_1, \ldots, u_m). \end{aligned}\right\} \tag{1}$$

On suppose que les fonctions $x_1(u), \ldots, x_n(u)$ sont dérivables un nombre suffisant de fois (combien, on le dira au début de chaque

paragraphe). Il est possible de remplacer les équations (1) par une équation vectorielle

$$r = r(u) = \{x_1(u), \ldots, x_n(u)\}.$$

Formellement, il est possible qu'à deux points paramétriques différents $u = (u_1, \ldots, u_m)$ et $u' = (u'_1, \ldots, u'_m)$ il corresponde un même point de la surface P. Mais, pour éviter des complications liées aux auto-intersections, nous ne considérons pas de telles situations. La plus simple supposition que nous acceptons et qui garantit la biunivocité locale de l'application $u \to r(u)$ consiste à ce que le rang de la matrice de Jacobi

$$\frac{\partial(x_1, \ldots, x_n)}{\partial(u_1, \ldots, u_m)} = \left\| \begin{array}{ccc} \frac{\partial x_1}{\partial u_1} & \cdots & \frac{\partial x_n}{\partial u_1} \\ \cdots & \cdots & \cdots \\ \frac{\partial x_1}{\partial u_m} & \cdots & \frac{\partial x_n}{\partial u_m} \end{array} \right\| \tag{2}$$

à un point u^0 du domaine U soit égal à m. Soit, par exemple,

$$\left\| \begin{array}{ccc} \frac{\partial x_1}{\partial u_1} & \cdots & \frac{\partial x_m}{\partial u_1} \\ \cdots & \cdots & \cdots \\ \frac{\partial x_1}{\partial u_m} & \cdots & \frac{\partial x_m}{\partial u_m} \end{array} \right\|$$

un mineur non nul. Alors on voit de *1.74b* que, dans un voisinage V du point $u^0 = (u_1^0, \ldots, u_m^0)$, l'application (1) est biunivoque, i.e. à deux points *différents* $(u_1, \ldots, u_m)$ et $(u'_1, \ldots, u'_m)$ du domaine V il correspond les points *différents* de la surface P. C'est dans ce voisinage V de variation des paramètres que nous étudions notre surface. Nous adoptons donc un point de vue local : la surface n'est considérée qu'au voisinage d'un point fixe. Mais cela suffit largement pour les premiers théorèmes de la théorie des surfaces.

5.12. Supposons que le rang de la matrice *5.11* (2), à un point donné $u^0 \in U$, vaille m. Il en résulte que les vecteurs

$$r_1 \equiv \frac{\partial r}{\partial u_1} = \left\{ \frac{\partial x_1}{\partial u_1}, \ldots, \frac{\partial x_n}{\partial u_1} \right\},$$

$$\ldots\ldots\ldots\ldots\ldots\ldots$$

$$r_m \equiv \frac{\partial r}{\partial u_m} = \left\{ \frac{\partial x_1}{\partial u_m}, \ldots, \frac{\partial x_n}{\partial u_m} \right\},$$

sont linéairement indépendants, au point correspondant M de la surface P. Ces vecteurs, du moment que leurs origines soient au point M, sont disposés dans le plan tangent à la surface P ; le vecteur r_j est tangent à la ligne sur laquelle la coordonnée u_j varie, les autres coordonnées u_i, $i \neq j$, étant constantes.

Construisons un vecteur tangent à *n'importe quelle* ligne L passant sur la surface P par le point M.

La ligne L peut être donnée par les équations paramétriques

$$\left.\begin{aligned} u_j &= u_j(t),\ j = 1, \ldots, m, \\ a \leqslant t \leqslant b;\ r &= r(u_1(t), \ldots, u_m(t)). \end{aligned}\right\} \tag{1}$$

Les fonctions $u_j(t)$ sont supposées dérivables. Alors, d'après le théorème sur la dérivée d'une fonction composée *1.33b*, on a

$$\frac{dr}{dt} = \sum_{j=1}^{m} \frac{\partial r}{\partial u_j} \frac{du_j}{dt} = \sum_{j=1}^{m} r_j u_j'(t). \tag{2}$$

L'égalité (2) peut être interprétée comme développement du vecteur $\frac{dr}{dt}$ suivant les vecteurs de base $r_1, \ldots, r_m$ au point M. D'une manière analogue, pour la différentielle dr du vecteur $r(t)$, on a

$$dr = \sum_{j=1}^{m} r_j u_j'(t)\, dt = \sum_{j=1}^{m} r_j\, du_j. \tag{3}$$

5.13. Comme nous avons vu dans *U16.19*, la longueur de la courbe L entre les points correspondant à des valeurs $t = a$ et $t = \tau$ du paramètre est donnée par l'intégrale

$$s = \int_a^{\tau} |r'(t)|\, dt,$$

par conséquent,

$$ds(t) = |r'(t)|\, dt = |dr(t)|.$$

En particulier, nous avons sur la surface P :

$$ds^2 = |dr|^2 = \Big(\sum_{j=1}^{m} r_j\, du_j,\ \sum_{k=1}^{m} r_k\, du_k\Big) = \sum_{j=1}^{m}\sum_{k=1}^{m} g_{jk}\, du_j\, du_k, \tag{1}$$

où $g_{jk} = (r_j, r_k)$. La forme quadratique dans le second membre de (1) s'appelle *première forme quadratique de la surface* P, on la désigne par $G = G(u; du)$. Si l'on connaît la première forme quadratique, i.e. les coefficients g_{jk} en tant que fonctions du point M de la surface P, ou bien, ce qui revient au même, des paramètres $u_1, \ldots, u_m$, alors on peut trouver la longueur de la ligne L entre deux points A et B correspondant aux valeurs a et b du paramètre t d'après la formule

$$s = \int_{t=a}^{b} ds = \int_a^b \sqrt{\sum_{j=1}^{m}\sum_{k=1}^{m} g_{jk}(u(t))\, u_j'(t)\, u_k'(t)}\, dt. \tag{2}$$

Si deux lignes L_1 et L_2 tracées sur la surface P et déterminées par les équations respectives $u_j = u_j^{(1)}(t)$, $u_j = u_j^{(2)}(t)$ $(j = 1, \ldots, m)$ se coupent au point M, alors on peut trouver l'angle qu'elles font (i.e. l'angle ω formé par les tangentes correspondantes) d'après la première forme quadratique:

$$\cos\omega = \frac{(dr^{(1)}, dr^{(2)})}{|dr^{(1)}|\,|dr^{(2)}|} = \frac{\sum_{j=1}^{m}\sum_{k=1}^{m} g_{jk}(M)\,du_j^{(1)}\,du_k^{(2)}}{\sqrt{\sum_{j,k=1}^{m} g_{jk}(M)\,du_j^{(1)}\,du_k^{(1)}}\ \sqrt{\sum_{j,k=1}^{m} g_{jk}(M)\,du_j^{(2)}\,du_k^{(2)}}}. \quad (3)$$

La première forme quadratique est, évidemment, définie positive (par sa construction même d'autant plus que ses mineurs principaux représentent les déterminants de Gram pour les vecteurs linéairement indépendants, donc sont positifs). Définissons, pour les vecteurs $r^{(i)} = \sum_{j=1}^{m} \xi_j^{(i)} r_j$ $(i = 1, 2)$ dans le plan tangent à la surface P, un produit scalaire $(r^{(1)}, r^{(2)})_g$ par la formule

$$(r^{(1)}, r^{(2)})_g = \sum_{j=1}^{m}\sum_{k=1}^{m} \xi_j^{(1)}\xi_k^{(2)} g_{jk}.$$

Nous affirmons que ce produit scalaire coïncide avec le produit scalaire ordinaire $(r^{(1)}, r^{(2)})$ des mêmes vecteurs dans l'espace R_n. En effet,

$$(r^{(1)}, r^{(2)}) = \Big(\sum_{j=1}^{m} \xi_j^{(1)} r_j,\ \sum_{k=1}^{m} \xi_k^{(2)} r_k\Big) = \sum_{j=1}^{m}\sum_{k=1}^{m} \xi_j^{(1)}\xi_k^{(2)}\cdot(r_j\ r_k) =$$

$$= \sum_{j=1}^{m}\sum_{k=1}^{m} \xi_j^{(1)}\xi_k^{(2)} g_{jk} = (r^{(1)}, r^{(2)})_g.$$

Ainsi, *la première forme quadratique restitue sur le plan tangent, en coordonnées* $\xi_1, \ldots, \xi_m$, *la métrique euclidienne initiale de l'espace* R_n.

5.14. Soient $P^{(1)}$ et $P^{(2)}$ deux surfaces, et supposons qu'il y ait une correspondance biunivoque entre leurs points, telle que la longueur de chaque ligne sur la surface $P^{(1)}$ soit égale à la longueur de son image sur la surface $P^{(2)}$. Une telle correspondance biunivoque entre deux surfaces $P^{(1)}$ et $P^{(2)}$ est appelée *isométrie*.

Théorème. *Pour que deux surfaces* $P^{(1)}$ *et* $P^{(2)}$ *soient isométriques, il faut et il suffit qu'il existe leurs représentations paramétriques respectives*

$$r^{(1)} = r^{(1)}(u_1, \ldots, u_m), \quad r^{(2)} = r^{(2)}(u_1, \ldots, u_m)$$

dans un même domaine de variation U des paramètres $u_1, \ldots, u_m$, *telles que les fonctions* $g_{jk}^{(1)} = (r_j^{(1)}, r_k^{(1)})$ *coïncident respectivement avec les fonctions* $g_{jk}^{(2)} = (r_j^{(2)}, r_k^{(2)})$.

Démonstration. Si les représentations paramétriques mentionnées des surfaces $P^{(1)}$ et $P^{(2)}$ existent, les longueurs de deux courbes correspondant l'une à l'autre sur les surfaces $P^{(1)}$ et $P^{(2)}$ sont données par les mêmes intégrales du type *5.13* (2), donc, coïncident, et les surfaces $P^{(1)}$ et $P^{(2)}$ sont isométriques. Inversement, si les surfaces $P^{(1)}$ et $P^{(2)}$ sont isométriques, alors, dans leurs représentations paramétriques correspondant l'une à l'autre, pour tout τ, a lieu l'égalité

$$\int_a^\tau \sqrt{\sum_{j,k} g_{jk}^{(1)}(u)\, u_j'(t)\, u_k'(t)}\, dt = \int_a^\tau \sqrt{\sum_{j,k} g_{jk}^{(2)}(u)\, u_j'(t)\, u_k'(t)}\, dt.$$

En dérivant par rapport à τ, nous trouvons:

$$\sum_{j,k} g_{jk}^{(1)}(u)\, u_j'(\tau)\, u_k'(\tau) = \sum_{j,k} g_{jk}^{(2)}(u)\, u_j'(\tau)\, u_k'(\tau)$$

quels que soient u_j', car il est possible de tracer sur la surface $P^{(1)}$ ou $P^{(2)}$ une courbe passant par le point M dans n'importe quelle direction. L'identité des formes implique l'égalité des coefficients respectifs de ces formes, ce qu'il fallait.

Les propriétés d'une surface qui sont conservées en passant à une surface isométrique se rapportent à la *géométrie intrinsèque* de la surface; analytiquement, pour qu'une propriété soit intrinsèque, il faut et il suffit qu'elle puisse être exprimée par les fonctions $g_{jk}(u)$. Les propriétés d'une surface qui *changent* en passant à une surface isométrique (par exemple, les courbures des lignes sur la surface) sont dites *extérieures*.

En cas de ligne, i.e. pour $m = 1$, la classification ci-dessus perd le sens: n'importe quelles deux lignes dérivables sont isométriques. Pour $m = 2$, elle a un sens réel; un plan et un cylindre sont isométriques (localement), tandis qu'un plan et une sphère ne le sont pas, ce qu'on verra dans la suite (*5.33*). Pour $m > 2$, l'isométrie de deux surfaces représente déjà un phénomène relativement rare (*5.35*).

5.15. Surfaces bidimensionnelles dans R_3. Désignons les coordonnées dans R_3, comme d'habitude, par x, y, z, et les paramètres u_1, u_2, qui déterminent la position du point M sur la surface P, par u, v, de sorte que le système *5.11* (1) soit de la forme

$$x = x(u, v), \quad y = y(u, v), \quad z = z(u, v)$$

ou, avec les vecteurs,

$$r = r(u, v) = \{x(u, v), \quad y(u, v), \quad z(u, v)\}.$$

Les dérivées du rayon vecteur de la surface $r(u, v)$ par rapport aux paramètres u et v sont désignées par r_u et r_v respectivement, et l'on a

$$r_u = \left\{ \frac{\partial x(u, v)}{\partial u}, \frac{\partial y(u, v)}{\partial u}, \frac{\partial z(u, v)}{\partial u} \right\},$$

$$r_v = \left\{ \frac{\partial x(u, v)}{\partial v}, \frac{\partial y(u, v)}{\partial v}, \frac{\partial z(u, v)}{\partial v} \right\}.$$

Les coefficients de la première forme quadratique, d'après Gauss, sont désignés comme suit

$$E = (r_u, r_u), \quad F = (r_u, r_v), \quad G = (r_v, r_v),$$

après quoi la première forme quadratique s'exprime:

$$ds^2 = E\,du^2 + 2F\,du\,dv + G\,dv^2. \tag{1}$$

Exemples. a. Soit P un plan qui coïncide avec le plan des coordonnées x, y. En choisissant pour paramètres du point M les coordonnées x et y

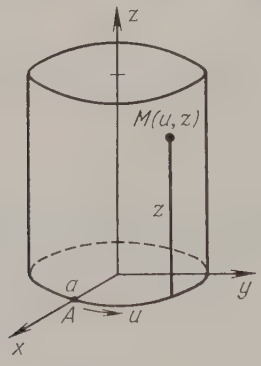


Fig. 5.1-1

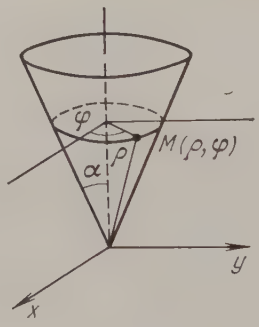


Fig. 5.1-2

nous trouvons:

$$r = \{x, y, 0\}, \quad r_z = \{1, 0, 0\}, \quad r_y = \{0, 1, 0\}, \quad ds^2 = dx^2 + dy^2. \tag{2}$$

b. Dans le même plan, on peut choisir pour paramètres les coordonnées polaires φ et ρ. Alors

$$r = \{\rho \cos \varphi, \rho \sin \varphi, 0\},$$
$$r_\varphi = \{-\rho \sin \varphi, \rho \cos \varphi, 0\},$$
$$r_\rho = \{\cos \varphi, \sin \varphi, 0\},$$
$$E = (r_\varphi, r_\varphi) = \rho^2, \quad F = 0, \quad G = 1,$$
$$ds^2 = \rho^2\,d\varphi^2 + d\rho^2. \tag{3}$$

c. Sur un cylindre C de rayon a (fig. 5.1-1), choisissons pour paramètres du point M la longueur d'arc u de la circonférence de la base inférieure comptée

du point A dans le sens indiqué et la coordonnée z. Alors

$$r(u, z) = \left\{ a\cos\frac{u}{a},\ a\sin\frac{u}{a},\ z \right\}$$

$$r_u = \left\{ -\sin\frac{u}{a},\ \cos\frac{u}{a},\ 0 \right\},$$

$$r_z = \{0, 0, 1\},$$

$$E = 1, \quad F = 0, \quad G = 1,$$

$$ds^2 = du^2 + dz^2, \tag{4}$$

Les coefficients de la première forme quadratique sont les mêmes que pour le plan en coordonnées rectangulaires ; cela confirme analytiquement qu'un plan est isométrique à un cylindre.

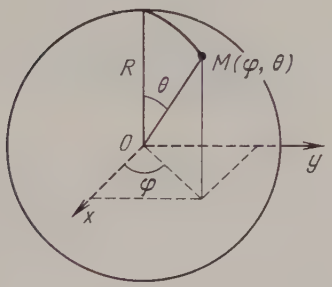


Fig. 5.1-3

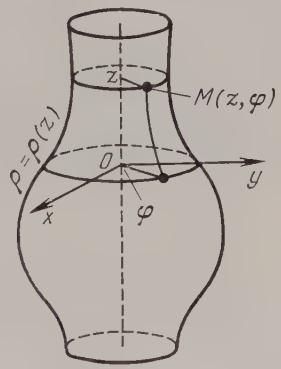


Fig. 5.1-4

d. Sur un cône K, choisissons pour paramètres du point M son angle polaire φ et sa distance ρ du sommet le long de la génératrice (fig. 5.1-2). Si l'angle formé par l'axe du cône et la génératrice vaut α, on a

$$r(\varphi, \rho) = \{\rho\sin\alpha\cos\varphi,\ \rho\sin\alpha\sin\varphi,\ \rho\cos\alpha\},$$

$$r_\varphi = \{-\rho\sin\alpha\sin\varphi,\ \rho\sin\alpha\cos\varphi,\ 0\}$$

$$r_\rho = \{\sin\alpha\cos\varphi,\ \sin\alpha\sin\varphi,\ \cos\alpha\},$$

$$E = (r_\varphi, r_\varphi) = \rho^2\sin^2\alpha, \quad F = (r_\varphi, r_\rho) = 0,$$

$$G = (r_\rho, r_\rho) = 1,$$

$$ds^2 = \rho^2\sin^2\alpha\, d\varphi^2 + d\rho^2. \tag{5}$$

En remplaçant ici la coordonnée φ par la nouvelle coordonnée $\psi = \dfrac{\varphi}{\sin\alpha}$, la forme ds^2 s'écrira $\rho^2\, d\psi^2 + d\rho^2$, ce qui coïncide avec la première forme quadratique du plan en coordonnées polaires. Ainsi, un cône et un plan sont eux aussi isométriques ; cela est bien connu dans la géométrie élémentaire.

e. Sur une sphère S de rayon R (fig. 5.1-3), choisissons les coordonnées sphériques ordinaires φ et θ. Dans ce cas, nous avons :

$$r(\varphi, \theta) = \{R \sin\theta \cos\varphi,\ R \sin\theta \sin\varphi,\ R \cos\theta\},$$
$$r_\varphi = \{-R \sin\theta \sin\varphi,\ R \sin\theta \cos\varphi,\ 0\},$$
$$r_\theta = \{R \cos\theta \cos\varphi,\ R \cos\theta \sin\varphi,\ -R \sin\theta\},$$
$$E = (r_\varphi, r_\varphi) = R^2 \sin^2\theta,\ F = (r_\varphi, r_\theta) = 0,$$
$$G = (r_\theta, r_\theta) = R^2,$$
$$ds^2 = R^2 \sin^2\theta\, d\varphi^2 + R^2\, d\theta^2 \qquad (6)$$

f. Considérons *une surface P de révolution* autour de l'axe des z, définie par l'équation $\rho = \rho(z)$ où ρ est la distance du point M à l'axe des z. Choisis-

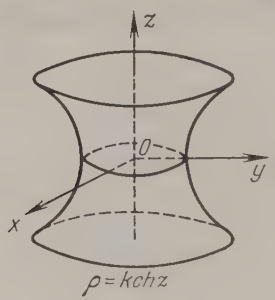


Fig. 5.1-5

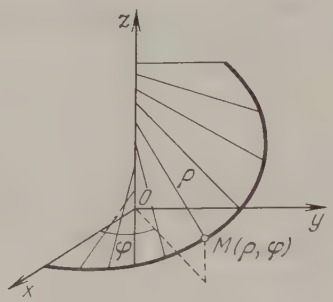


Fig. 5.1-6

sons pour paramètres du point M la quantité z et l'angle polaire φ (fig. 5.1-4). Nous avons

$$r(z, \varphi) = \{\rho(z) \cos\varphi,\ \rho(z) \sin\varphi,\ z\},$$
$$r_z = \{\rho_z \cos\varphi,\ \rho_z \sin\varphi,\ 1\},$$
$$r_\varphi = \{-\rho \sin\varphi,\ \rho \cos\varphi,\ 0\},$$
$$E = (r_z, r_z) = \rho_z^2 + 1,\quad F = (r_z, r_\varphi) = 0,$$
$$G = \{r_\varphi, r_\varphi\} = \rho^2,$$
$$ds^2 = (\rho_z^2 + 1)\, dz^2 + \rho^2\, d\varphi^2. \qquad (7)$$

g. Un cas spécial intéressant de surface de révolution est présenté par une *caténoïde*, i.e. la surface de révolution autour de l'axe des z d'une chaînette:

$$\rho(z) = k \operatorname{ch} \frac{z}{k}.$$

(voir fig. 5.1-5). En cas de caténoïde, la formule (7) prend la forme

$$ds^2 = \operatorname{ch}^2 \frac{z}{k} (dz^2 + k^2\, d\varphi^2). \qquad (8)$$

h. Un *hélicoïde* ou *surface hélicoïdale* est décrit par une demi-droite β parallèle au plan x, y et issue de l'axe des z, qui tourne autour de cet axe et, en même temps, monte à la vitesse proportionnelle en s'éloignant du plan (fig. 5.1-6). Ainsi, en choisissant pour les paramètres du point M de l'hélicoïde sa distance ρ de l'origine de β et l'angle polaire φ (i.e. l'angle de rotation de β compté de sa

position initiale), nous avons

$$r(\rho, \varphi) = \{\rho\cos\varphi, \rho\sin\varphi, k\varphi\},$$
$$r_\rho = \{\cos\varphi, \sin\varphi, 0\}, \quad r_\varphi = \{-\rho\sin\varphi, \rho\cos\varphi, k\},$$
$$E = (r_\rho, r_\rho) = 1, \quad F = (r_\rho, r_\varphi) = 0,$$
$$G = (r_\varphi, r_\varphi) = \rho^2 + k^2,$$
$$ds^2 = d\rho^2 + (\rho^2 + k^2)\, d\varphi^2.$$

Si l'on pose $\rho = k\,\mathrm{sh}\,\frac{v}{k}$, alors, en coordonnées v, φ, l'expression de ds^2 aura la forme

$$ds^2 = \mathrm{ch}^2\frac{v}{k}\, dv^2 + k^2\,\mathrm{ch}^2\frac{v}{k}\, d\varphi^2 = \mathrm{ch}^2\frac{v}{k}\,(dv^2 + k^2\, d\varphi^2). \tag{9}$$

Le second membre de (9) coïncide (aux désignations des paramètres près) avec la première forme quadratique de la caténoïde (8). Nous aboutissons à une conclusion un peu inattendue: *une caténoïde est isométrique à un hélicoïde*. La correspondance est telle que les lignes de φ constant de l'hélicoïde (i.e. les positions successives de la demi-droite en rotation) s'appliquent sur les lignes de φ constant de la caténoïde (i.e. sur ses génératrices ou méridiennes); les lignes $\rho = C$ de l'hélicoïde (lignes hélicoïdales) s'appliquent sur les lignes $z = C$ de la caténoïde (ses parallèles). A la ligne de striction de la caténoïde $z = 0$ il correspond l'axe des z sur l'hélicoïde. Soulignons qu'il s'agit de l'isométrie des parties de ces surfaces; l'isométrie en entier n'a pas lieu.

5.16. Dans les considérations précédentes, la forme métrique de la surface provient, en fin des comptes, de la métrique de l'espace euclidien contenant la surface. Or, ce n'est pas nécessaire. On peut définir autrement une métrique (i.e. une forme $ds^2 = \sum_{i,j} g_{ij}\, du_i\, du_j$) pour une surface m-dimensionnelle $r = r(u_1, \ldots, u_m)$, sans la lier à la métrique de l'espace contenant la surface. Il faut seulement que la forme $\sum_{i,j} g_{ij}\, du_i\, du_j$ soit symétrique et définie positive. Un exemple important d'une telle métrique sur l'hyperboloïde $x^2 + y^2 - z^2 = -\rho^2$ de l'espace R_3 sera donné dans *5.54*.

§ 5. 2. Deuxième forme quadratique

5.21. Courbure d'une ligne sur une surface. Ici nous nous bornons au cas $n = m + 1$ (surface n-dimensionnelle dans l'espace euclidien $(n+1)$-dimensionnel R_{n+1}), la métrique étant empruntée à R_{n+1}. Dans ces conditions on peut introduire *le vecteur normal* à la surface à un point donné; dans les mêmes suppositions sur la dérivabilité du rayon vecteur $r(u_1, \ldots, u_n)$ et sur le rang de la matrice jacobienne $\frac{\partial(x_1, \ldots, x_{n+1})}{\partial(u_1, \ldots, u_n)}$, il est défini d'une façon univoque, à un facteur numérique près. Analytiquement, le vecteur N normal à la surface P en un point donné M peut être déterminé comme produit vectoriel de n vecteurs de base $r_1 = \frac{\partial r}{\partial u_1}, \ldots, r_n = \frac{\partial r}{\partial u_n}$ du

plan tangent (*3.62c*) :

$$N = [r_1, \ldots, r_n] = \begin{Vmatrix} e_1 & \ldots & e_{n+1} \\ \frac{\partial x_1}{\partial u_1} & \ldots & \frac{\partial x_{n+1}}{\partial u_1} \\ \ldots & \ldots & \ldots \\ \frac{\partial x_1}{\partial u_n} & \ldots & \frac{\partial x_{n+1}}{\partial u_n} \end{Vmatrix}. \qquad (1)$$

On peut normaliser le vecteur N en le divisant par sa longueur, après quoi on aura un vecteur normal unitaire $m = m(u)$ défini déjà au facteur ± 1 près. Dans les limites d'un petit voisinage du point M, nous fixons le vecteur $m(u)$ en tout point de façon à avoir une fonction $m(u)$ continue.

5.22. A partir de ce moment nous supposons que le rayon vecteur $r(u)$ de la surface soit deux fois continûment dérivable (par rapport aux paramètres $u_1, \ldots, u_n$). Considérons sur la surface P une courbe L donnée par les équations

$$u_1 = u_1(t), \ldots, u_n = u_n(t), \quad a \leqslant t \leqslant b,$$
$$r = r(u_1(t), \ldots, u_n(t)),$$

où les fonctions $u_j(t)$ sont elles aussi deux fois continûment dérivables. C'est la courbure de cette ligne au point $A \in P$ répondant à une valeur $t = \alpha$ qui nous intéresse à présent. Pour la calculer, passons sur la courbe L au paramètre naturel s qui est la longueur d'arc comptée à partir d'un point fixe A_0. Rappelons (*U16.22*) que, pour une courbe paramétrisée avec s, on a la formule

$$\frac{dr(A)}{ds} = \tau(A),$$

et si $\frac{d^2r(A)}{ds^2} \neq 0$, alors

$$\frac{d^2r(A)}{ds^2} = \frac{d\tau(A)}{ds} = k(A)\,\nu(A).$$

Ici τ est le vecteur unitaire tangent à la courbe L, ν le vecteur unitaire orthogonal à τ et se trouvant dans le plan osculateur, k la courbure de la ligue L. La droite déterminée dans le plan osculateur par le vecteur ρ s'appelle *normale principale* à la courbe L.

Vu que la courbe L se trouve sur la surface P, nous avons

$$\frac{d^2r}{ds^2} = \frac{d}{ds}\left(\frac{dr}{ds}\right) = \frac{d}{ds}\left(\sum_{j=1}^{n} r_j \frac{du_j}{ds}\right) = \sum_{j,\,k=1}^{n} r_{jk} \frac{du_k}{ds}\frac{du_j}{ds} + \sum_{j=1}^{n} r_j \frac{d^2u_j}{ds^2}, \qquad (1)$$

avec la désignation $r_{jk} = \frac{\partial^2 r}{\partial u_j \partial u_k}$. En multipliant scalairement (1) par le vecteur m, on a

$$\left(\frac{d^2r}{ds^2}, m\right) = \sum_{j,k=1}^{n} (r_{jk}, m) \frac{du_j}{ds} \frac{du_k}{ds},$$

puisque $(r_j, m) = 0$. Soit $b_{jk} = (r_{jk}, m)$. La forme

$$B \equiv B(u, du) = \sum_{j,k=1}^{n} b_{jk}\, du_j\, du_k \tag{2}$$

s'appelle *deuxième forme quadratique de la surface P*. En se rappelant que $ds^2 = \sum_{j,k=1}^{n} g_{jk}\, du_j\, du_k = G(u, du)$ est la première forme quadratique de la surface P, on obtient définitivement:

$$\left(\frac{d^2r}{ds^2}, m\right) = \frac{B(u, du)}{G(u, du)} = \frac{\sum_{j,k=1}^{n} b_{jk}(u)\, du_j\, du_k}{\sum_{j,k=1}^{n} g_{jk}(u)\, du_j\, du_k}$$

(formule de Meusnier). La formule de Meusnier conduit aux conclusions suivantes:

a. Si, pour une direction $\{du_j\}$, la deuxième forme quadratique $B(u, du)$ est non nulle, alors, pour toute courbe sur la surface P qui est tangente au point A au vecteur du_j, a lieu l'inégalité $\left(\frac{d^2r(A)}{ds^2}, m\right) \neq 0$. Dans ce cas, $\frac{d^2r}{dS^2} = k\nu \neq 0$, donc, la courbe L a, au point A, une courbure non nulle, et la formule de Meusnier prend la forme

$$(k\nu, m) = k \cos(\widehat{\nu, m}) = \frac{B(u, du)}{G(u, du)}. \tag{3}$$

Ainsi, *pour les directions* $\{du_j\}$ *où* $B(u, du) \neq 0$, *la courbure de la ligne correspondante sur la surface est déterminée par les quantités*: u_j $(j = 1, \ldots, n)$ (*point de la surface*), du_j $(j = 1, \ldots, n)$ (*direction de la droite tangente*) *et* ν (*direction de la normale principale*). *En particulier, toutes les lignes sur la surface P qui passent par un point donné A dans une direction donnée* $\{du_j\}$, *qui ont un même plan osculateur, et pour lesquelles* $B(u, du) \neq 0$, *ont une même courbure.*

b. *Parmi toutes les lignes sur la surface P qui passent par un point donné A, dans une direction donnée* (u_j *et* du_j *sont fixes*), *avec* $B(u, du) \neq 0$, *la ligne dont le plan osculateur contient le vecteur m, de sorte que* $|\cos(\widehat{\nu, m})| = 1$, *a la moindre courbure.* En tant qu'une telle ligne, on peut choisir la section normale de la surface, i.e.

la ligne déterminée par l'intersection de la surface P et du plan (bidimensionnel) passant par les vecteurs m et $dr = \sum_j r_j\, du_j$. La courbure d'une telle ligne s'obtient d'après la formule

$$k_N = \frac{\| B(u, du) \|}{G(u, du)}.$$

La définition de la courbure d'une section normale est plus répandue sous la forme

$$k_N = \frac{B(u, du)}{G(u, du)}. \tag{4}$$

Cette expression s'appelle *courbure normale* de la surface P au point A suivant la direction $\{du_j\}$; elle est de signe + si les vecteurs m et ν sont de même sens, et de signe — dans le cas contraire ; autrement dit, la courbure normale est positive si la ligne est courbée dans le sens du vecteur normal, et négative dans le cas contraire.

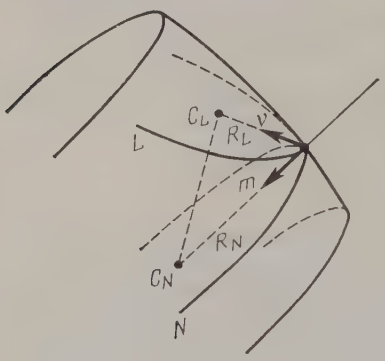


Fig. 5.2-1

Il résulte de (3) et (4) que la courbure de n'importe quelle ligne est liée à la courbure de la section normale correspondante par la relation

$$k \cos(\widehat{\nu, m}) = k_N. \tag{5}$$

Les rayons de courbure R_L et R_N de la ligne L et de la section normale correspondante sont reliés par l'égalité

$$R_N \cos(\widehat{\nu, m}) = R_L. \tag{6}$$

Il en résulte la propriété géométrique suivante (fig. 5.2-1) :

c. (Théorème de Meusnier) *Etant donnée une courbe L, son centre de courbure C_L est la projection sur le plan osculateur du centre de courbure C_N de la section normale correspondante N.*

Cette dernière propriété, qui découle directement de la formule (6), peut être appliquée à la détermination du centre de la courbure normale, si l'on connaît le centre de courbure d'une ligne passant par un point donné de la courbe donnée dans la même direction.

d. Toutes les propositions formulées sont valables pour les directions $\{du_j\}$ telles que $B(u, du) \neq 0$. Considérons une direction $\{du_j\}$ pour laquelle $B(u, du) = 0$. Appelons-la *direction asymptotique sur le plan tangent* Π (pour la raison qu'on verra dans *5.24f*). Pour une courbe $L = \{r = r(u),\ u = u(s)\}$ qui passe sur la surface

P par le point A dans une direction asymptotique, a lieu l'égalité

$$\left(\frac{d^2r(A)}{ds^2},\ m\right)=0.$$

Cela signifie : *ou bien la courbure de la courbe L au point A est nulle, ou bien, si* $\frac{d^2r(A)}{ds^2}\neq 0$, *le vecteur* $\frac{d^2r}{ds^2}=k\nu\ (A)$ *se trouve dans le plan* Π. Comme le vecteur $\tau\ (A)$ appartient lui aussi au plan Π, le plan osculateur de la courbe L au point A se trouve dans le plan Π tangent à la surface P. En particulier, *la courbure de toute ligne plane qui va dans une direction asymptotique et qui n'appartient pas au plan tangent* Π *est nulle* ; ainsi, une section normale dans une direction asymptotique est, nécessairement, de courbure nulle.

5.23. On peut attribuer à la deuxième forme quadratique un sens géométrique direct. En développant l'accroissement du rayon vecteur $r\ (u)$, pour le passage d'un point u à un point proche $u+du$ d'après la formule de Taylor, nous avons (aux infiniment petits du deuxième ordre près) :

$$\Delta r\equiv r\ (u+\Delta u)-r\ (u)=\sum_{j=1}^{n} r_j\,du_j+\frac{1}{2}\sum_{j,\,k=1}^{n} r_{jk}\,du_j\,du_k+\ldots,$$

d'où

$$(\Delta r,\ m)=\frac{1}{2}\sum_{j,\,k=1}^{n}(r_{jk},\ m)\,du_j\,du_k+\ldots=\frac{1}{2}B\ (u,\ du)+\ldots$$

Ainsi, la forme $B\ (u,\ du)$ avec le coefficient 1/2 coïncide avec la partie quadratique principale de la projection du vecteur Δr sur le vecteur m, i.e. avec l'écart de la surface P de son plan tangent. On choisit pour sens positif celui du vecteur m. Lorsqu'on remplace le vecteur m par son opposé, toute la forme $B\ (u,\ du)$ change de signe, les courbures de toutes les lignes sur la surface au point A en font autant. On peut dire que la forme $B\ (u,\ du)$ admet l'interprétation géométrique mentionnée au signe près.

Par exemple, si $n=2$ et la forme $B\ (u,\ du)$ en un point donné A est de signe défini, alors la surface P est disposée (dans un voisinage du point A) d'un côté du plan tangent ; et si la forme $B\ (u,\ du)$, au point A, n'est pas de signe défini (étant non dégénérée), alors la surface P (dans la proximité arbitraire du point A) a des parties se trouvant de chaque côté du plan tangent. Dans le premier cas, le point A est dit point *elliptique* de la surface, dans le deuxième, point *hyperbolique*, puisque, dans le premier cas, la surface a la forme, au voisinage du point A, d'un paraboloïde elliptique, et dans le deuxième, d'un paraboloïde hyperbolique (aux infiniment petits du deuxième ordre près).

Et si, au point A, la forme $B(u, du)$ se trouve dégénérée, cela signifie qu'au voisinage du point A la surface a la forme d'un cylindre parabolique. Un tel point de la surface est dit *parabolique*. Si la forme $B(u, du)$ est identiquement nulle au point A, la surface coïncide, à des infiniment petits du deuxième ordre près, avec son plan

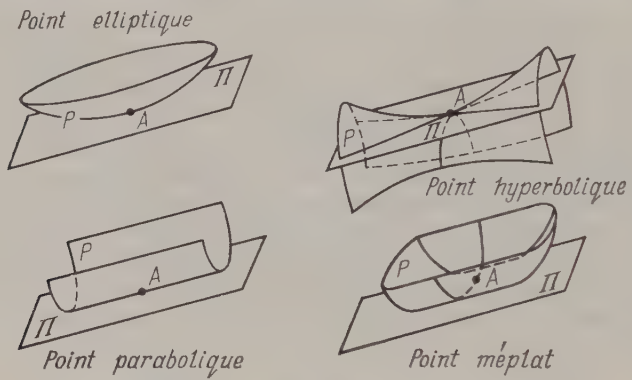


Fig. 5.2-2

tangent en s'en écartant d'un infiniment petit d'un ordre supérieur à 2. Un tel point est appelé point *méplat*.

Tous les types de points sont illustrés fig. 5.2-2.

5.24. Dépendance de la courbure des sections normales de la direction sur le plan tangent.

a. Nous essayons de simplifier la formule de la courbure d'une section normale

$$k_N = \frac{B(u, du)}{G(u, du)} \tag{1}$$

en passant à un nouveau système de coordonnées sur le plan tangent (n-dimensionnel) à un point donné A de la surface P. Trouvons dans ce plan une base orthonormée $g_1, \ldots, g_n$ dans laquelle la forme B prenne sa forme canonique

$$B(u, du) = \sum_{j=1}^{n} \lambda_j \xi_j^2,$$

où $\xi_1, \ldots, \xi_n$ sont les nouvelles coordonnées du vecteur $dr = \sum_{j=1}^{n} r_j\, du_j$. Comme la forme $G(u, du)$ représente tout simplement le carré de la longueur du vecteur dr, dans la nouvelle base elle sera :

$$G(u, du) = \sum_{j=1}^{n} \xi_j^2.$$

La formule (1), dans la nouvelle base, aura la forme

$$k_N = \frac{\sum_{j=1}^{n} \lambda_j \xi_j^2}{\sum_{j=1}^{n} \xi_j^2} = \sum_{j=1}^{n} \lambda_j \cos^2 \varphi_j, \qquad (2)$$

où φ_j est l'angle que le vecteur dr fait avec le vecteur de base g_j. Les directions des vecteurs g_j sont appelées *directions principales* sur le plan tangent Π. Les nombres λ_j sont les courbures des sections normales correspondantes; elles s'appellent *courbures principales* de la surface Π au point A. *La formule* (2) porte le nom *d'Euler.*

b. Si, au point A, tous les nombres λ_j sont d'un même signe, la forme $B(u, du)$ est de signe défini; nous appellerons un tel point (comme pour $n = 2$) *elliptique.* En un point elliptique, toutes les sections normales sont courbées dans un même sens (par rapport au vecteur m). En un point A où, parmi les nombres λ_j, il y a des positifs et des négatifs (mais il n'y a pas de nuls), certaines sections normales sont courbées dans un sens, d'autres dans le sens opposé; un tel point sera appelé (comme pour $n = 2$) *hyperbolique.* Si l'un des nombres λ_j, au moins, est nul et l'un, au moins, ne l'est pas, le point A est dit *parabolique.* Enfin, si tous les λ_j sont nuls, le point A s'appelle *point méplat.*

Le nombre $H = \frac{1}{n} \sum_{j=1}^{n} \lambda_j$ s'appelle *courbure moyenne de la surface* P au point A. Le nombre $K = \prod_{j=1}^{n} \lambda_j$ est appelé *courbure totale de la surface* P au point A. Pour $n = 2$, en un point elliptique on a $K > 0$, en un point hyperbolique $K < 0$, en un point parabolique $K = 0$.

c. Pour un plan n-dimensionnel $P \subset R_{n+1}$, toutes ses sections normales à tout son point A représentent des droites; par conséquent, *la deuxième forme quadratique d'un plan est identiquement nulle.*

Inversement, *si la deuxième forme quadratique d'une surface n-dimensionnelle* $P \subset R_{n+1}$ *est identiquement nulle, alors* P *est un plan n-dimensionnel.* En effet, si $(r_{kj}, m) = 0$ pour $k, j = 1, \ldots, n$, alors en dérivant l'égalité $(r_j, m) = 0$ nous obtenons $(r_j, m_k) = -(r_{jk}, m) = 0$, d'où $m_k = 0$ pour tous les $k = 1, \ldots, n$; par conséquent, le vecteur m ne varie pas sur la surface P. Soient r_0 le vecteur menant à un point fixe de la surface P, et r le vecteur menant à un point arbitraire de cette surface. Alors $(r - r_0, m)_j = (r_j, m) = 0$ $(j = 1, \ldots, n)$, de sorte que la quantité $(r - r_0, m)$ ne varie pas sur la surface P; comme elle est nulle pour $r = r_0$, on a $(r - r_0, m) = 0$ pour toutes les valeurs de r. La dernière équation est celle du plan passant par l'extrémité du vecteur r_0 et orthogonal au vecteur m.

En particulier, pour un plan n-dimensionnel $P \subset R_{n+1}$, toutes les courbures principales sont nulles, de même que la courbure moyenne et la courbure totale.

d. Pour une partie de la sphère n-dimensionnelle dans R_{n+1}, de rayon R, de centre r_0 et de normale dirigée du centre vers la sphère, nous avons $r - r_0 = Rm$, $r_j = Rm_j$; il en résulte $(r_j, m) = R(m_j, m) = 0$, donc,

$$b_{jk} = (r_{jk}, m) = (r_j, m)_k - (r_j, m_k) = \\ = -(r_j, m_k) = -\frac{1}{R}(r_j, r_k) = -\frac{1}{R} g_{jk}.$$

Et si la normale à la sphère est dirigée vers le centre, on a $b_{jk} = \frac{1}{R} g_{jk}$. Ainsi, *dans tout système de coordonnées, sur une partie de la sphère de rayon R, les coefficients de la deuxième forme quadratique sont proportionnels* (de coefficient $\pm 1/R$) *aux coefficients respectifs de la première forme quadratique.* Il s'avère que c'est une propriété caractérisant la sphère. Précisément, si pour une surface $P \subset R_{n+1}$, on a la relation $b_{jk} = \pm \frac{1}{R} g_{jk}$ ($j, k = 1, \ldots, n$, R est une constante, le signe étant fixe), alors:

$$b_{jk} = (r_{jk}, m) = -(r_j, m_k) = \pm \frac{1}{R} g_{jk} = \\ = \pm \frac{1}{R}(r_j, r_k) = \pm \left(r_j, \frac{1}{R} r_k\right),$$

donc, $m_k = \mp \frac{1}{R} r_k$ ($k = 1, \ldots, n$). Ceci est équivalent au fait que le vecteur $m \pm \frac{1}{R} r$ ne varie pas sur la surface P. Désignons-le par $\pm \frac{1}{R} r_0$. Alors $| r - r_0 | = R | m | = R$, et nous voyons que la surface P représente effectivement une partie de la sphère de rayon R et de centre au point r_0.

e. Pour une partie de la sphère d'un rayon R, toutes les sections normales représentent des circonférences de rayon R centrées au centre de la sphère. Donc, toutes les courbures normales de la sphère coïncident l'une avec l'autre et sont égales à la courbure du grand cercle $1/R$. La même quantité $1/R$ est la courbure moyenne en tout point de la sphère. La courbure totale de la sphère, en tant que produit de n courbures principales, vaut $1/R^n$. Notons que les deux dernières quantités restent constantes sur toute la sphère.

f. Indicatrice de Dupin. Si, dans le plan tangent Π, sur toute demi-droite $\{du_j\}$ issue du point A et formant les angles φ_j avec les vecteurs $g_j(a)$, nous portons un segment $\rho = \sqrt{R_N}$ (où $R_N = 1/|k_N|$ est le rayon de courbure de la section normale cor-

respondante), alors nous avons une surface d'équation

$$\rho = \sqrt{R_N} = \frac{1}{\sqrt{|k_N|}} = \frac{1}{\sqrt{\left|\sum_{j=1}^{n} \lambda_j \cos^2 \varphi_j\right|}}$$

ou, ce qui est le même,

$$\left|\rho^2 \sum_{j=1}^{n} \lambda_j \cos^2 \varphi_j\right| = \left|\sum_{j=1}^{n} \lambda_j \xi_j^2\right| \equiv |B(u, du)| = 1,$$

$$B(u, du) = \pm 1.$$

Cette surface du deuxième ordre, ou un couple de telles surfaces, (*indicatrice de Dupin*) permet d'interpréter géométriquement la dépendance de la courbure des sections normales de la direction du vecteur tangent $\{du_j\}$ sur le plan Π. Lorsque $n = 2$, l'indicatrice de Dupin représente une ellipse pour un point elliptique, un couple d'hyperboles pour un point hyperbolique, un couple de droites parallèles pour un point parabolique qui n'est pas un point méplat.

Les directions asymptotiques dans le plan Π, i.e. les directions dans lesquelles la deuxième forme quadratique est nulle, correspondent aux points à l'infini de l'indicatrice de Dupin. Ce sont les directions des asymptotes de l'indicatrice, d'où la dénomination même de ces directions.

5.25.a. Montrons comment calculer les courbures principales et trouver les directions principales en coordonnées initiales $u_1, \ldots, u_n$. Ce problème peut être interprété comme problème de réduction simultanée de la forme $G(u, du)$ à la somme des carrés des coordonnées, et de la forme $B(u, du)$ à sa forme canonique. L'algèbre linéaire nous dit (*L10.32*) qu'il faut pour cela considérer le système d'équations

$$\left.\begin{aligned} (b_{11} - \mu g_{11})\, du_1 + \ldots + (b_{1n} - \mu g_{1n})\, du_n = 0, \\ (b_{n1} - \mu g_{n1})\, du_1 + \ldots + (b_{nn} - \mu g_{nn})\, du_n = 0. \end{aligned}\right\} \qquad (1)$$

Les racines de l'équation

$$\begin{vmatrix} b_{11} - \mu g_{11} & \ldots & b_{1n} - \mu g_{1n} \\ \ldots & \ldots & \ldots \\ b_{n1} - \mu g_{n1} & \ldots & b_{nn} - \mu g_{nn} \end{vmatrix} = 0 \qquad (2)$$

fournissent les coefficients canoniques de la forme $B(u, du)$, i.e. les nombres $\lambda_1, \ldots, \lambda_n$ dans *5.24* (2). Si l'on remplace μ par λ_j dans (1), le système obtenu aura une solution non nulle $\{du_1^{(j)}, \ldots, du_n^{(j)}\}$ qui représente le vecteur de base g_j correspondant (à un facteur près).

On sait que l'équation (2) possède n racines réelles (chaque racine ρ-uple étant comptée ρ fois), et que le système (1), pour chaque racine ρ-uple λ_j, admet ρ solutions linéairement indépendantes. En cas de n racines *différentes* $\lambda_1, \ldots, \lambda_n$, on a n directions principales définies d'une façon univoque; et si une racine quelconque, mettons λ_1, a une multiplicité ρ_1 supérieure à 1, alors on peut choisir pour système de directions principales une base quelconque de ρ_1 vecteurs orthogonaux dans le sous-espace correspondant ρ_1-dimensionnel.

b. Soit

$$a_0\mu^n + a_1\mu^{n-1} + \ldots + a_n = a_0(\mu - \lambda_1)\ldots(\mu - \lambda_n)$$

la factorisation du polynôme dans le premier membre de (2). Nous avons

$$(-1)^n a_n = \lambda_1 \ldots \lambda_n a_0 = K a_0,$$

où K est la courbure totale de la surface P au point A (*5.24b*).

Le terme libre a_n du polynôme (2) s'obtient en y posant $\mu = 0$; nous voyons que a_n coïncide avec le déterminant de la deuxième forme quadratique $\det B$. Le coefficient a_0 s'obtient en divisant le polynôme (2) par μ^n et puis passant à la limite pour $\mu \to \infty$. Si l'on divise par μ chaque colonne et que l'on passe à la limite, nous trouvons que a_0 coïncide avec le déterminant $\det G$ de la première forme quadratique précédé du signe $(-1)^n$. Par conséquent, on arrive à la formule suivante pour la courbure totale K:

$$K = (-1)^n \frac{a_n}{a_0} = \frac{\det B}{\det G}. \tag{3}$$

5.26. Cas $n = 2$; exemples. Pour $n = 2$, on adopte les désignations suivantes:

$$r_{11} = \frac{\partial^2 r}{\partial u^2} = r_{uu},\quad r_{12} = \frac{\partial^2 r}{\partial u\,\partial v} = r_{uv},\quad r_{22} = \frac{\partial^2 r}{\partial v^2} = r_{vv},$$

$$b_{11} = (r_{11}, m) = L,\quad b_{12} = (r_{12}, m) = M,\quad b_{22} = (r_{22}, m) = N,$$

de sorte que

$$B(u, du) = L\,du^2 + 2M\,du\,dv + N\,dv^2. \tag{1}$$

Le système *5.25* (1) prend la forme

$$\left.\begin{aligned}(L - \mu E)\,du + (M - \mu F)\,dv = 0,\\ (M - \mu F)\,du + (N - \mu G)\,dv = 0,\end{aligned}\right\} \tag{2}$$

et l'équation *5.25* (2) la forme

$$\begin{vmatrix} L - \mu E & M - \mu F \\ M - \mu F & N - \mu G \end{vmatrix} = 0. \tag{3}$$

Calculons la deuxième forme quadratique pour les surfaces des exemples *5.15*.

a. Pour le plan $z = 0$, dans tout système de coordonnées curvilignes u, v, on a : $(r_{uu}, m) = (r_{uv}, m) = (r_{vv}, m) = 0$, de sorte que $L = M = N = 0$, $B(u, du) \equiv 0$. La courbure de toute section normale est nulle (cf. *5.24c*).

b. Dans les exemples *5.16b-g* on étudiait des surfaces de révolution. Envisageons maintenant la surface de révolution générale (*5.15f*). En coordonnées z, φ, nous avons

$$r = \{\rho(z)\cos\varphi,\ \rho(z)\sin\varphi,\ z\},$$

$$r_z = \{\rho_z\cos\varphi,\ \rho_z\sin\varphi,\ 1\},\qquad r_\varphi = \{-\rho\sin\varphi,\ \rho\cos\varphi,\ 0\},$$

$$N = \begin{vmatrix} t & j & k \\ \rho_z\cos\varphi & \rho_z\sin\varphi & 1 \\ -\rho\sin\varphi & \rho\cos\varphi & 0 \end{vmatrix} = -\rho\{\cos\varphi,\ \sin\varphi,\ -\rho_z\},$$

$$m = \frac{N}{|N|} = \frac{\{-\cos\varphi,\ -\sin\varphi,\ \rho_z\}}{\sqrt{1+\rho_z^2}},$$

$$r_{zz} = \{\rho_{zz}\cos\varphi,\ \rho_{zz}\sin\varphi,\ 0\},\qquad r_{z\varphi} = \{-\rho_z\sin\varphi,\ \rho_z\cos\varphi,\ 0\},$$

$$r_{\varphi\varphi} = \{-\rho\cos\varphi,\ -\rho\sin\varphi,\ 0\},$$

$$L = (r_{zz}, m) = -\frac{\rho_{zz}}{\sqrt{1+\rho_z^2}},$$

$$M = (r_{z\varphi}, m) = 0,$$

$$N = (r_{\varphi\varphi}, m) = \frac{\rho}{\sqrt{1+\rho_z^2}},$$

$$B(u, du) = -\frac{\rho_{zz}}{\sqrt{1+\rho_z^2}}\,dz^2 + \frac{\rho}{\sqrt{1+\rho_z^2}}\,d\varphi^2.$$

Comme, en coordonnées z, φ, la première forme quadratique représente elle aussi une somme de carrés (*5.15f*) :

$$G(u, du) = (\rho_z^2 + 1)\,dz^2 + \rho^2\,d\varphi^2,$$

les directions des lignes des coordonnées z, φ sont des directions principales. La section normale menée dans la direction de la ligne $\varphi = \text{const}$ est, évidemment, une méridienne. La section normale menée dans la direction de la parallèle $z = \text{const}$ est, en général, non pas cette parallèle même, mais une autre courbe qui ne coïncide avec la parallèle que si le vecteur m, au point donné, soit orthogonal à l'axe des z. Le centre de courbure de la parallèle se trouve sur l'axe des z, celui de la section normale, d'après le théorème de Meusnier (*5.22c*), a pour projection sur le plan de la parallèle le centre de celle-ci, et se trouve donc lui-même sur l'axe de révolution (théorème de Monge) (voir fig. 5.2-3). L'equation (3) devient

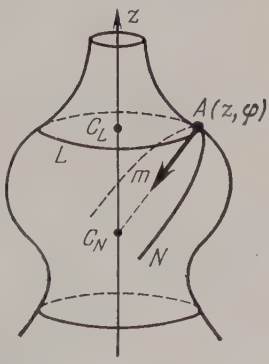


Fig. 5.2-3

$$\left(-\frac{\rho_{zz}}{\sqrt{1+\rho_z^2}} - \mu(1+\rho_z^2)\right)\left(\frac{\rho}{\sqrt{1+\rho_z^2}} - \mu\rho^2\right) = 0,$$

d'où

$$\mu_1 = -\frac{\rho_{zz}}{(1+\rho_z^2)^{3/2}},\qquad \mu_2 = \frac{1}{\rho(1+\rho_z^2)^{1/2}}. \tag{4}$$

Ainsi,

$$K=\mu_1\mu_2=-\frac{\rho_{zz}}{\rho\,(1+\rho_z^2)^2}, \tag{5}$$

$$2H=\mu_1+\mu_2=\frac{-\rho\rho_{zz}+(1+\rho_z^2)}{\rho\,(1+\rho_z^2)^{3/2}}. \tag{6}$$

La courbure totale K est positive aux points où $\rho_{zz}<0$, i.e. aux points où la courbe $\rho=\rho\,(z)$ est concave vers l'axe de révolution, et négative au points où $\rho_{zz}>0$, i.e. aux points où la courbe $\rho=\rho\,(z)$ est convexe vers cet axe.

c. Quelles sont les surfaces de révolution dont la courbure totale est nulle? De la formule (5) nous tirons $\rho_{zz}=0$, d'où $\rho=az+b$; c'est un cône pour $a\neq 0$, et un cylindre, pour $a=0$, $b>0$. Dans les deux cas, la génératrice de la surface présente une section normale de courbure nulle.

d. Quelles sont les surfaces de révolution dont la courbure moyenne est nulle? Ici nous avons à résoudre l'équation $\rho\rho_{zz}=1+\rho_z^2$. Avec la substitution $\rho_z=u\,(\rho)$, nous trouvons

$$\rho_{zz}=u_z=u_\rho\rho_z\equiv u_\rho u,\qquad \rho u_\rho u=1+u^2,$$

donc,

$$\frac{u\,du}{1+u^2}=\frac{d\rho}{\rho}.$$

En intégrant l'équation différentielle obtenue, on a $\ln\sqrt{1+u^2}=\ln\rho+\ln C$, d'où $\sqrt{1+u^2}=C\rho$ et ensuite $u\equiv\rho_z=\sqrt{C^2\rho^2-1}$. En posant $C\rho=\operatorname{ch} t$; $C\,d\rho=\operatorname{sh} t\,dt$, on a $dt=C\,dz$; en intégrant encore une fois, on trouve $t=C\,(z-z_0)$, d'où, définitivement,

$$\rho=\frac{1}{C}\operatorname{ch} C\,(z-z_0).$$

La surface cherchée s'avère une caténoïde (*5.15g*).

e. Voyons encore l'exemple *5.15h*, i.e. l'hélicoïde. En coordonnées ρ, φ qu'on y a introduites, nous avons

$$r_\rho=\{\cos\varphi,\ \sin\varphi,\ 0\},\qquad r_\varphi=\{-\rho\sin\varphi,\ \rho\cos\varphi,\ k\},$$

$$ds^2=d\rho^2+(\rho^2+k^2)\,d\varphi^2.$$

Il en résulte

$$N=\begin{vmatrix} i & j & k\\ \cos\varphi & \sin\varphi & 0\\ -\rho\sin\varphi & \rho\cos\varphi & k\end{vmatrix}=\{k\sin\varphi,\ -k\cos\varphi,\ \rho\},$$

$$m=\frac{N}{|N|}=\frac{\{k\sin\varphi,\ -k\cos\varphi,\ \rho\}}{\sqrt{k^2+\rho^2}}.$$

Ensuite,

$$r_{\rho\rho}=0,\quad r_{\rho\varphi}=\{-\sin\varphi,\ \cos\varphi,\ 0\},\quad r_{\varphi\varphi}=\{-\rho\cos\varphi,\ -\rho\sin\varphi,\ 0\},$$

$$L=(r_{\rho\rho},\,m)=0,\quad M=(r_{\rho\varphi},\,m)=-\frac{k}{\sqrt{k^2+\rho^2}},\quad N=(r_{\varphi\varphi},\,m)=0.$$

L'équation (3) en cas d'hélicoïde, prend la forme:

$$\begin{vmatrix} -\mu & -\dfrac{k}{\sqrt{k^2+\rho^2}} \\ -\dfrac{k}{\sqrt{k^2+\rho^2}} & -\mu(\rho^2+k^2) \end{vmatrix} = 0;$$

en la résolvant on a

$$\mu = \pm \frac{k}{\rho^2+k^2}.$$

On voit que la courbure moyenne d'un hélicoïde est également nulle. L'indicatrice de Dupin représente un couple d'hyperboles équilatères $\xi_2^2 - \xi_1^2 = \pm \frac{\rho^2+k^2}{k}$. La courbure de la section normale dans la direction de la demi-droite génératrice est nulle, de sorte que cette demi-droite représente une asymptote de l'indicatrice; par conséquent, les directions principales, à chaque point de l'hélicoïde (qui n'est pas sur l'axe des z), font les angles de 45° avec la direction de la demi-droite génératrice.

Remarquons que l'hélicoïde est la seule surface réglée (i.e. surface obtenue par le déplacement d'une droite) de courbure moyenne nulle (voir les exercices 8-10).

La courbure totale de l'hélicoïde est négative:

$$K = \mu_1\mu_2 = -\frac{k^2}{(\rho^2+k^2)^2}.$$

5.27. Interprétation géométrique de la courbure moyenne. D'après la formule d'Euler *5.24* (2), la courbure k_N de chaque section normale N de la surface P à un point A est donnée par la formule

$$k_N = \sum_{j=1}^{n} \lambda_j \cos^2 \varphi_j.$$

Trouvons la moyenne intégrale de toutes les valeurs k_N suivant toutes les directions des sections normales dans le plan tangent Π au point A. Pour la raison de symétrie, la valeur moyenne des quantités $\cos^2 \varphi_j$ est la même pour toutes les valeurs $j = 1, \ldots, n$. Comme $\sum_{j=1}^{n} \cos^2 \varphi_j \equiv 1$, la valeur moyenne de chaque fonction $\cos^2 \varphi_j$ vaut, évidemment, $1/n$. Définitivement, la moyenne intégrale des courbures est égale à $\frac{1}{n}\sum_{j=1}^{n} k_j$, i.e. à la courbure moyenne de la surface P au point A.

En particulier, la plus grande de toutes les courbures *principales* est en même temps la plus grande de *toutes* les courbures *normales*

en un point donné. Pour $n = 2$, lorsqu'il n'y a que deux courbures principales, l'une d'elles peut être caractérisée comme la plus grande, l'autre comme la plus petite de toutes les courbures normales. Pour $n > 2$, les courbures qui restent peuvent elles aussi être caractérisées en termes extrémaux; cf. *L10.24*.

5.28. Interprétation géométrique de la courbure totale. Pour une courbe plane, la courbure en un point donné A peut être définie comme vitesse de rotation du vecteur tangent, ou bien, ce qui est le même, comme vitesse de rotation du vecteur normal au point A. C'est la limite du rapport de la rotation du vecteur normal sur un petit arc Δs contenant le point A à cet arc même, lorsque ce dernier se contracte vers le point A. L'angle de rotation de la normale, à son tour, peut être défini comme longueur $\Delta\varphi$ de l'arc correspondant de la circonférence unitaire (fig. 5.2-4) centrée en un point arbitraire O.

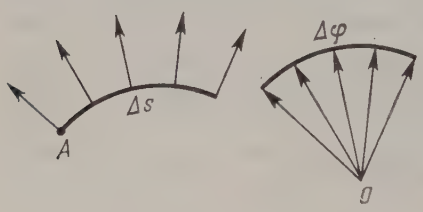


Fig. 5.2-4

Cette construction se généralise au cas de surface n-dimensionnelle dans l'espace $(n+1)$-dimensionnel de la façon suivante. Soit ΔS un petit domaine de la surface P contenant le point A et déterminé par un petit domaine $\Delta G \subset R_n$ de variation des paramètres. En tout point $M \in \Delta S$, considérons le vecteur unitaire de la normale $m(M)$; transportons les origines de tous ces vecteurs en un point fixe O et désignons par $\Delta\Omega$ le domaine sur la sphère unitaire déterminé par leurs extrémités. Ce domaine s'appelle *application sphérique du domaine* ΔS. Plus les directions des vecteurs $m(M)$ aux points $M \in \Delta S$ sont variées, i.e. plus la surface P est courbée au voisinage du point A, plus ce domaine est grand. Gauss définit en son temps (1828) la courbure $\varkappa(A)$ de la surface P (bidimensionnelle) au point A comme limite du rapport des aires $|\Delta\Omega| / |\Delta S|$ pour $\Delta S \to A$.

Essayons de relier, pour tout n, *la courbure de Gauss* $\varkappa = \lim_{\Delta S \to A} \frac{|\Delta\Omega|}{|\Delta S|}$ aux caractéristiques de la courbure de la surface P au point A que nous avons introduites. On sait (*3.62c*) que l'aire $|\Delta S|$ de la partie ΔS s'exprime d'après la formule

$$|\Delta S| = \int_{\Delta G} |r'(u)|\, du = \int_{\Delta G} |[r_1, \ldots, r_n]|\, du.$$

Comme le rayon vecteur du domaine $\Delta\Omega$ est le vecteur $m(u)$, avec le même domaine ΔG de variation des paramètres $u = (u_1, \ldots, u_n)$, l'aire $|\Delta\Omega|$ du domaine $\Delta\Omega$ s'exprime d'après la formule

analogue :

$$|\Delta\Omega| = \int_{\Delta G} |m'(u)|\,du = \int_{\Delta G} |[m_1, \ldots, m_n]|\,du,$$

où $m_j = \dfrac{\partial m}{\partial u_j}$.

Conformément à la définition de Gauss, nous avons

$$\varkappa(A) = \lim_{\Delta S \to A} \frac{|\Delta\Omega|}{|\Delta S|} = \lim_{\Delta S \to A} \frac{\int_{\Delta G} |[m_1, \ldots, m_n]|\,du}{\int_{\Delta G} |[r_1, \ldots, r_n]|\,du} = \frac{|[m_1(A), \ldots, m_n(A)]|}{|[r_1(A), \ldots, r_n(A)]|}.$$

Calculons $[m_1(A), \ldots, m_n(A)]$. Comme le vecteur m_j, en tant que dérivée d'un vecteur unitaire, est orthogonal à m, on peut noter

$$m_j = \sum_{k=1}^{n} b_j^k r_k.$$

En utilisant la propriété linéaire du déterminant qui exprime le produit vectoriel (*3.62c*), on trouve

$$\begin{aligned}[m_1, \ldots, m_n] &= \Big[\sum_{k_1} b_1^{k_1} r_{k_1}, \ldots, \sum_{k_n} b_n^{k_n} r_{k_n}\Big] = \\ &= \sum_{k_1, \ldots, k_n} b_1^{k_1} \ldots b_n^{k_n} [r_{k_1}, \ldots, r_{k_n}] = \\ &= \sum_{k_1 \ldots k_n} b_1^{k_1} \ldots b_n^{k_n} \varepsilon(k_1, \ldots, k_n) [r_1, \ldots, r_n] = \\ &= \det \| b_j^k \| [r_1, \ldots, r_n].\end{aligned}$$

Or, en dérivant l'identité $(r_i, m) = 0$, on a

$$0 = (r_i, m)_j = (r_i, m_j) + (r_{ij}, m) = \Big(r_i, \sum_{k=1}^{n} b_j^k r_k\Big) + b_{ij} = \sum_{k=1}^{n} b_j^k g_{ik} + b_{ij}.$$

Par conséquent,

$$b_{ij} = -\sum_{k=1}^{n} b_j^k g_{ik}, \quad i, j = 1, \ldots, n. \tag{1}$$

D'après le théorème sur le déterminant d'un produit de matrices, on a

$$\det \| b_{ij} \| = (-1)^n \det \| b_j^k \| \det \| g_{ij} \|,$$

d'où

$$\frac{[m_1(A), \ldots, m_n(A)]}{[r_1(A), \ldots, r_n(A)]} = \det \| b_j^k \| = (-1)^n \frac{\det \| b_{ij} \|}{\det \| g_{ij} \|} = (-1)^n K(A),$$

en vertu de *5.25b*. Définitivement,

$$\varkappa(A) = |K(A)|,$$

et la courbure de Gauss d'une surface P en un point A s'avère égale au module de la courbure totale de la surface P au même point A.

5.29. Lignes de courbure.

a. Si, en un point donné A d'une surface P, les racines $\lambda_1, \ldots, \lambda_n$ sont deux à deux distinctes, elles le restent dans un voisinage du point A (les racines de l'équation dépendant continûment de ses coefficients, *1.59c*); par conséquent, les directions principales restent définies d'une façon univoque au voisinage du point A. Ainsi, au voisinage du point A, sont définis n champs de directions orthogonales l'une à l'autre. Une racine $\mu = \lambda_j$ étant fixe, le système *5.25* (1) représente un système d'équations différentielles pour le j-ème champ. Les courbes intégrales de ce j-ème champ sont appelées *lignes de courbure* répondant à *la racine* λ_j. Les lignes de courbure forment n familles orthogonales l'une à l'autre. Il résulte de *2.64*, pour $n = 2$, qu'il est possible de passer, au voisinage du point A sur la surface P, à un nouveau système de coordonnées u, v dans lequel les lignes de courbure deviennent lignes de coordonnées. Dans ce système, la première et la deuxième formes quadratiques prennent leurs formes canoniques, et c'est à un voisinage tout entier du point A :

$$G = \tilde{g}_{11}(u, v)\, du^2 + \tilde{g}_{22}(u, v)\, dv^2,$$
$$B = \tilde{b}_{11}(u, v)\, du^2 + \tilde{b}_{22}(u, v)\, dv^2.$$

Pour $n > 2$, une réduction simultanée des formes G et B à la forme canonique dans tout un voisinage du point A n'est en général pas possible.

b. Exemple. Pour une surface de révolution, comme nous l'avons vu dans *5.26b*, les directions principales sont celles d'une méridienne et d'une parallèle; par conséquent, le réseau des méridiennes et parallèles est le réseau des lignes de courbure. Aux sommets de la surface, i.e. à ses points d'intersection avec l'axe de révolution, ce réseau a des singularités, c'est-à-dire cesse d'être réseau des lignes de coordonnées; si la surface reste dérivable à de tels points (comme, par exemple, l'ellipsoïde de révolution) les deux courbures principales s'y confondent.

§ 5. 3. Relations entre la première et la deuxième formes quadratiques

5.31. Formules de dérivation de Gauss et Weingarten. Les formules ainsi dénommées décrivent la variation des vecteurs r_j et m lorsque le point se déplace sur la surface, tout comme les formules de Frénet (*U16.27*) décrivent la variation des vecteurs de la base naturelle causée par le déplacement du

point d'une courbe. Nous avons

$$r_{ij}(A) \equiv \frac{\partial r_i(A)}{\partial u_j} = \sum_{k=1}^{n} \Gamma_{ij}^k(A)\, r_k(A) + \beta_{ij}(A)\, m(A) \tag{1}$$

$$(i, j = 1, \ldots, n),$$

$$m_j(A) \equiv \frac{\partial m(A)}{\partial u_j} = \sum_{k=1}^{n} b_j^k(A)\, r_k(A) \tag{2}$$

$$(j = 1, \ldots, n),$$

où $\Gamma_{ij}^k(A)$, $\beta_{ij}(A)$, $b_j^k(A)$ sont certains coefficients. Il se trouve que *tous ces coefficients s'expriment par les coefficients de la première et de la deuxième formes quadratiques de la surface au point A, ainsi que par leurs dérivées.* Nous avons déjà vu les coefficients b_j^k dans *5.28*; ils satisfont au système d'équations *5.28* (1):

$$b_{ij} = -\sum_{k=1}^{n} b_j^k g_{ik}. \tag{3}$$

Pour résoudre ce système, nous utilisons la matrice $\| g^{is} \|$ qui est l'inverse de la matrice $\| g_{ij} \|$.

Multiplions (3) par g^{is} et sommons sur i; comme

$$\sum_{i=1}^{n} g_{ik} g^{is} = \begin{cases} 1 \text{ pour } k = s, \\ 0 \text{ pour } k \neq s, \end{cases} \tag{4}$$

la sommation du résultat obtenu sur k ne fournira que le terme correspondant à la valeur $k = s$, et nous aurons:

$$\sum_{i=1}^{n} b_{ij} g^{is} = -b_j^s. \tag{5}$$

Ainsi, *les coefficients b_j^s s'expriment par les coefficients des première et deuxième formes quadratiques.*

Il est aisé de trouver les coefficients β_{ij} dans l'égalité (1) en la multipliant scalairement par m:

$$(r_{ij}, m) = b_{ij} = \beta_{ij}(m, m) = \beta_{ij}. \tag{6}$$

Ainsi, *les quantités β_{ij} coïncident avec les coefficients respectifs b_{ij} de la deuxième forme quadratique.*

Quant aux coefficients Γ_{ij}^k, la situation est plus compliquée. En multipliant scalairement (1) par r_s nous trouvons:

$$\Gamma_{ij,\,s} \equiv (r_{ij}, r_s) = \sum_{k=1}^{n} \Gamma_{ij}^k (r_k, r_s) = \sum_{k=1}^{n} \Gamma_{ij}^k g_{ks}. \tag{7}$$

La dérivation de l'égalité $(r_i, r_s) = g_{is}$ par rapport à u_j fournit:

$$(r_{ij}, r_s) + (r_i, r_{sj}) = \frac{\partial g_{is}}{\partial u_j}. \tag{8}$$

En effectuant le changement cyclique des indices, nous obtenons successivement:

$$(r_{js}, r_i) + (r_j, r_{is}) = \frac{\partial g_{ji}}{\partial u_s}, \tag{9}$$

$$(r_{si}, r_j) + (r_s, r_{ji}) = \frac{\partial g_{sj}}{\partial u_i}. \tag{10}$$

En additionnant (8) et (10) et puis soustrayant (9), trouvons:

$$\Gamma_{ij,\,s} = (r_{ij}, r_s) = \frac{1}{2}\left(\frac{\partial g_{is}}{\partial u_j} + \frac{\partial g_{js}}{\partial u_i} - \frac{\partial g_{ij}}{\partial u_s}\right). \tag{11}$$

Pour résoudre les équations (7) par rapport aux quantités Γ_{ij}^k, remplaçons dans (7) l'indice de sommation k par p, multiplions (7) par g^{ks} et sommons sur s. Compte tenu de (4), on obtient:

$$\Gamma_{ij}^k = \sum_{s=1}^{n} \Gamma_{ij,\,s}\, g^{hs} = \frac{1}{2} \sum_{s=1}^{n} \left(\frac{\partial g_{is}}{\partial u_j} + \frac{\partial g_{js}}{\partial u_i} - \frac{\partial g_{ij}}{\partial u_s}\right) g^{ks}. \tag{12}$$

Ainsi, *les coefficients Γ_{ij}^k s'expriment par les seuls coefficients de la première forme quadratique et leurs dérivées.* C'est un fait bien important qui montre que les coefficients Γ_{ij}^k, contrairement à β_{ij} et b_j^k, *appartiennent à la géométrie intrinsèque de la surface.*

Les coefficients $\Gamma_{ij,\,s}$ sont appelés *symboles de Christoffel de 1-ère espèce*, les coefficients Γ_{ij}^k, *symboles de Christoffel de 2-ème espèce.* Les uns et les autres sont également appelés *coefficients de connexion*, puisqu'ils servent à *relier* des caractéristiques géométriques de la surface en des points voisins à l'aide de certaines équations différentielles, ce qu'on verra dans la suite. Les coefficients $\Gamma_{ij,\,s}$, de même que Γ_{ij}^k, sont symétriques par rapport aux indices i et j; pour les quantités $\Gamma_{ij,\,s} = (r_{ij}, r_s)$ ceci résulte de l'égalité $r_{ij} = r_{ji}$, pour Γ_{ij}^k, de la symétrie de $\Gamma_{ij,\,s}$ et de l'équation (12).

Les formules (1), avec les valeurs signalées de Γ_{ij}^k et β_{ij}, s'appellent *formules de Gauss*; les formules (2), avec les valeurs signalées des coefficients b_j^k, sont appelées *formules de Weingarten.*

5.32. Relations entre les coefficients de la première et de la deuxième formes quadratiques. Les coefficients de la première et de la deuxième formes quadratiques ne sont pas indépendants. Les relations entre eux découlent de l'égalité des dérivées mixtes supérieures des vecteurs $r(u)$ et $m(u)$ (en supposant leur existence et continuité) et des formu-

les de dérivation de Gauss et Weingarten. En effet,

$$r_{ikj} \equiv \frac{\partial r_{ik}}{\partial u_j} = \frac{\partial}{\partial u_j}\left(\sum_{s=1}^{n} \Gamma_{ik}^{s} r_s + \beta_{ik} m\right), \tag{1}$$

$$r_{jki} \equiv \frac{\partial r_{jk}}{\partial u_i} = \frac{\partial}{\partial u_i}\left(\sum_{s=1}^{n} \Gamma_{jk}^{s} r_s + \beta_{jk} m\right). \tag{2}$$

Les premiers membres étant identiques, on a

$$\sum_{s=1}^{n} \frac{\partial \Gamma_{ik}^{s}}{\partial u_j} r_s + \sum_{p=1}^{n} \Gamma_{ik}^{p} r_{jp} + \frac{\partial \beta_{ik}}{\partial u_j} m + \beta_{ik} m_j =$$

$$= \sum_{s=1}^{n} \frac{\partial \Gamma_{jk}^{s}}{\partial u_i} r_s + \sum_{p=1}^{n} \Gamma_{jk}^{p} r_{ip} + \frac{\partial \beta_{jk}}{\partial u_i} m + \beta_{jk} m_i. \tag{3}$$

En y portant les quantités r_{jp}, r_{ip} et m_j, m_i des formules de Gauss et Weingarten on obtient:

$$\sum_{s=1}^{n} \frac{\partial \Gamma_{ik}^{s}}{\partial u_j} r_s + \sum_{p=1}^{n} \Gamma_{ik}^{p} \left(\sum_{s=1}^{n} \Gamma_{jp}^{s} r_s + \beta_{jp} m\right) +$$

$$+ \frac{\partial \beta_{ik}}{\partial u_j} m + \beta_{ik} \sum_{s=1}^{n} b_j^s r_s =$$

$$= \sum_{s=1}^{n} \frac{\partial \Gamma_{jk}^{s}}{\partial u_i} r_s + \sum_{p=1}^{n} \Gamma_{jk}^{p} \left(\sum_{s=1}^{n} \Gamma_{ip}^{s} r_s + \beta_{ip} m\right) +$$

$$+ \frac{\partial \beta_{jk}}{\partial u_i} m + \beta_{jk} \sum_{s=1}^{n} b_i^s r_s. \tag{4}$$

Les vecteurs r_s et m étant linéairement indépendants, (4) implique l'égalité pour chaque composante:

$$\frac{\partial \Gamma_{ik}^{s}}{\partial u_j} + \sum_{p=1}^{n} \Gamma_{ik}^{p} \Gamma_{jp}^{s} + \beta_{ik} b_j^s = \frac{\partial \Gamma_{jk}^{s}}{\partial u_i} + \sum_{p=1}^{n} \Gamma_{jk}^{p} \Gamma_{ip}^{s} + \beta_{jk} b_i^s \tag{5}$$

(formule de Gauss),

$$\sum_{p=1}^{n} \Gamma_{ik}^{p} \beta_{jp} + \frac{\partial \beta_{ik}}{\partial u_j} = \sum_{p=1}^{n} \Gamma_{jk}^{p} \beta_{ip} + \frac{\partial \beta_{jk}}{\partial u_i} \tag{6}$$

(formule de Peterson-Codazzi).

La formule (5) peut être mise sous la forme:

$$\beta_{jk}b_i^s-\beta_{ik}b_j^s=\frac{\partial\Gamma_{ik}^s}{\partial u_j}-\frac{\partial\Gamma_{jk}^s}{\partial u_i}+\sum_{p=1}^{n}(\Gamma_{ik}^p\Gamma_{jp}^s-\Gamma_{jk}^p\Gamma_{ip}^s).$$

Multiplions encore par g_{sl} et additionnons sur l'indice s. Alors, en désignant

$$\sum_s b_i^s g_{sl}=c_{il},$$

nous avons

$$\beta_{jk}c_{il}-\beta_{ik}c_{jl}=\sum_{s=1}^{n}\left[\frac{\partial\Gamma_{ik}^s}{\partial u_j}-\frac{\partial\Gamma_{jk}^s}{\partial u_i}+\sum_{p=1}^{n}(\Gamma_{ik}^p\Gamma_{jp}^s-\Gamma_{jk}^p\Gamma_{ip}^s)\right]g_{sl}. \quad (7)$$

Portons dans cette formule les valeurs de β_{jk} (β_{ik}) et c_{il} (c_{jl}) trouvées dans *5.31* (3) et (6):

$$\beta_{jk}=b_{jk},\quad c_{il}=\sum_{s=1}^{n}b_i^s g_{sl}=-b_{il}.$$

Désignons encore $B_{ij,\,kl}=b_{ik}b_{jl}-b_{jk}b_{il}$, nous aboutissons à la formule

$$B_{ij,\,kl}=\sum_{s=1}^{n}\left[\frac{\partial\Gamma_{ik}^s}{\partial u_j}-\frac{\partial\Gamma_{jk}^s}{\partial u_i}+\sum_{p=1}^{n}(\Gamma_{ik}^p\Gamma_{jp}^s-\Gamma_{jk}^p\Gamma_{ip}^s)\right]g_{sl} \quad (8)$$

qui porte elle aussi le nom *de Gauss.*

L'expression $B_{ij,\,kl}=b_{ik}b_{jk}-b_{jk}b_{il}$ est le mineur du deuxième ordre de la matrice symétrique $B=\|\,b_{ik}\,\|$ construit sur les lignes de numéros i et j et les colonnes de numéros k et l.

Dans le second membre de (8) on voit une expression qui ne dépend que des coefficients de la première forme et de leurs dérivées (5.31). Ainsi, *tous les mineurs du deuxième ordre de la matrice de la deuxième forme quadratique sont déterminés d'une façon univoque par la première forme quadratique.* Inversement, les égalités (8) et (6) impliquent les égalités (4) et (3) et, en fin des comptes, l'égalité des seconds membres de (1) et (2).

A son tour, l'égalité $m_{ij}\equiv\dfrac{\partial^2 m}{\partial u_i\partial u_j}=\dfrac{\partial^2 m}{\partial u_j\partial u_i}=m_{ji}$ amène elle aussi à deux relations entre les quantités g_{ij} et b_{ij}. Nous avons:

$$\frac{\partial^2 m\,(u)}{\partial u_i\partial u_j}=\frac{\partial m_i}{\partial u_j}=\frac{\partial}{\partial u_j}\sum_{s=1}^{n}b_i^s r_s=\sum_{s=1}^{n}\frac{\partial b_i^s}{\partial u_j}r_s+\sum_{s=1}^{n}b_i^s r_{sj}=$$

$$=\sum_{s=1}^{n}\frac{\partial b_i^s}{\partial u_j}r_s+\sum_{p=1}^{n}b_i^p\left(\sum_{s=1}^{n}\Gamma_{pj}^s r_s+\beta_{pj}m\right)=$$

$$=\sum_{s=1}^{n}\left(\frac{\partial b_i^s}{\partial u_j}+\sum_{p=1}^{n}b_i^p\Gamma_{pj}^s\right)r_s+\sum_{p=1}^{n}b_j^p\beta_{pj}m.$$

En permutant i et j et égalant les composantes des expressions obtenues, nous avons les égalités

$$\frac{\partial b_i^s}{\partial u_j} + \sum_{p=1}^{n} b_i^p \Gamma_{pj}^s = \frac{\partial b_j^s}{\partial u_i} + \sum_{p=1}^{n} b_j^p \Gamma_{pi}^s, \tag{9}$$

$$\sum_{p=1}^{n} b_i^p \beta_{pj} = \sum_{p=1}^{n} b_j^p \beta_{pi}. \tag{10}$$

En fait, les deux relations ne sont pas nouvelles, mais peuvent être déduites algébriquement de celles déjà établies. Pour obtenir la relation (9), nous partons de l'égalité *5.31* (8)

$$\frac{\partial g_{kq}}{\partial u_j} = \Gamma_{jk,\,q} + \Gamma_{jq,\,k}$$

qui implique

$$\sum_{k=1}^{n} b_i^k \frac{\partial g_{kq}}{\partial u_j} - \sum_{p=1}^{n} b_i^p \Gamma_{pj,\,q} = \sum_{p=1}^{n} b_i^p \Gamma_{qj,\,p},$$

ou bien, en permutant les indices i et j,

$$\sum_{k=1}^{n} b_j^k \frac{\partial g_{kq}}{\partial u_i} - \sum_{p=1}^{n} b_j^p \Gamma_{pi,\,q} = \sum_{p=1}^{n} b_j^p \Gamma_{qi,\,p}.$$

La formule (6) de Peterson-Codazzi peut être écrite comme suit

$$\sum_{p=1}^{n} b_i^p \Gamma_{jq,\,p} + \frac{\partial b_{iq}}{\partial u_j} = \sum_{p=1}^{n} b_j^p \Gamma_{iq,\,p} + \frac{\partial b_{jq}}{\partial u_i},$$

d'où

$$\sum_{k=1}^{n} b_i^k \frac{\partial g_{kq}}{\partial u_j} - \sum_{p=1}^{n} b_i^p \Gamma_{pj,\,q} + \frac{\partial b_{iq}}{\partial u_j} = \sum_{k=1}^{n} b_j^k \frac{\partial g_{kq}}{\partial u_i} - \sum_{p=1}^{n} b_j^p \Gamma_{pi,\,q} + \frac{\partial b_{jq}}{\partial u_i}. \tag{11}$$

D'autre part, l'égalité *5.31* (3)

$$b_{lq} = -\sum_{k=1}^{n} b_l^k g_{kq}$$

conduit à la formule

$$\frac{\partial b_{lq}}{\partial u_m} = -\sum_{k=1}^{n} \frac{\partial b_l^k}{\partial u_m} g_{kq} - \sum_{k=1}^{n} b_l^k \frac{\partial g_{kq}}{\partial u_m}$$

qui permet de mettre l'égalité (11) sous la forme

$$-\sum_{k=1}^{n} \frac{\partial b_i^k}{\partial u_j} g_{kq} - \sum_{p=1}^{n} b_i^p \Gamma_{pj,\,q} = -\sum_{k=1}^{n} \frac{\partial b_j^k}{\partial u_i} g_{kq} - \sum_{p=1}^{n} b_j^p \Gamma_{pi,\,q}.$$

En la multipliant par $-g^{sq}$ et sommant sur q nous aboutissons à l'égalité (9).

Pour établir la relation (10), il suffit d'y porter les expressions *5.31* (5) et (6):

$$b_i^p = -\sum_{k=1}^{n} b_{ik}g^{kp}, \quad \beta_{pj} = b_{pj},$$

après quoi elle prend la forme

$$\sum_{k=1}^{n}\sum_{p=1}^{n} b_{ik}g^{kp}b_{pj} = \sum_{k=1}^{n}\sum_{p=1}^{n} b_{jk}g^{kp}b_{pi}. \tag{12}$$

La relation (12) exprime, évidemment, le fait que la matrice $BG^{-1}B$ est symétrique; or, ceci résulte immédiatement de la symétrie des matrices B et G.

5.33. Dans le cas général, la deuxième forme quadratique n'est pas déterminée d'une façon univoque par la première forme quadratique: par exemple, pour un plan et un cylindre, les premières formes quadratiques coïncident (dans certains systèmes de coordonnées), les deuxièmes étant différentes (la deuxième forme du plan est nulle, la deuxième forme du cylindre non).

Tout de même, les théorèmes démontrés permettent de tirer de la première forme quadratique une information de valeur sur la deuxième forme.

a. Considérons d'abord le cas $n = 2$ où les deux formes sont données par des matrices du deuxième ordre. Conformément à la formule de Gauss, nous trouvons le déterminant de la deuxième forme quadratique d'après la première forme:

$$\det B = B_{12,12} = \sum_{s=1}^{2}\left[\frac{\partial\Gamma_{11}^s}{\partial u_2} - \frac{\partial\Gamma_{12}^s}{\partial u_1} + \sum_{p=1}^{2}\left(\Gamma_{11}^p\Gamma_{2p}^s - \Gamma_{21}^p\Gamma_{1p}^s\right)\right] g_{s2}. \tag{1}$$

En sachant $\det B$ nous trouvons d'après *5.25* (3) la courbure totale de la surface:

$$K = \frac{\det B}{\det G} = \frac{\sum_{s=1}^{2}\left[\frac{\partial\Gamma_{11}^s}{\partial u_2} - \frac{\partial\Gamma_{12}^s}{\partial u_1} + \sum_{p=1}^{2}\left(\Gamma_{11}^p\Gamma_{2p}^s - \Gamma_{21}^p\Gamma_{1p}^s\right)\right] g_{s2}}{g_{11}g_{22} - g_{12}^2}. \tag{2}$$

Ainsi, *la courbure totale d'une surface bidimensionnelle peut être déterminée d'une façon univoque d'après la seule première forme quadratique.* Il en résulte que *la courbure totale en un point donné d'une surface bidimensionnelle ne change pas lors d'une transformation isométrique de la surface* (théorème de Gauss).

b. Il résulte immédiatement du théorème de Gauss qu'il est impossible de transformer isométriquement une partie du plan ($K = 0$) en une partie de la sphère ($K > 0$) ou de la caténoïde ($K < 0$).

c. Si l'on considère une surface bidimensionnelle $P = P_2$ munie d'une métrique non pas empruntée à l'espace euclidien R_3 qui contient la surface, mais donnée d'une manière indépendante (par exemple, pour une surface P_2 disposée dans R_n avec $n > 3$ on peut choisir la métrique induite par ce R_n), alors toute la théorie de la courbure des lignes sur P_2 (§ 5.2) ne marche plus, il n'y a plus de deuxième forme quadratique. Nous sommes pourtant en mesure de calculer la quantité K d'après la formule (2). Nous l'appelons « *courbure totale formelle* », elle appartient à la géométrie intrinsèque de la surface. La question de l'interprétation géométrique de la « courbure totale formelle » se pose tout naturellement. Comme les considérations liées aux normales à la surface sont dans ce cas impossibles, on ne peut plus interpréter la quantité K à l'aide d'une application sphérique (*5.28*), ni comme produit des courbures totales, faute de ces courbures totales. Néanmoins, on peut attribuer un sens géométrique à la quantité K; il sera découvert dans *5.63*.

d. Soit maintenant $n > 2$. Supposons que l'un des mineurs du troisième ordre de la matrice B, par exemple le déterminant de la matrice

$$B_3 = \left\| \begin{matrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{matrix} \right\|,$$

soit non nul. Nous affirmons que, dans ce cas, *tous les éléments* b_{11}, b_{12}, . . ., b_{33} *se déterminent par la forme G d'une façon univoque, au signe* (commun à tous les éléments) *près*. En effet, d'après le théorème de Gauss *a*, nous connaissons tous les compléments algébriques B_{11}, B_{12}, . . ., B_{33} des éléments respectifs b_{11}, b_{12}, . . ., b_{33}. Comme ces compléments algébriques divisés par $\det B_3$ fournissent les éléments de la matrice inverse B_3^{-1}, nous appliquons à l'égalité

$$B_3 B_3^{-1} = E_3$$

le théorème sur le produit de déterminants et trouvons

$$\det B_3 \det \| B_{jk} \| (\det B_3)^{-3} = 1 \quad (j, k = 1, 2, 3),$$

d'où

$$(\det B_3)^2 = \det \| B_{jk} \|,$$

et nous connaissons $\det B_3$ au signe près. Or, $\det B_3$ et les nombres B_{jk} déterminent la matrice B_3^{-1}, donc, aussi la matrice B_3.

Dans le cas considéré, sans nuire à la généralité, on peut admettre que la matrice

$$B_2 = \left\| \begin{matrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{matrix} \right\|$$

soit elle aussi non dégénérée et que $b_{11} \neq 0$. Fixons le signe de l'élément b_{11}; alors les signes de tous les autres éléments de la matrice B_3

sont fixés automatiquement. Considérons la matrice

$$\begin{Vmatrix} b_{11} & b_{12} & b_{1k} \\ b_{21} & b_{22} & b_{2k} \\ b_{j1} & b_{j2} & b_{jk} \end{Vmatrix}$$

avec quelques j et k. Si elle n'est pas dégénérée, ces éléments sont eux aussi déterminés d'une façon univoque (cette fois, le choix du signe est fixé par le signe de b_{11}). Et si elle est dégénérée, i.e. son déterminant est nul, alors sa troisième ligne représente une combinaison linéaire de deux premières (puisque $\det B_2 \neq 0$), les coefficients de la combinaison linéaire s'exprimant, selon les formules de Cramer, par les mineurs du deuxième ordre; par conséquent, ces coefficients, donc aussi les deux premiers éléments de la troisième ligne, sont connus. En appliquant un raisonnement analogue aux colonnes et utilisant les éléments connus de deux premières colonnes, nous trouvons tous les trois éléments de la troisième colonne. Nous voyons ainsi que *s'il y a un mineur non nul du troisième ordre, la matrice B est déterminée d'une façon univoque, au signe près, par la matrice G (u).*

e. Le résultat *d* peut être formulé d'une autre façon. Notamment, *la matrice B peut être restituée d'après la matrice G (u) d'une façon univoque, au signe près, dès qu'il y ait, au point donné de la surface P, trois courbures principales non nulles au moins.* Alors, dans une base canonique formée par les vecteurs principaux $g_1, \ldots, g_n$ (*5.24a*), le mineur principal correspondant du troisième ordre de la matrice B est non nul, donc le rang de la matrice vaut 3 au moins, et il ne reste que d'appliquer le résultat *d*.

5.34. Construction d'une surface d'après ses première et deuxième formes quadratiques.

Théorème (de Bonnet). *Soient* $G = \| g_{ij}(u) \|$, $B = \| b_{ij}(u) \|$ *deux* $n \times n$ *— fonctions matricielles données dans un domaine* $V \subset R_n$ *et y possédant leurs dérivées continues de trois premiers ordres. La matrice G (u) est supposée définie positive, la matrice B (u) liée à G (u) par les formules de Gauss 5.32* (8) *et de Peterson-Codazzi 5.32* (6). *Alors, pour tout point* $u^0 \in V$, *il existe un voisinage* $U = U(u^0) \subset V$ *dans lequel est définie une fonction vectorielle* $r = r(u)$ *telle que, pour la surface correspondante* $P = \{r \in R_{n+1} : r = r(u), u \in U\}$, *les matrices G (u) et B (u) représentent les matrices de la première et de la deuxième formes quadratiques respectivement (5.13 et 5.22). La surface P est définie à la position dans l'espace près; plus exactement, deux surfaces* $r = r^{(1)}(u)$ et $r = r^{(2)}(u)$ *vérifiant les hypothèses du théorème peuvent être rendues identiques, dans un voisinage du point* u^0 *au moins, à l'aide d'une transformation orthogonale dans l'espace* R_{n+1}.

Démonstration. Ecrivons un système d'équations différentielles pour les fonctions vectorielles inconnues $r_1(u), \ldots, r_n(u)$,

$m(u)$:

$$\left.\begin{aligned} \frac{\partial r_i(u)}{\partial u_j} &= \sum_{k=1}^{n} \Gamma_{ij}^k(u)\, r_k(u) + b_{ij}(u)\, m(u), \\ \frac{\partial m(u)}{\partial u_j} &= \sum_{k=1}^{n} b_j^k(u)\, r_k(u), \end{aligned}\right\} \tag{1}$$

où les coefficients $b_{ij}(u)$ sont les éléments donnés de la matrice B, et $\Gamma_{ij}^k(u)$ et $b_j^k(u)$ se déduisent des éléments des matrices G et B d'après les règles signalées dans *5.31*.

Appliquons à la résolution de ce système le théorème de Frobenius (*2.55*). Les conditions d'intégrabilité exigées par ce théorème sont ici remplies, elles coïncident avec les conditions de Gauss et de Peterson-Codazzi qui, à propos, s'obtiennent en égalant les dérivées mixtes. En vertu du théorème de Frobenius, le système (1) admet l'unique solution pour n'importe quel système des données initiales r_i^0 et m^0. Choisissons arbitrairement un point $u^0 \in V$ et trouvons n vecteurs $r_i^0 \in R_{n+1}$ satisfaisant aux conditions

$$(r_i^0, r_j^0) = g_{ij}(u^0) \quad (i, j = 1, \ldots, n). \tag{2}$$

De tels vecteurs r_i^0 $(i = 1, \ldots, n)$ peuvent être trouvés même sur le plan n-dimensionnel $R_n = \{x \in R_{n+1} : x_{n+1} = 0\}$. En effet, la matrice $\| g_{ij}(u^0) \|$ est définie positive et symétrique ; on peut donc en extraire une racine carrée, i.e. une $n \times n$-matrice $\| \xi_{ij} \|$ telle que $\| \xi_{ij} \|^2 = \| g_{ij}(u^0) \|$.* On voit maintenant qu'il est possible de choisir pour vecteurs cherchés r_i^0 $(i = 1, \ldots, n)$ les lignes de la matrice $\| \xi_{ij} \|$ complétées par zéro à la $(n+1)$-ème place, i.e.

$$r_i^0 = (\xi_{i1}, \ldots, \xi_{in}, 0).$$

Après avoir déterminé les vecteurs initiaux $r_1^0, \ldots, r_n^0$, posons $m^{(0)} = (0, 0, \ldots, 0, 1) \in R_{n+1}$. Soit $r_i(u)$ $(i = 1, \ldots, n)$, $m(u)$ $(u \in U \subset V)$ la solution du système (1), pour les conditions initiales signalées, définie dans un voisinage U_1 du point u^0. Il résulte de la symétrie des coefficients $\Gamma_{ij}^k(u)$ et $b_{ij}(u)$ par rapport aux indices i et j que $\frac{\partial r_i}{\partial u_j} = \frac{\partial r_j}{\partial u_i}$; cela veut dire à son tour que, en vertu de *2.56*, il existe dans un voisinage $U \subset U_1$ du point u^0 une fonction vectorielle $r = r(u) = r(u_1, \ldots, u_n)$ pour laquelle

$$\frac{\partial r(u)}{\partial u_i} = r_i(u_1, \ldots, u_n), \tag{3}$$

* La matrice $\| g_{ij}(u^0) \|$, dans la base orthonormée des vecteurs propres $e_1, \ldots, e_n$, se réduit à la forme diagonale avec les éléments positifs $\lambda_1, \ldots, \lambda_n$ sur la diagonale. En tant que matrice $\| \xi_{ij} \|$ on peut choisir la matrice qui, dans la même base, a la forme diagonale avec les éléments $\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n}$ sur la diagonale.

et on peut poser $r(u^0) = 0$. Nous affirmons que la surface P définie par l'équation $r = r(u)$ est une surface cherchée. Pour le prouver, il faut montrer que les matrices de la première et de la deuxième formes quadratiques de la surface P coïncident avec les matrices données $G(u)$ et $B(u)$ respectivement. Nous tirons des équations (1):

$$\left.\begin{aligned}
&\frac{\partial (r_i, r_s)}{\partial u_j} = \left(\frac{\partial r_i}{\partial u_j}, r_s\right) + \left(r_i, \frac{\partial r_s}{\partial u_j}\right) = \\
&= \sum_{k=1}^{n} [\Gamma_{ij}^k (r_k, r_s) + \Gamma_{sj}^k (r_i, r_k)] + b_{ij}(m, r_s) + b_{sj}(m, r_i), \\
&\frac{\partial (r_i, m)}{\partial u_j} = \left(\frac{\partial r_i}{\partial u_j}, m\right) + \left(r_i, \frac{\partial m}{\partial u_j}\right) = \\
&= \sum_{k=1}^{n} \Gamma_{ij}^k (r_k, m) + b_{ij}(m, m) + \sum_{k=1}^{n} b_j^k (r_i, r_k), \\
&\frac{\partial (m, m)}{\partial u_j} = 2\left(\frac{\partial m}{\partial u_j}, m\right) = 2\sum_{k=1}^{n} b_j^k (m, r_k).
\end{aligned}\right\} \quad (4)$$

Considérons ces égalités comme système d'équations différentielles par rapport aux fonctions (r_i, r_s), (r_i, m), (m, m) $(i, s = 1, \ldots, n)$. Le système (4) possède la solution

$$(r_i, r_s) \equiv g_{is}(u), \quad (r_i, m) \equiv 0, \quad (m, m) \equiv 1. \quad (5)$$

En effet, d'après *5.31* (7) et *5.31* (8) on a:

$$\sum_{k=1}^{n} (\Gamma_{ij}^k g_{ks} + \Gamma_{sj}^k g_{ki}) = \Gamma_{ij, s} + \Gamma_{sj, i} = \frac{\partial g_{is}}{\partial u_j},$$

de sorte que la première des équations (4) est satisfaite par les fonctions (5). Il en est de même de la deuxième équation en vertu de la définition de b_j^k (*5.31* (5)), pour la troisième, d'une façon évidente. Les fonctions (5) vérifient les conditions initiales $(r_i^0, r_j^0) = g_{is}(u^0)$, $(r_i^0, m^0) = 0$, $(m^0, m^0) = 1$. Vu le théorème d'unicité *2.52*, nous avons $(r_i, r_s) \equiv g_{is}(u)$, $(r_i, m) \equiv 0$, $(m, m) \equiv 1$ à tous les points u d'un voisinage du point u^0.

Ainsi, $(G(u)\, du, du)$ est la première forme quadratique de la surface P et $m(u)$ son vecteur unitaire normal. Le coefficient (r_{ij}, m) de la deuxième forme quadratique se déduit maintenant de la première équation (1):

$$(r_{ij, m}) = \left(\frac{\partial r_i}{\partial u_j}, m\right) = \sum_{k=1}^{n} \Gamma_{ij}^k (u)(r_k, m) + b_{ij}(u)(m, m) = b_{ij}(u),$$

et nous voyons que la forme $(B\, du, du)$ coïncide avec la deuxième forme quadratique de la surface P. L'existence d'une surface cherchée

est donc établie ; il nous reste à établir son unicité à la position dans l'espace près.

Ce fait découlera de l'unicité de solution du système (1) qui satisfait par les vecteurs r_i et m pour chacune des deux surfaces $P^{(1)}$ et $P^{(2)}$ aux mêmes matrices données G et B si, en plus des coefficients Γ_{ij}^k, b_{ij} et b_i^k, on fixe les valeurs initiales r_i^0 et m^0. Or, si chacune des surfaces $P^{(1)} = \{r = r^{(1)}(u)\}$ et $P^{(2)} = \{r = r^{(2)}(u)\}$ avec les formes quadratiques respectivement égales subit une translation et une transformation orthogonale dans l'espace R_{n+1} de sorte que, pour un u^0, les points $r^{0(1)}$ et $r^{0(2)}$ coïncident, de même que les vecteurs $r_i^{0(1)}$, $m^{0(1)}$ avec les vecteurs $r_i^{0(2)}$, $m^{0(2)}$ respectivement, alors les données initiales des solutions $r_i^{(1)}$, $m^{(1)}$ et $r_i^{(2)}$, $m^{(2)}$ du système (3) s'avèrent égales, ce qui permet d'appliquer le théorème d'unicité. On voit donc que les vecteurs $r_i^{(1)}(u)$ et $r_i^{(2)}(u)$, de même que les vecteurs $m^{(1)}(u)$ et $m^{(2)}(u)$, coïncident pour toutes les valeurs de u d'un voisinage du point u^0. Dans ce cas, les vecteurs $r^{(1)}(u)$ et $r^{(2)}(u)$ coïncident dans le même voisinage, la valeur initiale $r^{0(1)} = r^{0(2)}$ étant fixe. Pour finir, remarquons que l'on peut faire coïncider deux systèmes de vecteurs $r_i^{(1)}$ et $r_i^{(2)}$, ayant les mêmes valeurs des produits scalaires $(r_i^{(1)}, r_j^{(1)}) = (r_i^{(2)}, r_j^{(2)})$, à l'aide d'une transformation orthogonale *. Le théorème est démontré.

5.35. Rigidité des surfaces multidimensionnelles. L'une des conséquences intéressantes du théorème de Bonnet est la *rigidité* des surfaces multidimensionnelles d'une vaste classe, c'est-à-dire l'impossibilité de transformation isométrique d'une telle surface, autre que transformation orthogonale complétée éventuellement par une translation. Il s'agit là de la classe des surfaces n-dimensionnelles dans l'espace $(n+1)$-dimensionnel, pour $n > 2$, qui possèdent en un point donné trois courbures principales non nulles au moins. Pour de telles surfaces, selon *5.33*, la deuxième forme quadratique se détermine d'après la première forme quadratique au signe près, donc la surface elle-même se détermine, d'après *5.34*, d'une façon univoque à une translation et une transformation orthogonale près.

5.36. Le théorème de Bonnet montre que la notion géométrique de surface n-dimensionnelle (suffisamment lisse) dans l'espace R_{n+1}

---

* Si $\{e_i^{(1)}\}$ et $\{e_i^{(2)}\}$ sont deux systèmes obtenus par orthonormalisation des systèmes donnés $\{r_i^{(1)}\}$ et $\{r_i^{(2)}\}$, alors la transformation orthogonale cherchée est celle qui transforme $\{e_i^{(1)}\}$ en $\{e_i^{(2)}\}$. En effet, les éléments de la matrice d'orthonormalisation, donc, aussi de la matrice inverse sont définis d'une façon univoque et ne dépendent que des produits scalaires $(r_i^{(1)}, r_j^{(1)}) = (r_i^{(2)}, r_j^{(2)})$ (*L7.52*). Une application linéaire qui transforme chaque $e_i^{(1)}$ en $e_i^{(2)}$ ($i = 1, \ldots, n$), transforme également l'une en l'autre les combinaisons linéaires correspondantes, en particulier les vecteurs $r_i^{(1)}$ en les vecteurs respectifs $r_i^{(2)}$.

a pour équivalent analytique un couple de $n \times n$-fonctions matricielles $\| g_{ij}(u) \|$ et $\| b_{ij}(u) \|$ liées par les formules de Gauss et de Peterson.

Est-ce qu'on peut donner arbitrairement l'une de ces matrices, par exemple la matrice $G = \| g_{ij}(u) \|$ (naturellement, symétrique et définie positive), puis trouver l'autre, $B = \| b_{ij}(u) \|$, de façon que l'hypothèse du théorème de Bonnet soit remplie? En d'autres termes, est-ce que chaque $n \times n$-matrice $G = \| g_{ij}(u) \|$ (symétrique, définie positive et dérivable un nombre convenable de fois) peut servir de matrice de la première forme quadratique à une surface dans l'espace R_{n+1}?

A ce propos, il existe un théorème de Janet (1926)-Cartan (1927): en supposant l'analyticité des coefficients $g_{ij}(u)$ on démontre qu'il existe une surface n-dimensionnelle voulue (définie localement, i.e. dans un domaine suffisamment petit des valeurs de $u = (u_1, \ldots, u_n)$), mais, en général, dans l'espace euclidien $\frac{n(n+1)}{2}$-dimensionnel et non pas dans l'espace $(n+1)$-dimensionnel, et qu'il est en général impossible d'abaisser ce nombre $\frac{n(n+1)}{2}$. En particulier, pour une forme de deux variables u_1, u_2, une surface bidimensionnelle cherchée est déjà placée dans l'espace de dimension $\frac{2 \cdot 3}{2} = 3$, ce qui est conforme à la position du problème, et il existe donc la deuxième forme quadratique cherchée; mais déjà pour une forme de trois variables u_1, u_2, u_3, une surface tridimensionnelle cherchée est placée dans l'espace de dimension $\frac{3 \cdot 4}{2} = 6$, et il y a des formes pour lesquelles, comme montre Cartan, il est impossible de mettre la surface correspondante dans l'espace de dimension 4 ou 5, donc, la deuxième forme quadratique cherchée n'existe pas. Voir à ce propos le livre: E. Cartan. « Les systèmes différentiels extérieurs et leurs applications géométriques », Paris, Hermann, 1945 et l'article: М. Л. Громов и В. А. Рохлин, Вложения и погружения в римановой геометрии, УМН 25, № 5 (1970).

§ 5. 4. Lignes géodésiques et systèmes de coordonnées qui y sont liés

5.41. Lignes géodésiques.

a. Soit $P_n \subset R_{n+1}$ une surface donnée par son rayon vecteur $r = r(u)$, $u = (u_1, \ldots, u_n) \in V \subset R_n$. Considérons la courbe $L = \{r \in R_{n+1}: r = r(u(s)), a \leqslant s \leqslant b\}$ paramétrisée par la longueur d'arc s comptée à partir d'une point fixe. Développons le vecteur de courbure $\frac{d^2r}{ds^2}$ en un point M de la courbe L suivant la base $r_1, \ldots, r_n, m$. Nous avons:

$$\frac{d^2r}{ds^2} = \frac{d}{ds}\left(\frac{dr}{ds}\right) = \frac{d}{ds}\sum_{i=1}^{n}\frac{\partial r}{\partial u_i}\frac{du_i}{ds} =$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2 r}{\partial u_i \partial u_j}\frac{du_i}{ds}\frac{du_j}{ds} + \sum_{i=1}^{n}\frac{\partial r}{\partial u_i}\frac{d^2u_i}{ds^2}.$$

En exprimant les quantités $\frac{\partial^2 r}{\partial u_i \partial u_j} \equiv r_{ij}$ d'après la formule de Gauss *5.31* (1), nous obtenons

$$\frac{d^2r}{ds^2} = \sum_{i,j=1}^{n} \left(\sum_{k=1}^{n} \Gamma_{ij}^k r_k + b_{ij} m \right) \frac{du_i}{ds} \frac{du_j}{ds} + \sum_{k=1}^{n} \frac{d^2u_k}{ds^2} r_k =$$

$$= \sum_{k=1}^{n} \left(\sum_{i,j=1}^{n} \Gamma_{ij}^k \frac{du_i}{ds} \frac{du_j}{ds} + \frac{d^2u_k}{ds^2} \right) r_k + \sum_{i,j=1}^{n} b_{ij} \frac{du_i}{ds} \frac{du_j}{ds} m. \quad (1)$$

Le premier terme à droite dans (1) se trouve dans le plan tangent Π_n et s'appelle *vecteur de courbure géodésique de la courbe L au point M* ; le deuxième terme qui est normal à Π_n s'appelle *vecteur de courbure forcée*. Les composantes du vecteur de courbure géodésique contiennent, en plus de $\frac{du_i}{ds}$ et $\frac{d^2u_i}{ds^2}$, les symboles de Christoffel de deuxième espèce Γ_{ij}^k qui s'expriment, à leur tour, par les coefficients de la première forme quadratique, donc ne varient pas en transformant isométriquement la surface P_n. Les composantes du vecteur de courbure normale s'expriment par les coefficients de la deuxième forme quadratique, donc peuvent varier en transformant isométriquement la surface P_n.

b. *La courbe L s'appelle ligne géodésique sur la surface P_n si, en tout son point, sa courbure géodésique est nulle.* Ainsi, la propriété d'une courbe d'être ligne géodésique n'est pas violée en transformant la surface isométriquement. Evidemment, la définition formulée est équivalente à la suivante: *la courbe L est une ligne géodésique si, en tout son point où la courbure n'est pas nulle, sa normale principale coïncide avec la normale à la surface au même point.*

c. La courbure forcée de chaque ligne sur un plan est nulle; par conséquent, la courbure géodésique de chaque telle ligne vaut sa courbure totale. Pour une ligne géodésique, cela signifie que sa courbure totale est identiquement nulle. Ce sont les droites, et seules les droites, qui possèdent cette propriété sur un plan.

d. Les lignes géodésiques sur un cylindre et sur un cône sont faciles à trouver en développant chacune de ces surfaces sur un plan avec son réseau des droites, i.e. de ses lignes géodésiques, et en utilisant l'invariance des géodésiques par rapport à une transformation isométrique. Pour le cylindre, ce sont ses génératrices, parallèles et lignes hélicoïdales qui se développeront en les droites du plan (fig. 5.4-1). Pour le cône, ce sont ses génératrices (elles deviendront droites passant par l'image du sommet) et certaines courbes qui descendent jusqu'à une parallèle puis remontent (fig. 5.4-2).

e. Sur une sphère, les arcs des grands cercles possèdent une normale principale dirigée vers le centre de la sphère, donc colinéaire à la

normale de la sphère. Par conséquent, les arcs des grands cercles sont des géodésiques sur la sphère, et un peu plus bas (*5.42b*) nous verrons que toute géodésique sur la sphère est un arc de grand cercle.

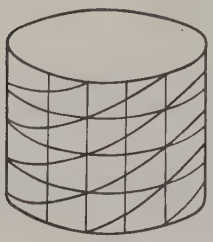


Fig. 5.4-1

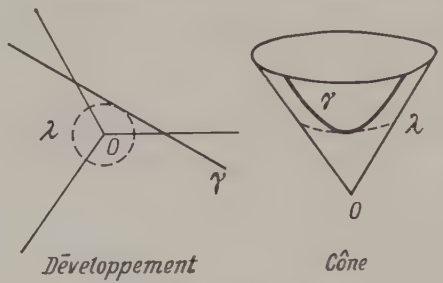


Fig. 5.4-2

Encore un exemple sera vu dans *5.47* (géodésiques sur une surface de révolution).

5.42. Equations différentielles des lignes géodésiques.

a. La définition des lignes géodésiques donnée dans *5.41* est remarquable par le fait qu'on peut l'étendre aux surfaces n-dimensionnelles P_n munies d'une métrique qui n'est pas forcément empruntée à l'espace euclidien R_{n+1}, mais qui est donnée par une matrice symétrique et définie positive quelconque (dérivable):

$$G = \| g_{ik}(u) \|, \quad i, k = 1, \ldots, n.$$

Dans ce cas, nous *définissons* les lignes géodésiques comme lignes sur la surface P_n qui annulent les expressions

$$\sum_{i,\, j=1}^{n} \Gamma_{ij}^{k} \frac{du_i}{ds} \frac{du_j}{ds} + \frac{d^2 u_k}{ds^2} \tag{1}$$

appartenant à la géométrie intrinsèque de la surface.

Ainsi, dans tous les cas, sur une ligne géodésique sont satisfaites les équations

$$\frac{d^2 u_k(s)}{ds^2} = - \sum_{i,\, j=1}^{n} \Gamma_{ij}^{k}(u) \frac{du_i(s)}{ds} \frac{du_j(s)}{ds} \quad (k = 1, \ldots, n) \tag{2}$$

qui, inversement, peuvent servir de définition d'une ligne géodésique en cas de $P_n \subset R_{n+1}$. Le système (2) représente un système de n équations différentielles ordinaires du deuxième ordre par rapport à n fonctions $u_1(s), \ldots, u_n(s)$, ce système étant résolu par rapport aux dérivées secondes et ses deuxièmes membres représentant des polynômes du second degré par rapport aux dérivées premières.

Théorème. *Si, au voisinage d'un point donné* $M_0 = (u_1^0, \ldots, u_n^0)$, *les fonctions* $g_{ij}(u)$ *admettent les dérivées continues, alors il existe une et une seule géodésique issue du point* M_0 *dans chaque direction donnée* $\{du_1, \ldots, du_n\}$.

Démonstration. Pour le point donné $M_0 = (u_1^0, \ldots, u_n^0)$ et pour la direction fixe $\{du_1, \ldots, du_n\}$, envisageons de façon formelle le système (2) complété des conditions initiales

$$u_k(0) = u_k^0, \quad \frac{du_k(0)}{ds} = v_k^0, \; k = 1, \ldots, n, \tag{3}$$

où les nombres v_k^0 sont proportionnels aux nombres du_k et normalisés par la condition

$$\sum_{i,j=1}^{n} g_{ij}(u^0) v_i^0 v_j^0 = 1.$$

Les fonctions $\Gamma_{ij}^k(u) = \frac{1}{2} g^{ks}\left(\frac{\partial g_{is}}{\partial u_j} + \frac{\partial g_{js}}{\partial u_i} - \frac{\partial g_{ij}}{\partial u_s}\right)$ étant continues par hypothèse, et les seconds membres du système (2), qui sont quadratiques par rapport aux dérivées, vérifiant la condition de Lipschitz par rapport aux variables $\frac{du_i}{ds}$, on peut appliquer le théorème d'existence et d'unicité de solution (cf. *U13.42* et *U13.51*) en vertu duquel le système (2) avec les conditions initiales (3) admet une et une seule solution $u_i = u_i(s)$, $0 \leqslant s \leqslant s_0$, $i = 1, \ldots, n$. Il lui correspond une courbe L sur la surface P_n qui passe par le point M dans la direction $\{du_i\}$. Si nous montrons que le paramètre formel s sur la courbe L représente la longueur d'arc, il en résultera, vu (2), que L est une ligne géodésique. Nous avons donc à vérifier l'égalité

$$I(s) \equiv \sum_{i,j=1}^{n} g_{ij}(u(s)) \frac{du_i}{ds} \frac{du_j}{ds} = 1.$$

Le long de la ligne L nous avons

$$I'(s) = \sum_{i,j,k=1}^{n} \frac{\partial g_{ij}}{\partial u_k} \frac{du_i}{ds} \frac{du_j}{ds} \frac{du_k}{ds} + \\ + \sum_{i,j=1}^{n} g_{ij}\left(\frac{d^2u_i}{ds^2} \frac{du_j}{ds} + \frac{du_i}{ds} \frac{d^2u_j}{ds^2}\right).$$

En substituant aux dérivées secondes leurs expressions du système (2) et appliquant les formules *5.31* (7) et (8)

$$\frac{\partial g_{ij}}{\partial u_k} = \Gamma_{ik,j} + \Gamma_{jk,i} \text{ et } \Gamma_{ij,l} = \sum_{k=1}^{n} \Gamma_{ij}^k g_{kl},$$

nous trouvons

$$I'(s) = \sum_{i,j,k=1}^{n} \frac{\partial g_{ij}}{\partial u_k} \frac{du_i}{ds} \frac{du_j}{ds} \frac{du_k}{ds} -$$

$$- \sum_{k,l=1}^{n} g_{kl} \Big[\sum_{i,j=1}^{n} \Big(\Gamma_{ij}^{k} \frac{du_i}{ds} \frac{du_j}{ds} \frac{du_l}{ds} + \Gamma_{ij}^{l} \frac{du_i}{ds} \frac{du_j}{ds} \frac{du_k}{ds} \Big) \Big] =$$

$$= \sum_{i,j,k=1}^{n} \Gamma_{jk,\,i} \frac{du_i}{ds} \frac{du_j}{ds} \frac{du_k}{ds} + \sum_{i,j,k=1}^{n} \Gamma_{jk,\,i} \frac{du_i}{ds} \frac{du_j}{ds} \frac{du_k}{ds} -$$

$$- \sum_{i,j,l=1}^{n} \Gamma_{ij,\,l} \frac{du_i}{ds} \frac{du_j}{ds} \frac{du_l}{ds} - \sum_{i,j,k=1}^{n} \Gamma_{ij,\,k} \frac{du_i}{ds} \frac{du_j}{ds} \frac{du_k}{ds} = 0$$

puisque les sommes ne se diffèrent que par les désignations des indices. Ainsi $I'(s) \equiv 0$, $I(s) \equiv I(0) = \sum_{i,j=1}^{n} g_{ij}(u(0)) \frac{du_i(0)}{ds} \frac{du_j(0)}{ds} =$ $= 1$, ce qui achève la démonstration.

b. Il résulte du fait établi, en particulier, que les arcs des grands cercles épuisent toutes les lignes géodésiques sur une sphère. En effet, si nous fixons sur une ligne géodésique L un point M et menons par M un arc Γ de grand cercle dans la direction de la ligne L, alors, d'après le théorème d'unicité démontré, l'arc Γ coïncidera entièrement avec L.

c. On peut montrer qu'il existe un voisinage U du point M dans lequel, par tout son point A, il passe une et une seule géodésique issue de M; qui plus est, n'importe quels deux points A et B de U sont liés par une et une seule géodésique dont tout l'arc de A à B se trouve entièrement dans U (cf. exercices 12 et 13).

5.43. Parallèles géodésiques.

a. Soit L une ligne sur une surface bidimensionnelle P_2. Menons par chaque point M de la ligne L, dans la direction orthogonale à L sur la surface P_2, une ligne géodésique $\gamma(M)$. En vertu du théorème *5.42a*, une telle géodésique existe et est unique pour tout point M. Sur chacune de ces géodésiques, portons, dans une direction fixe, un même arc w compté du point M. Le lieu géométrique des extrémités des arcs géodésiques ainsi obtenus représente une ligne L_w qui s'appelle *parallèle géodésique à la ligne L à la distance w*. Il s'avère que *toutes les parallèles géodésiques* L_w, $0 \leqslant w \leqslant \varepsilon$ *sont orthogonales aux lignes géodésiques* $\gamma(M)$. Nous démontrons cette proposition pour le cas de n dimensions.

b. Supposons que, sur une surface n-dimensionnelle $P_n = \{r \in R_{n+1}: r = r(u), u \in U \in R_n\}$, on ait une surface $(n-1)$-dimensionnelle suffisamment lisse L_{n-1}. En tout point $M \in L_{n-1}$ est définie une direction orthogonale sur la surface P_n. Soit $\gamma(M)$ la ligne géodésique issue de M et orthogonale à L_{n-1}. Portons sur $\gamma(M)$, dans une direction fixe, un arc fixe w *. Le lieu géométrique des points ainsi obtenus s'appelle *surface géodésiquement parallèle* à L_{n-1} à *la distance w et se désigne par* L_{n-1}^{w}.

Théorème. *Pour tout point $M_0 \in L_{n-1}$, il existe un voisinage $V(M_0)$ dans lequel chaque surface L_{n-1}^{w} coupe orthogonalement toutes les géodésiques $\gamma(M)$.*

Démonstration. Dans le cas général, la surface L_{n-1} est donnée par un système d'équations à $n-1$ paramètres $\tau_1, \ldots, \tau_{n-1}$:

$$\left.\begin{aligned} u_1 &= u_1(\tau_1, \ldots, \tau_{n-1}), \\ &\cdots\cdots\cdots\cdots \\ u_n &= u_n(\tau_1, \ldots, \tau_{n-1}), \end{aligned}\right\} \tag{1}$$

le rang de la matrice de Jacobi $\frac{\partial(u_1, \ldots, u_n)}{\partial(\tau_1, \ldots, \tau_{n-1})}$ valant $n-1$ au point M_0, donc aussi à un voisinage de ce point. Supposons que le mineur

$$\begin{vmatrix} \frac{\partial u_2}{\partial \tau_1} & \cdots & \frac{\partial u_2}{\partial \tau_{n-1}} \\ \cdots & \cdots & \cdots \\ \frac{\partial u_n}{\partial \tau_1} & \cdots & \frac{\partial u_n}{\partial \tau_{n-1}} \end{vmatrix}$$

ne soit pas nul au point M_0. Alors, d'après le théorème de la fonction inverse, nous sommes en mesure de résoudre les $n-1$ dernières équations du système (1) par rapport aux paramètres $\tau_1, \ldots, \tau_{n-1}$ en les présentant comme fonctions de $u_2, \ldots, u_n$, puis de les porter dans la première équation; l'équation obtenue de la surface L_{n-1} sera de la forme

$$u_1 = \varphi(u_2, \ldots, u_{n-1}). \tag{2}$$

D'après le théorème *2.36a*, les fonctions (2) sont autant lisses que les seconds membres du système (1). A présent, on peut choisir pour

* Le théorème *1.61* nous garantit l'existence d'un $\delta > 0$ et d'un voisinage U du point donné M_0 sur la surface P_n tels que les lignes géodésiques $\gamma(M)$ issues de n'importe quel point M dans le voisinage U orthogonalement à L_{n-1} seront définies, au moins, pour toutes les valeurs de w avec $|w| < \delta$. (Dans le théorème *1.61*, il s'agit d'une équation du premier ordre; or, une équation d'ordre supérieur se réduit bien à une équation du premier ordre comme c'est indiqué dans *U13.51*.)

paramètres déterminant la position de n'importe quel point $M \in L_{n-1}$ au voisinage du point M_0 les quantités

$$v_1 = u_1 - \varphi(u_2, \ldots, u_{n-1}), \quad v_2 = u_2, \ldots, v_n = u_n,$$

puisque le jacobien de la matrice $\frac{\partial(v_1, \ldots, v_n)}{\partial(u_1, \ldots, u_n)}$ est identiquement égal à l'unité. Les vecteurs $\frac{\partial r(M_0)}{\partial v_2}, \ldots, \frac{\partial r(M_0)}{\partial v_n}$ sont tangents aux lignes de coordonnées $v_2, \ldots, v_n$ respectivement, qui passent sur la surface $v_1 = 0$, donc sur L_{n-1}, et se trouvent, par conséquent, dans le plan tangent à L_{n-1}.

Examinons maintenant, pour un $\varepsilon > 0$, tous les ensembles numériques $w_1, \ldots, w_n$ soumis aux conditions:

$$|w_1| < \varepsilon, \quad |w_2 - v_2^0| < \varepsilon, \ldots, |w_n - v_n^0| < \varepsilon$$

où $v_2^0 = u_2^0, \ldots, v_n^0 = u_n^0$. Faisons correspondre à chaque ensemble $w_1, \ldots, w_n$ un point M de la surface P_n selon la règle suivante: en posant $v_1 = 0$, $v_2 = w_2, \ldots, v_n = w_n$ choisissons un point M' sur la surface L_{n-1}, puis trouvons sur la ligne géodésique $\gamma(M')$ le point M à la distance w_1 de M (suivant l'arc orienté). Les coordonnées $v_1, \ldots, v_n$ du point M sont des fonctions continues et dérivables de la variable w_1 et des paramètres $w_2, \ldots, w_n$, donc le jacobien de la matrice $\frac{\partial(v_1, \ldots, v_n)}{\partial(w_1, \ldots, w_n)}$ est une fonction continue de $w_1, \ldots, w_n$. Montrons qu'il est non nul au point M_0. A cette fin, il faut prouver que les vecteurs $\frac{\partial r(v)}{\partial w_1}, \ldots, \frac{\partial r(v)}{\partial w_n}$ sont linéairement indépendants à ce point. Or, au point M_0, les vecteurs $\frac{\partial r}{\partial w_2}, \ldots, \frac{\partial r}{\partial w_n}$ coïncident avec les vecteurs respectifs $\frac{\partial r}{\partial v_2}, \ldots, \frac{\partial r}{\partial v_n}$ qui sont linéairement indépendants et se trouvent dans le plan tangent à la surface L_{n-1}; quant au vecteur $\frac{\partial r(v)}{\partial w_1}$, il est unitaire, orthogonal à L_{n-1} et, par cela, linéairement indépendant de $\frac{\partial r}{\partial w_2}, \ldots, \frac{\partial r}{\partial w_n}$. Il en découle que les quantités $w_1, \ldots, w_n$, elles aussi, peuvent être choisies pour nouveaux paramètres au voisinage du point M_0. Dans ce système de coordonnées, nous avons $g_{11}(w) = \left|\frac{\partial r}{\partial w_1}\right|^2 = 1$, car w_1 est la longueur d'arc. Ensuite, à l'équation générale des géodésiques *5.42* (2)

$$\frac{d^2 w_k}{ds^2} = -\sum_{i,j=1}^{n} \Gamma_{ij}^k \frac{dw_i}{ds} \frac{dw_j}{ds}, \tag{3}$$

satisfait, dans notre cas, le système des fonctions $w_1 = s$, $w_2 = C_2, \ldots, w_n = C_n$; en les portant dans les équations (3) nous

voyons que, *dans le système des paramètres* $w_1, \ldots, w_n$, *les quantités* Γ_{11}^k $(k = 1, \ldots, n)$ *sont nulles.* Mais alors on a aussi

$$\Gamma_{11,\,s} = \sum_{k=1}^{n} \Gamma_{11}^k g_{ks} = 0 \quad (s = 1, \ldots, n).$$

Comme $\Gamma_{11,\,s} = \frac{1}{2}\left(\frac{\partial g_{1s}}{\partial w_1} + \frac{\partial g_{1s}}{\partial w_1} - \frac{\partial g_{11}}{\partial w_s}\right)$ et $\frac{\partial g_{11}}{\partial w_s} = 0$, la fonction $g_{1s}(w)$ s'avère constante sur chaque ligne où sont constants $w_2, \ldots, w_n$, c'est-à-dire sur chaque géodésique $\gamma(M')$. La courbe $\gamma(M')$ étant, par construction, orthogonale à L_{n-1} au point M', de sorte que $g_{1s}(0) = 0$, on a $g_{1s}(w) = 0$ pour tous les w_1. C'est ce qui veut dire que la géodésique $\gamma(M')$ est orthogonale à la parallèle géodésique L_{n-1}^w. Le théorème est démontré.

Le système de coordonnées $w_1, \ldots, w_n$ que nous venons de construire sur la surface P_n, au voisinage du point M_0, est dit *semi-géodésique de base* L_{n-1}.

5.44. Utilisons le système semi-géodésique de coordonnées pour établir la propriété extrémale suivante des lignes géodésiques:

a. Théorème. *Considérons sur la surface* P_n *une ligne géodésique* γ *qui passe par deux points suffisamment proches* A *et* B, *de même que toutes les autres courbes* β *(rectifiables) passant sur* P_n *par les points* A *et* B *dans un petit voisinage de la ligne* γ. *Nous affirmons que parmi toutes ces lignes la ligne géodésique* γ *a la moindre longueur.*

Démonstration. Menons par le point A une surface L_{n-1} orthogonale à la ligne γ et choisissons L_{n-1} pour base d'un système semi-géodésique de coordonnées au voisinage du point A. La ligne γ devient l'une des lignes de coordonnées du système semi-géodésique. En admettant que le point B et la ligne β se trouvent dans le petit domaine sur la surface P_n où le système semi-géodésique construit est valable, écrivons l'expression de la longueur de β:

$$s(\beta) = \int_A^B ds(\beta) = \int_A^B \sqrt{\sum_{i,\,j=1}^{n} g_{ij}(w)\, dw_i\, dw_j} =$$

$$= \int_A^B \sqrt{dw_1^2 + \sum_{i,\,j=2}^{n} g_{ij}(w)\, dw_i\, dw_j}.$$

La dernière égalité est due au fait que $g_{11} = 1$, $g_{1s} = 0$ $(s = 2, \ldots, n)$. Notons que la $(n-1) \times (n-1)$-matrice $\| g_{ij}(w) \|$ $(i, j = 2, \ldots, n)$ est définie positive. Par conséquent,

$$s(\beta) \geqslant \int_A^B dw_1 = w_1(B) - w_1(A) = s(\gamma),$$

ce qu'il fallait.

b. Remarque. Sur le cylindre C (fig. 5.4-3), en plus de la géodésique hélicoïdale liant les points A et B, il en existe d'autres, plus courtes (par exemple, le segment AB). Il n'y a pas de contradiction avec le théorème *a*, puisque, dans ce théorème, on ne compare la longueur de la géodésique qu'avec celles des lignes suffisamment proches.

c. Remarque. Sur la sphère S (fig. 5.4-4), considérons une géodésique méridienne qui va du pôle sud A par le pôle nord O et aboutit à un point B. Désignons par C le premier point d'intersection de la méridienne AOB avec la parallèle β contenant le point B. *Dans tout voisinage de cette ligne géodésique,*

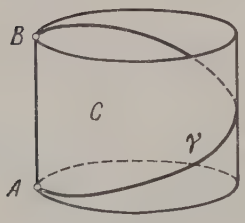


Fig. 5.4-3

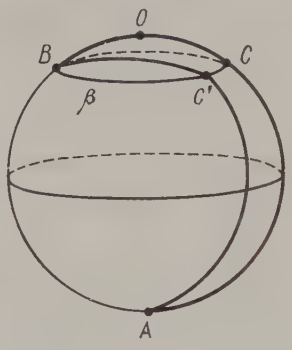


Fig. 5.4-4

il y a une ligne joignant A et B, et plus courte que l'arc AOB. Telle est, par exemple, la ligne $AC'B$ où C' se trouve sur la parallèle β dans la proximité fixée à l'avance du point C, AC' est l'arc de la méridienne, $C'B$ l'arc du grand cercle. Les arcs AC' et AC sont, évidemment, d'une même longueur, mais l'arc $C'B$ est plus court que CB, la corde $C'B$ étant plus courte que la corde CB qui est le diamètre de la parallèle β.

Ici la contradiction apparente avec le théorème *a* s'explique par le fait que la géodésique ACB est trop « longue » pour qu'on puisse l'inclure dans le système semi-géodésique de coordonnées qui ne se construit, comme nous le savons, que localement, au voisinage du point donné.

5.45. Calcul des caractéristiques principales d'une surface bidimensionnelle dans un système semi-géodésique des paramètres.

a. Nous avons déjà vu dans *5.43b* que les quantités Γ^k_{11} et $\Gamma_{11,\,s}$ sont nulles dans un système semi-géodésique de coordonnées. Calculons, pour une surface bidimensionnelle P_2, les quantités qui restent, Γ^k_{ij} et $\Gamma_{ij,\,s}$. Nous avons établi dans *5.43b* que la première forme quadratique peut être écrite, dans un système semi-géodésique des paramètres w_1, w_2, comme suit

$$ds^2 = dw_1^2 + g_{22}(w_1, w_2)\, dw_2^2.$$

Par conséquent, il ne reste dans les expressions des symboles de Christoffel de première espèce

$$\Gamma_{ij,\,s} = \frac{1}{2}\left(\frac{\partial g_{is}}{\partial w_j} + \frac{\partial g_{js}}{\partial w_i} - \frac{\partial g_{ij}}{\partial w_s}\right)$$

que les termes où, parmi les indices $i, j, s = 1, 2$, il y a deux au moins égaux à 2. Ainsi, en désignant pour abréger $g_{22}(w_1, w_2) = G$, $w_1 = w$, $w_2 = u$, nous avons:

$$\Gamma_{12,1} = \Gamma_{21,1} = 0, \quad \Gamma_{12,2} = \Gamma_{21,2} = \frac{1}{2} G_w,$$

$$\Gamma_{22,1} = -\frac{1}{2} G_w, \quad \Gamma_{22,2} = \frac{1}{2} G_u.$$

Ensuite, vu la formule $\Gamma_{ij}^k = \sum_{s=1}^{2} \Gamma_{ij,s} g^{sk}$ et le fait que l'inverse de la matrice $\| g_{ij} \| = \begin{Vmatrix} 1 & 0 \\ 0 & G \end{Vmatrix}$ est la matrice $\| g^{sk} \| = \begin{Vmatrix} 1 & 0 \\ 0 & 1/G \end{Vmatrix}$, nous trouvons:

$$\Gamma_{12}^1 = \Gamma_{21}^1 = 0, \quad \Gamma_{12}^2 = \Gamma_{21}^2 = \Gamma_{12,2} g^{22} = \frac{1}{2} \frac{Gw}{G},$$

$$\Gamma_{22}^1 = -\frac{1}{2} G_w, \quad \Gamma_{22}^2 = \frac{1}{2} \frac{G_u}{G}.$$

b. Calculons à présent, dans le système semi-géodésique, la courbure totale de la surface P_2. Pour le faire, trouvons d'abord le déterminant $B_{12,12}$ de la deuxième forme quadratique d'après la formule *5.33* (1):

$$B_{12,12} = \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix} = \sum_{s=1}^{2} \left[\frac{\partial \Gamma_{11}^s}{\partial u} - \frac{\partial \Gamma_{12}^s}{\partial w} + \sum_{p=1}^{2} (\Gamma_{11}^p \Gamma_{2p}^s - \Gamma_{21}^p \Gamma_{1p}^s) \right] g_{s2}.$$

Ici il ne reste que les termes avec $s = 2$ et $p = 2$. On obtient:

$$B_{12,12} = -G \left(\frac{1}{2} \frac{G_w}{G} \right)_w - G \left(\frac{1}{2} \frac{G_w}{G} \right)^2 =$$

$$= -\frac{G}{2} \frac{G G_{ww} - G_w^2}{G^2} - \frac{G_w^2}{4G} = -\frac{G_{ww}}{2} + \frac{1}{4} \frac{G_w^2}{G}.$$

La courbure totale K de la surface P_2 se trouve donc égale à

$$K = \frac{B_{12,12}}{G_{12,12}} = -\frac{G_{ww}}{2G} + \frac{1}{4} \frac{G_w^2}{G^2}. \tag{1}$$

Cette expression peut être rendue encore plus compacte si l'on remarque que

$$(\sqrt{G})_w = \frac{1}{2} G^{-1/2} G_w, \quad (\sqrt{G})_{ww} = -\frac{1}{4} G^{-3/2} G_w^2 + \frac{1}{2} G^{-1/2} G_{ww},$$

$$\frac{(\sqrt{G})_{ww}}{\sqrt{G}} = -\frac{1}{4} \frac{G_w^2}{G^2} + \frac{1}{2} \frac{G_{ww}}{G} = -K,$$

de sorte que

$$K = -\frac{(\sqrt{G})_{ww}}{\sqrt{G}}. \tag{2}$$

Ainsi, dans un système semi-géodésique sur une surface bidimensionnelle P_2, la fonction $\sqrt{G}$ est reliée à la courbure par l'équation différentielle

$$(\sqrt{G})_{ww} + K\sqrt{G} = 0. \tag{3}$$

5.46. Restitution de la première forme quadratique d'après la courbure totale donnée. Considérons une surface $P = P_2$ en tout point M de laquelle on connaît sa courbure totale $K = K(M)$. Fixons sur la surface un point M_0 et construisons dans son voisinage un système semi-géodésique spécial de coordonnées, ayant une ligne géodésique L pour base et la longueur d'arc $w_2 = u$ comptée d'un point fixe (par exemple de M_0) pour paramètre sur cette ligne. La première forme quadratique est

$$ds^2 = dw^2 + G(w, u)\, du^2,$$

où

$$G(0, u) = 1, \quad \sqrt{G(0, u)} = 1. \tag{1}$$

En portant dans l'équation générale des géodésiques *5.42* (2)

$$\frac{d^2 w_k}{ds^2} = -\sum_{i,\, j=1}^{2} \Gamma_{ij}^k \frac{dw_i}{ds} \frac{dw_j}{ds}$$

les fonctions connues $w_1 = 0$, $w_2 = s$, nous obtenons $\Gamma_{22}^1(0, u) = \Gamma_{22}^2(0, u) = 0$. D'où $\Gamma_{22,\,1}(0, u) = g_{11}\Gamma_{22}^1(0, u) + g_{12}\Gamma_{22}^2(0, u) = 0$. Comme (d'après *5.45a*) $\Gamma_{22}^1(0, u) = -\frac{1}{2} G_w(0, u)$, on a encore

$$G_w(0, u) = 0, \quad (\sqrt{G(0, u)})_w = 0. \tag{2}$$

Dans l'équation

$$(\sqrt{G})_{ww} + K\sqrt{G} = 0 \tag{3}$$

à laquelle satisfait la courbure totale (*5.45* (3)), considérons K comme fonction connue (de w et u), et $\sqrt{G}$ comme fonction inconnue des mêmes paramètres. Pour l'équation différentielle du deuxième ordre (3), nous avons les conditions initiales (1), (2) qui permettent de restituer la fonction $G(w, u)$ d'une façon univoque dans le voisinage indiqué du point M_0. Nous arrivons, en particulier, au théorème d'unicité suivant :

Théorème. *Si, pour deux surfaces P_2 et $\widetilde{P}_2$, dans des systèmes semi-géodésiques spéciaux, la courbure totale K s'exprime par une même fonction des coordonnées, alors ces surfaces sont isométriques ; une isométrie est donnée par l'application qui conserve les coordonnées semi-géodésiques spéciales.*

La réciproque est également immédiate: si deux surfaces P_2 et $\widetilde{P}_2$ sont isométriques, alors, dans les systèmes semi-géodésiques correspondant l'un à l'autre, la courbure totale des deux surfaces représente une même fonction *5.45* (2) des coefficients de la première forme quadratique, donc des coordonnées.

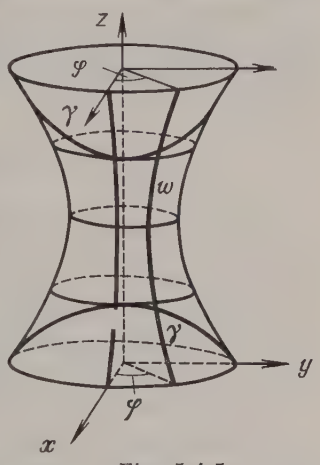


Fig. 5.4-5

5.47. Pour conclure le présent paragraphe, considérons un exemple important.

Géodésiques sur une surface de révolution. Sur une surface de révolution de génératrice $\rho = \rho(z)$ (*5.15f*), les méridiennes sont, évidemment, lignes géodésiques, puisque la normale principale à la méridienne coïncide avec la normale à la surface. La famille des parallèles et des méridiennes de la surface de révolution représente un système semi-géodésique naturel avec n'importe quelle parallèle pour base; le rôle des coordonnées est détenu par la longueur d'arc le long de la méridienne et l'angle polaire (fig. 5.4-5).

Dans ces coordonnées, la première forme quadratique est (*5.15f*):

$$ds^2 = dw^2 + G\,d\varphi^2,$$

où $G = \rho^2$. Les équations des géodésiques

$$\frac{d^2w_k}{ds^2} = -\sum_{i,\,j=1}^{2} \Gamma_{ij}^k \frac{dw_i}{ds}\frac{dw_j}{ds},$$

vu les valeurs des coefficients Γ_{ij}^k (*5.45a*), prennent la forme:

$$w_{ss} = \frac{1}{2} G_w w_s^2, \tag{1}$$

$$\varphi_{ss} = -\frac{G_w}{G} w_s\varphi_s = -\frac{2\rho\rho_w}{\rho^2} w_s\varphi_s = -\frac{2\rho_s\varphi_s}{\rho}. \tag{2}$$

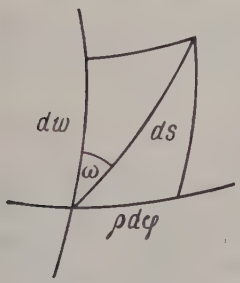


Fig. 5.4-6

Désignons par ω l'angle que fait la ligne géodésique avec la méridienne. En considérant le triangle infiniment petit avec ds pour hypoténuse et dw et $\rho\,d\varphi$ pour côtés (fig. 5.4-6), nous obtenons directement:

$$\sin\omega = \rho\varphi_s. \tag{3}$$

Théorème (de Clairaut). *Le long d'une ligne géodésique sur la surface de révolution, on a* $\rho \sin\omega = \text{const}$.

Démonstration. Le long de n'importe quelle géodésique (en vertu des équations (2) et (3)), on a:

$$(\rho\sin\omega)_s = (\rho^2\varphi_s)_s = 2\rho\rho_s\varphi_s + \rho^2\varphi_{ss} = 0,$$

ce qui démontre le théorème.

Le comportement géométrique des géodésiques peut être caractérisé, d'après le théorème de Clairaut, comme suit : chaque géodésique γ, différente de la méridienne, tourne en diminuant la valeur de ρ de façon à ce que son angle avec la méridienne croît ; si, pour un $\rho = \rho_{\min}$, cet angle devient droit conformément au théorème de Clairaut $\rho \sin \omega = c$, alors, en général, la géodésique γ devient tangente à la parallèle correspondante, puis revient dans le domaine de grandes valeurs de ρ (fig. 5.4-5). Il y a tout de même des exceptions (cf. exercices 6 et 7).

§ 5. 5. Surfaces bidimensionnelles de courbure constante

5.51. Les surfaces bidimensionnelles de courbure totale constante, $K = \text{const}$, possèdent plusieurs propriétés singulières. Avec le théorème *5.46*, on peut affirmer que *toutes les surfaces d'une même courbure totale constante K sont localement isométriques l'une à l'autre.* De plus, comme on peut choisir arbitrairement le point initial M_0 et la direction de la géodésique de base, l'isométrie peut être réalisée en fixant arbitrairement sur les surfaces P et $\widetilde{P}$ un couple de points M_0 et $\widetilde{M}_0$ correspondant l'un à l'autre et un couple de directions issues de ces points. Chaque partie suffisamment petite d'une surface de courbure totale constante peut être appliquée isométriquement sur une autre partie de la même surface, étant donnés un couple de points correspondants et un couple de directions correspondantes.

L'équation *5.45* (3) qui relie la courbure K et la fonction $G(w, u)$ dans le système semi-géodésique de coordonnées,

$$(\sqrt{G})_{ww} + K\sqrt{G} = 0,$$

avec les données initiales sur la géodésique de base *5.45* (1), (2),

$$\sqrt{G(0, u)} = 1, \quad (\sqrt{G(0, u)})_w = 0,$$

admet, pour $K = \text{const}$, les solutions :

$$K = 0: \; G(w, u) \equiv 1, \; ds^2 = dw^2 + du^2; \tag{1}$$

$$K > 0: \; \sqrt{G(w, u)} = \cos \sqrt{K} w,$$
$$ds^2 = dw^2 + \cos^2 \sqrt{K} w \, du^2; \tag{2}$$

$$K < 0: \; \sqrt{G(w, u)} = \operatorname{ch} \sqrt{-K} w,$$
$$ds^2 = dw^2 + \operatorname{ch}^2 \sqrt{-K} w \, du^2. \tag{3}$$

Le plus simple exemple de surface de courbure totale nulle est le plan ; nous voyons maintenant que *toute surface de courbure totale nulle est localement isométrique au plan.*

Le plus simple exemple de surface de courbure totale constante positive est la sphère bidimensionnelle d'un rayon R ; on a $K = 1/R^2$ (*5.24e*). Nous voyons donc que *toute surface de courbure constante positive K est localement isométrique à une sphère* (de rayon $1/\sqrt{K}$).

Un exemple de surface bidimensionnelle de courbure constante négative, i.e. d'une surface à laquelle toute surface de la même courbure négative est localement isométrique est donné dans *5.52*.

5.52. Nous donnerons ici un exemple de surface de courbure constante négative qui représente une surface de révolution. Comme nous l'avons vu dans *5.26b*, la courbure totale d'une surface de révolution de génératrice $\rho = \rho(z)$ s'écrit d'après la formule

$$K = -\frac{\rho_{zz}}{\rho(1+\rho_z^2)^2}.$$

En posant $-K = Q > 0$ constant nous avons pour $\rho(z)$ l'équation différentielle suivante du deuxième ordre:

$$\rho_{zz} = Q\rho(1+\rho_z^2)^2. \tag{1}$$

Soit $\rho_z = u(\rho)$, $\rho_{zz} = u_\rho \rho_z = u_\rho u$; alors l'équation (1) s'écrit:

$$u_\rho u = Q\rho(1+u^2)^2 \quad \text{ou} \quad \frac{u\,du}{(1+u^2)^2} = Q\rho\,d\rho.$$

En l'intégrant nous trouvons

$$-\frac{1}{2}\frac{1}{1+u^2} = \frac{1}{2}Q\rho^2 - \frac{C}{2} \quad \text{ou bien} \quad \frac{1}{1+u^2} = C - Q\rho^2.$$

Poursuivons l'intégration en cas de $C = 1$. Nous avons alors:

$$u^2 = \frac{Q\rho^2}{1-Q\rho^2}, \quad u = \frac{d\rho}{dz} = \pm\frac{\sqrt{Q}\,\rho}{\sqrt{1-Q\rho^2}}, \quad dz = \pm\frac{\sqrt{1-Q\rho^2}}{\sqrt{Q}\,\rho}\,d\rho.$$

Faisons la substitution $\sqrt{Q}\rho = \sin\theta$, $\sqrt{Q}\,d\rho = \cos\theta\,d\theta$. Nous obtenons

$$dz = \pm\frac{1}{\sqrt{Q}}\frac{\cos^2\theta}{\sin\theta}d\theta = \pm\frac{1}{\sqrt{Q}}\left[\frac{d\theta}{\sin\theta} - \sin\theta\,d\theta\right],$$

$$z - z_0 = \pm\frac{1}{\sqrt{Q}}\left(\ln\left|\operatorname{tg}\frac{\theta}{2}\right| + \cos\theta\right).$$

Comme le choix de la constante z_0 ne change pas la forme de la surface (il modifie seulement sa position par rapport à l'axe des z), on peut la choisir nulle. Deux équations

$$z = \pm\frac{1}{Q}\left(\ln\left|\operatorname{tg}\frac{\theta}{2}\right| + \cos\theta\right), \quad \rho = \frac{1}{\sqrt{Q}}\sin\theta \tag{2}$$

nous fournissent alors une représentation paramétrique de la méridienne de la surface cherchée. Cette surface même (fig. 5.5-1) est disposée dans les limites du cylindre $\rho \leqslant 1/\sqrt{Q}$ en ne s'approchant de la surface du cylindre que pour $\sin\theta = 1$, de sorte que $\cos\theta = 0$, $\ln\left|\operatorname{tg}\frac{\theta}{2}\right| = 0$, $z = 0$. En outre, si $\theta \to 0$, alors $\rho \to 0$, $z \to \pm\infty$.

La courbe (2) s'appelle *tractrice* *. On sait que la partie de sa tangente limitée par l'axe des z conserve la longueur constante.

* Figurons-nous que les axes des z et des x de la fig. 5.5-1 soient tracés sur le sol et qu'un tracteur marche le long de l'axe des z en partant du point O. Alors une voiture qu'il remorque et qui se trouve au départ au point A décrira une tractrice (fig. 5.5-2).

La surface de révolution que nous obtenons ainsi s'appelle *pseudo-sphère*. Sur cette surface (dans ses petites parties au moins) est réalisée la géométrie de Lobatchevski (les droites de la géométrie de Lobatchevski sont les géodésiques sur la pseudo-sphère). La pseudo-sphère a des singularités pour $z = 0$. Il s'avère (théorème de Hilbert) que, dans l'espace tridimensionnel, il n'existe pas de surface de courbure constante $K < 0$ sans singularité (ni sans bord). Il n'en existe pas même si l'on suppose $K \leqslant \text{const} < 0$ au lieu de $K = \text{const} < 0$ (N. V. Efimov, 1965, voir son article dans УМН, XXI, № 5 (1966), 3-58).

5.53. Donc, la famille de surfaces composée du plan, des sphères de tous les rayons R et des pseudo-sphères avec tous les Q possibles nous fournit les surfaces « canoniques » de courbure totale constante; toute surface de courbure

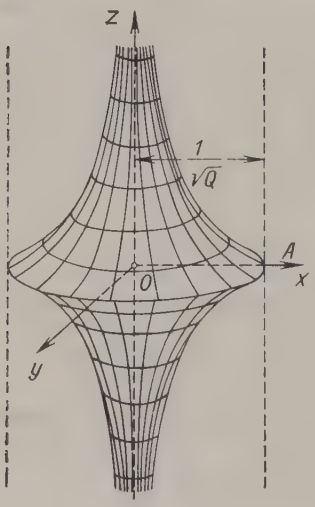


Fig. 5.5-1

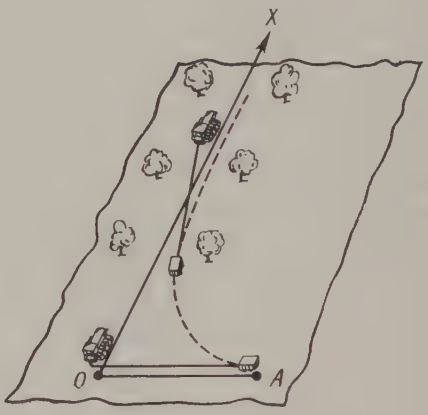


Fig. 5.5-2

totale constante est localement isométrique à la surface canonique correspondante (i.e. ayant la même courbure totale). On peut aussi bien formuler le problème de description complète des surfaces bidimensionnelles de courbure totale constante, i.e. non pas à une isométrie, mais à la position dans l'espace près. Dans le cas $K \neq 0$, c'est un problème assez compliqué et nous n'y touchons pas ici *. Mais en ce qui concerne le cas $K = 0$, la situation est beaucoup plus simple; les surfaces de courbure totale nulle dans R_3 admettent une description complète. De telles surfaces sont appelées *développables* (on sous-entend « développable sur le plan »). Nous savons que le cylindre et le cône circulaires sont des développables (*5.27c*). De plus, *tout cylindre de n'importe quelle courbe directrice* $r = r(t)$ *et de n'importe quel vecteur générateur* e *est une développable*; en effet, le vecteur peut être choisi unitaire et la courbe orthogonale à ce vecteur, avec la longueur d'arc pour paramètre; alors la première forme quadratique du cylindre coïncidera avec celle du plan en coordonnées cartésiennes. *Tout cône de sommet O et de directrice arbitraire est lui aussi une développable*; en effet, le vecteur générateur e du cône qui a le sommet pour origine peut être choisi

* Pour les surfaces de révolution de courbure constante dans R_3 voir В. Ф. Каган, Основы теории поверхностей, т. 2, ГИТТЛ, 1948.

unitaire, on peut le considérer comme fonction de l'arc σ que son extrémité décrit sur le cône; alors l'expression analytique du cône sera $r = r(\sigma, t) = te(\sigma)$, avec $ds^2 = dt^2 + t^2 d\sigma^2$ pour première forme quadratique. Ainsi, toutes les surfaces coniques se trouvent isométriques l'une à l'autre, donc, en particulier, au plan (qui est, évidemment, une surface conique).

Encore un type de développables est construit comme suit. Considérons une courbe gauche quelconque L et la surface P_2 formée par toutes les tangentes à L; nous affirmons que *la surface P_2 a elle aussi la courbure totale nulle.*

En effet, soit $r = r(\sigma)$ une équation de la courbe L avec le paramètre naturel σ; alors la surface P_2 sera donnée par la représentation paramétrique $\rho(\sigma, t) = r(\sigma) + tr_\sigma$. Construisons la première forme quadratique de la surface P_2. Nous avons $\rho_\sigma = r_\sigma + tr_{\sigma\sigma} = \tau + tk(\sigma)\nu$, $\rho_t = r_\sigma = \tau$; d'où $E = (r_\sigma, r_\sigma) = 1 + t^2k^2(\sigma)$, $F = (\rho_\sigma, \rho_\tau) = 1$, $G = 1$. Nous voyons que la première forme quadratique de P_2 ne dépend que de la courbure $k(\sigma)$ de la ligne L. Or, il existe une courbe plane $\widetilde{L}$ avec la même courbure $k(\sigma)$ en tant que fonction de l'arc σ (*U16.42*). Il lui correspond la surface formée des tangentes qui est plane et en même temps isométrique à la surface P_2 puisque possède la même première forme quadratique; par conséquent, P_2 est isométrique au plan, donc, sa courbure totale est nulle.

Notons que ce raisonnement reste valable si la courbe L se trouve dans n'importe quel espace euclidien (et même hilbertien).

Montrons que, dans R_3, la description donnée épuise toutes les surfaces de courbure totale nulle.

Théorème. *Toute surface P de courbure totale nulle se trouvant dans R_3 est un cylindre, ou un cône, ou la surface des tangentes à une courbe gauche.*

Démonstration. Comme notre surface P se trouve dans R_3, il est possible d'introduire sur n'importe quelle sa partie un réseau de coordonnées formé des lignes de courbure (*5.29a*). Par hypothèse, le produit k_1k_2 de deux courbures principales est nulle partout sur P. Si, sur la partie en question, les deux courbures principales k_1 et k_2 sont nulles, alors chaque section normale est de courbure nulle, donc chaque réseau orthogonal est composé par des lignes de courbure. Et si, en un point, l'une des courbures n'est pas nulle, elle le reste dans un voisinage de ce point, auquel cas le réseau des lignes de courbure se détermine d'une façon univoque (*5.29a*).

Supposons que le paramètre u varie le long des lignes de courbure répondant aux directions principales avec la courbure $k_1 = 0$, et le paramètre v varie le long des lignes orthogonales. La première forme quadratique en coordonnées u, v est $E\,du^2 + G\,dv^2$, la deuxième $L\,du^2 + N\,dv^2$. Comme les courbures principales $k_1 = 0$ et k_2 sont des racines de l'équation $(L - \mu E)(N - \mu G) = 0$, nous avons $L = k_1E = 0$, d'où $(m_u, r_u) = -(m, r_{uu}) = -L = 0$, $(m_u, r_v) = -(m, r_{uv}) = -M = 0$, donc $m_u = 0$. Ainsi, le vecteur m reste constant sur chaque ligne $v = \text{const}$. Alors le plan tangent à la surface P a une même orientation le long de la ligne $v = \text{const}$; par conséquent, il est fixe, car les relations $(m, r)_u = (m, r_u) = 0$ impliquent $(m, r) = \text{const} = (m, r_0)$ (r_0 est un vecteur initial fixe), ce qui signifie que $r - r_0$ se trouve dans le plan tangent Π_0 passant par le point r_0. Donc, chaque ligne $v = \text{const}$ est plane. Montrons que c'est une droite. Comme le vecteur m ne varie pas sur la ligne $v = \text{const}$, i.e. $m_u = 0$, on a $(m_v)_u = (m_u)_v = 0$, et le vecteur m_v est lui aussi constant sur la ligne $v = \text{const}$. Ce vecteur m_v se trouve dans le plan tangent, i.e. dans le plan Π_0, étant orthogonal à la ligne $v = \text{const}$, puisque $(m_v, r_u) = (m, r_u)_v - (m, r_{uv}) = 0 - M = 0$. Nous voyons que la normale à la ligne $v = \text{const}$ conserve sa direction dans le plan Π_0; or, ce n'est possible que si la ligne $v = \text{const}$ est une droite.

Nous concluons que toute la surface P est composée par les lignes droites qui représentent les lignes de coordonnées $v = \text{const}$. De telles surfaces sont dites *réglées*. On ne peut pas affirmer que toute surface réglée a la courbure nulle (par exemple, un hélicoïde est une surface réglée de courbure totale négative,

5.27*e*). Utilisons encore une fois le fait que les lignes de coordonnées $u = \text{const}$, $v = \text{const}$ déterminent en chaque point les directions principales. Fixons un point B sur la surface P. Soient $\rho = \rho(v) = r(u, v)$ l'équation de la ligne de coordonnées l ($u = \text{const}$) qui passe par B orthogonalement à la droite de coordonnées $v = \text{const}$, et $e = e(v)$ le vecteur unitaire qui en tout point de la ligne l a le même sens que le vecteur $r_u = r_u(v)$. Il est clair que $e(v) = \alpha(v) r_u(v)$. Par conséquent, $e_v = \alpha r_{uv} + \alpha_v r_u$, $(e_v, m) = \alpha(r_{uv}, m) + \alpha_v(r_u, m) = 0$, et le vecteur e_v se trouve dans le plan tangent à la surface P.

Si $e_v = 0$, alors $e(v) \equiv e$ est un vecteur constant, et toute la surface P est le cylindre de directrice $\rho = \rho(v)$ et de vecteur générateur e. Si $e_v \neq 0$, le vecteur e_v se trouve dans le plan tangent à la surface P et admet la décomposition

$$e_v = \lambda(v)\, e(v) + \mu(v)\, \rho_u(v).$$

Ici le coefficient $\mu(v)$ n'est pas identiquement nul. Sinon le vecteur e_v serait colinéaire au vecteur e, et, vu que la dérivée d'un vecteur unitaire lui est orthogonale, cela signifierait que $e_v \equiv 0$ en contradiction avec l'hypothèse. Donc, nous avons $\mu(v) \neq 0$. Fixons une fonction arbitraire $g(v)$ (dérivable) et considérons la courbe $\beta = \{ q(v) = \rho(v) + g(v)\, e(v) \}$ se trouvant sur la surface P. Nous avons

$$q_v = \rho_v + g_v e + g e_v = (1 + g\mu)\, \rho_v + (g_v + \lambda g)\, e. \qquad (1)$$

Posons $g(v) = -\dfrac{1}{\mu(v)}$. Alors le premier terme dans le second membre de (1) s'annule. S'il en est de même du deuxième terme, alors le vecteur $q(v)$ est constant, i.e. $q(v) = q_0$. Dans ce cas, chaque génératrice $\rho(v) + te(v)$ $(-\infty < t < \infty)$ passe par le point fixe q_0 pour $t = g(v) = -\dfrac{1}{\mu(v)}$; c'est un cône. Il reste le cas où $g_v + \lambda g \neq 0$. Alors la courbe β, pour notre choix de la fonction $g(v)$, est tangente en tout point à la génératrice mentionnée. Ceci veut dire que chaque droite $v = \text{const}$ sur la surface P est tangente à la ligne β, autrement dit que P représente la surface des tangentes à la courbe β. Le théorème est démontré.

Notons que déjà dans R_4, il existe des surfaces bidimensionnelles isométriques au plan, mais n'appartenant pas aux types considérés (cf. exercice 14).

5.54. Signalons encore un exemple de surface bidimensionnelle de courbure constante négative. Contrairement à tous les exemples précédents, la métrique sur cette surface n'est pas empruntée à l'espace euclidien contenant la surface.

a. Examinons dans l'espace tridimensionnel euclidien R_3 la forme quadratique indéfinie

$$\langle r, r\rangle = x_1^2 + x_2^2 - x_3^2 \qquad (r = x_1 e_1 + x_2 e_2 + x_3 e_3)$$

qui correspond à la forme bilinéaire symétrique

$$\langle r^{(1)}, r^{(2)}\rangle = x_1^{(1)} x_1^{(2)} + x_2^{(1)} x_2^{(2)} - x_3^{(1)} x_3^{(2)}.$$

Un vecteur r est dit *pseudo-réel* si $\langle r, r\rangle > 0$, *pseudo-imaginaire* si $\langle r, r\rangle < 0$, et *isotrope* si $\langle r, r\rangle = 0$. Un vecteur pseudo-réel (pseudo-imaginaire, isotrope) reste pseudo-réel (pseudo-imaginaire, isotrope) en le multipliant par n'importe quel nombre réel ($\neq 0$). Un vecteur pseudo-réel r est dit *normalisé* si $\langle r, r\rangle = 1$; un vecteur pseudo-imaginaire r est dit *normalisé* si $\langle r, r\rangle = -1$. Tout vecteur pseudo-

réel (pseudo-imaginaire) r admet une normalisation moyennant la multiplication par un nombre positif.

Deux vecteurs $r^{(1)}$, $r^{(2)}$ sont dits *pseudo-orthogonaux* si $\langle r^{(1)}, r^{(2)}\rangle = 0$; en particulier, un vecteur isotrope est pseudo-orthogonal à lui-même. Si un vecteur non isotrope r est pseudo-orthogonal à certains vecteurs p, q, . . ., il n'en dépend pas linéairement, car la relation $r = C_1 p + C_2 q + \ldots$ implique $\langle r, r\rangle = C_1\langle r, p\rangle + C_2\langle r, q\rangle + \ldots = 0$.

Si un vecteur $r = r(t)$ dépend (d'une façon dérivable) d'un paramètre t, alors le vecteur $dr = r_t\, dt$ se trouve sur la droite tangente à la courbe correspondante (au point de dérivation). Si un vecteur dépend de deux paramètres u, v et détermine une surface P, alors le vecteur $dr = r_u\, du + r_v\, dv$ se trouve dans le plan tangent à la surface P. Si deux vecteurs $r^{(1)}$ et $r^{(2)}$ sont des fonctions dérivables de certains paramètres, alors

$$d\langle r^{(1)}, r^{(2)}\rangle = \langle r^{(1)}, dr^{(2)}\rangle + \langle dr^{(1)}, r^{(2)}\rangle,$$

en particulier,

$$d\langle r, r\rangle = 2\langle r, dr\rangle.$$

La surface

$$\langle r, r\rangle \equiv x_1^2 + x_2^2 - x_3^2 = -Q^2 \quad (Q > 0) \tag{1}$$

qu'il conviendrait d'appeler « *sphère pseudo-imaginaire* » et que nous appellerons « *pseudo-sphère* » tout court représente un hyperboloïde à deux nappes dont l'une est dans le demi-espace $x_3 \geqslant Q$, l'autre, dans le demi-espace $x_3 \leqslant -Q$. Nous n'en considérons que la supérieure en la désignant par P. En dérivant l'équation $\langle r, r\rangle = -Q^2$ nous trouvons $\langle r, dr\rangle = 0$. Cela signifie que le vecteur r est pseudo-orthogonal au plan tangent à la surface P et passant par son extrémité. On peut dire que le vecteur r est *pseudo-normal* à P. Le vecteur $m = r/Q$ est pseudo-normal normalisé car $\langle m, m\rangle = \frac{1}{Q^2}\langle r, r\rangle = -1$.

Nous affirmons que le *vecteur dr est pseudo-réel*, i.e. que $\langle dr, dr\rangle > 0$. En effet, on a $x_1\, dx_1 + x_2\, dx_2 - x_3\, dx_3 = \langle r, dr\rangle = 0$; ensuite:

$$(x_3\, dx_3)^2 = (x_1\, dx_1 + x_2\, dx_2)^2 \leqslant (x_1^2 + x_2^2)(dx_1^2 + dx_2^2) = \\ = (x_3^2 - Q^2)(dx_1^2 + dx_2^2) \leqslant x_3^2 (dx_1^2 + dx_2^2),$$

d'où

$$dx_3^2 \leqslant dx_1^2 + dx_2^2, \quad \langle dr, dr\rangle = dx_1^2 + dx_2^2 - dx_3^2 \geqslant 0,$$

ce qu'il fallait.

Ainsi, la *forme quadratique* $\langle dr, dr\rangle$ *sur la surface* P (plus exactement, dans le plan tangent à P en tout point fixe) *est définie positive*. Nous la choisissons pour *forme métrique de la surface* P. Ceci veut

dire qu'en introduisant arbitrairement sur la surface P les paramètres u_1 et u_2 nous posons

$$ds^2 = g_{11}\,du_1^2 + 2g_{12}\,du_1\,du_2 + g_{22}\,du_2^2,$$

où

$$g_{ij} = \langle r_i,\ r_j\rangle \quad (i,\ j = 1,\ 2).$$

b. Soit maintenant P n'importe quelle surface pour laquelle la forme $\langle r,\ r\rangle$ induise la forme positive ds^2. En tout point de la surface P considérons les vecteurs de base r_1, r_2 dans le plan tangent et le vecteur unitaire pseudo-normal m. Comme dans *5.31*, on peut écrire les formules de dérivation

$$\left.\begin{aligned} r_{ij} &= \sum_{k=1}^{2} \widetilde{\Gamma}_{ij}^k r_k + \widetilde{\beta}_{ij} m, \\ m_j &= \sum_{k=1}^{2} \widetilde{\beta}_j^k r_k, \end{aligned}\right\} \tag{2}$$

où le symbole $\sim$ montre que le coefficient correspondant est calculé pour la surface dont la métrique est induite par la forme $\langle r,\ r\rangle$. En pseudomultipliant ces égalités par r_s et m nous trouvons

$$\langle r_{ij},\ r_s\rangle = \sum_{k=1}^{2} \widetilde{\Gamma}_{ij}^k \langle r_k,\ r_s\rangle = \sum_{k=1}^{2} \widetilde{\Gamma}_{ij}^k g_{ks},$$

$$\langle r_{ij},\ m\rangle = -\widetilde{\beta}_{ij},$$

$$\langle m_j,\ r_s\rangle = \sum_{k=1}^{2} \widetilde{\beta}_j^k \langle r_k,\ r_s\rangle = \sum_{k=1}^{2} \widetilde{\beta}_j^k g_{ks} = -\langle m,\ r_{js}\rangle,$$

puisque le vecteur r_s se trouve dans le plan tangent et que la relation $\langle m, r_s\rangle = 0$ implique

$$0 = \langle m,\ r_s\rangle_j = \langle m_j,\ r_s\rangle + \langle m,\ r_{js}\rangle.$$

Les quantités $\langle r_{ij},\ r_s\rangle = \widetilde{\Gamma}_{ij,\ s}$ s'expriment par les coefficients de la première forme quadratique g_{ij} comme dans *5.31* (11):

$$\widetilde{\Gamma}_{ij,\,s} = \langle r_{ij},\ r_s\rangle = \frac{1}{2}\left(\frac{\partial g_{is}}{\partial u_j} + \frac{\partial g_{js}}{\partial u_i} - \frac{\partial g_{ij}}{\partial u_s}\right).$$

Il en résulte que les coefficients $\widetilde{\Gamma}_{ij}^k$ s'expriment eux aussi par les coefficients de la première forme quadratique:

$$\widetilde{\Gamma}_{ij}^k = \sum_{s=1}^{2} \widetilde{\Gamma}_{ij,\,s} g^{ks}, \quad \| g^{ks} \| = \| g_{ij} \|^{-1}.$$

Dans la formule générale pour la courbure totale « formelle » d'une surface bidimensionnelle P (*5.33* (2))

$$K = \frac{\sum_{s=1}^{2} \left[\frac{\partial \Gamma_{11}^{s}}{\partial u_2} - \frac{\partial \Gamma_{12}^{s}}{\partial u_1} + \sum_{p=1}^{2} (\Gamma_{11}^{p}\Gamma_{2p}^{s} - \Gamma_{21}^{p}\Gamma_{1p}^{s}) \right] g_{s2}}{g_{11}g_{22} - g_{12}^2}, \tag{3}$$

les coefficients Γ_{ij}^{k} sont calculés d'après g_{ij} à l'aide des formules:

$$\Gamma_{ij}^{k} = \frac{1}{2} \sum_{s=1}^{2} \left(\frac{\partial g_{is}}{\partial u_j} - \frac{\partial g_{js}}{\partial u_i} - \frac{\partial g_{ij}}{\partial u_s} \right) g^{hs}.$$

Nous voyons que, dans notre cas, on peut poser $\Gamma_{ij}^{k} = \widetilde{\Gamma}_{ij}^{k}$ dans la formule (3) pour calculer K. Rappelons à présent que, pour une surface euclidienne, le numérateur dans (3) s'obtient en égalant les expressions r_{ikj} et r_{jki}, utilisant les formules de dérivation et séparant la composante suivant les vecteurs de base dans le plan tangent. Toutes ces opérations s'effectuent sans modification pour les formules de dérivation (2), après quoi on obtient, tout comme dans *5.32* (7),

$$\widetilde{\beta}_{jk}\widetilde{c}_{il} - \widetilde{\beta}_{ik}\widetilde{c}_{jl} = \sum_{s=1}^{2} \left[\frac{\partial \Gamma_{ik}^{s}}{\partial u_j} - \frac{\partial \Gamma_{jk}^{s}}{\partial u_i} + \sum_{p=1}^{2} (\Gamma_{ik}^{p}\Gamma_{jp}^{s} - \Gamma_{jk}^{p}\Gamma_{ip}^{s}) \right] g_{ls},$$

où

$$\widetilde{\beta}_{ij} = -\langle r_{ij}, \quad m \rangle, \quad \widetilde{c}_{ij} = \sum_{k=1}^{2} b_i^k g_{jk}.$$

Désignons encore $\widetilde{b}_{ij} = \langle r_{ij}, m \rangle$. Alors on a:

$$K = \frac{\widetilde{b}_{12}^2 - \widetilde{b}_{11}\widetilde{b}_{22}}{g_{11}g_{22} - g_{12}^2}.$$

(Rappelons que, dans le cas euclidien, nous avions $\widetilde{\beta}_{ij} = \langle r_{ij}, m \rangle$ à la place de $\widetilde{\beta}_{ij} = -\langle r_{ij}, m \rangle$, ce qui conduisait à l'expression $b_{11}b_{12} - b_{12}^2$ dans le numérateur du résultat).

c. Jusqu'à ce moment, le calcul est valable pour n'importe quelle surface bidimensionnelle sur laquelle la forme $\langle r, r \rangle$ induit la forme positive ds^2.

Calculons maintenant les coefficients $\widetilde{b}_{ij}$ pour la pseudo-sphère. Dans ce cas on a $r = Qm$ et

$$r_j = Qm_j = Q \sum \widetilde{b}_j^k r_k,$$

de sorte que

$$\widetilde{b}_j^k = \begin{cases} 1/Q & \text{pour} \quad j = k, \\ 0 & \text{pour} \quad j \neq k. \end{cases}$$

Il en résulte

$$\widetilde{b}_{ij} = \langle r_{ij}, \quad m\rangle = -\sum_{k=1}^{2} \widetilde{b}_j^k g_{ik} = -\frac{1}{Q} g_{ij}.$$

(Rappelons que, pour une sphère de rayon R, dans tout système de coordonnées, nous avions $b_{ij} = \frac{1}{R} g_{ij}$ (*5.24d*).)

Définitivement,

$$K = \frac{1}{Q^2} \frac{g_{12}^2 - g_{11}g_{22}}{g_{11}g_{22} - g_{12}^2} = -\frac{1}{Q^2},$$

et la surface P avec la métrique $\langle dr, dr\rangle$ est effectivement une surface de courbure totale constante négative.

§ 5. 6. Translation des vecteurs et théorème de Levi-Civita

5.61. Translation d'un vecteur sur une surface. En translatant un vecteur tangent à une surface dans l'espace, ce vecteur cesse, en général, d'être tangent. Dans la géométrie différentielle, on définit une autre notion de translation, non pas relativement à l'espace contenant une surface, mais relativement à cette surface même.

Supposons que sur une surface $P_n \subset R_{n+1}$ soient donnés une ligne dérivable $L = \{r = r(u),\ u = (u_1, \ldots, u_n) = u(t),\ a \leqslant t \leqslant b\}$ et, en tout point de cette ligne, un vecteur $a(t)$ dans le plan tangent Π_n, ce vecteur étant dérivable en t, de sorte que $a(t) = \sum_{i=1}^{n} a_i(t)\, r_i(u(t))$, où $a_i(t)$ sont des fonctions dérivables. Trouvons la différentielle du vecteur $a(t)$ correspondant au passage au point $t + dt$. Nous avons $da = \sum_{k=1}^{n} r_k\, da_k + \sum_{i,j=1}^{n} a_i r_{ij}\, du_j$ et, d'après la formule de Gauss *5.31* (1):

$$da = \sum_{k=1}^{n} r_k\, da_k + \sum_{i,j=1}^{n} \Big(\sum_{k=1}^{n} \Gamma_{ij}^k r_k + b_{ij} m\Big) a_i\, du_j =$$

$$= \sum_{k=1}^{n} \sum_{i,j=1}^{n} (\Gamma_{ij}^k a_i\, du_j + da_k)\, r_k + \Big(\sum_{i,j=1}^{n} b_{ij} a_i\, du_j\Big) m. \quad (1)$$

Le premier terme de cette somme s'appelle *différentielle géodésique* du vecteur $a(t)$ et se désigne par Da. Le second terme s'appelle *différentielle forcée* du vecteur $a(t)$. Les composantes de la différentielle géodésique, comme on le voit de leur structure, tout en dépendant de a_i, du_i, da_i, ne dépendent plus que des symboles Γ_{ij}^k, donc

ne varient pas lors d'une isométrie de la surface P_n. La composante suivant la normale, i.e. la différentielle forcée, dépend des coefficients de la deuxième forme quadratique et, en général, varie avec une isométrie de la surface.

5.62.a. Nous dirons que le vecteur $a\ (t)$ subit une *translation*, ou bien, plus exactement, est *géodésiquement translaté* le long de la ligne L, si sa différentielle géodésique est nulle en tout point de la ligne L. En d'autres termes, le vecteur $a\ (t)$ est translaté si sa différentielle totale se réduit à une différentielle forcée. Ceci dit, il est clair que la notion de translation d'un vecteur le long d'une ligne appartient à la géométrie intrinsèque de la surface et ne dépend pas de transformation isométrique. Les composantes d'un vecteur translaté, comme on le voit de l'égalité (1) et de la définition de translation, satisfont au système suivant d'équations différentielles:

$$da_k = -\sum_{i,\ j=1}^{n} \Gamma_{ij}^k a_i\, du_j, \quad k=1,\ \ldots,\ n,$$

soit

$$\frac{da_k}{dt} = -\sum_{i,\ j=1}^{n} \Gamma_{ij}^k a_i\ (t)\ \frac{du_j}{dt}, \quad k=1,\ \ldots,\ n.$$

Ce système représente un système d'équations différentielles ordinaires par rapport aux fonctions $a_k\ (t)$ $(k = 1, \ldots, n)$, avec les coefficients connus (le long d'un chemin donné). En vertu du théorème fondamental d'existence et d'unicité de solution d'un tel système, les données initiales $a_k\ (t_0)$ $(k = 1, \ldots, n)$ déterminent d'une façon univoque le résultat de la translation, au moins dans un voisinage du point initial.

b. Nous avons donné la définition d'une translation pour le cas de surface n-dimensionnelle P_n dans l'espace R_{n+1}; à présent, vu que la définition est formulée en termes de géométrie intrinsèque, elle peut être répétée pour une surface P_n dont la métrique est donnée par une forme quelconque ds^2; par exemple, si la surface P_n se trouve dans un espace euclidien R_q, $q > n$, sa métrique peut être empruntée à cet espace (comme dans *5.16*), et la translation se définie de la façon naturelle.

c. Soient $L \subset P_n$ une ligne géodésique et a le vecteur unitaire tangent à la ligne L. Si le rayon vecteur r de la ligne L est donné en fonction du paramètre naturel s, i.e. de la longueur d'arc, $r = r\ (s)$, on peut noter $a = a\ (s) = \frac{dr}{ds} = \sum r_k \frac{du_k}{ds}$, de sorte que $a_k\ (s) = \frac{du_k(s)}{ds}$. En portant cette expression dans l'équation de la géodé-

sique 5.42 (2)

$$\frac{da_k(s)}{ds} = \frac{d^2u_k(s)}{ds^2} = -\sum_{i,\,j=1}^{n} \Gamma_{ij}^k \frac{du_i}{ds}\frac{du_j}{ds} = -\sum_{i,\,j=1}^{n} \Gamma_{ij}^k a_i \frac{du_j}{ds},$$

nous voyons que le vecteur $a(s)$ satisfait à la condition de translation. Donc, le *vecteur unitaire tangent à une ligne géodésique L est translaté le long de cette ligne.*

d. En général, si deux vecteurs $a(t)$ et $b(t)$ sont translatés le long d'une ligne $L \subset P_n$, *leur produit scalaire reste constant.* Ceci résulte du fait que $d(a, b) = (da, b) + (a, db) = 0$, puisque les

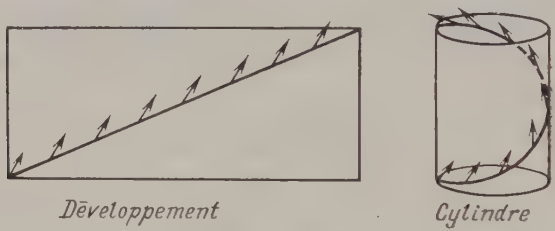


Fig. 5.6-1

différentielles des vecteurs subissant une translation se réduisent aux différentielles forcées et sont normales au plan tangent.

Il en découle qu'une translation conserve la longueur d'un vecteur $|a(t)| = \sqrt{(a(t), a(t))}$ et l'angle formé par deux vecteurs qui vaut $\arccos \frac{(a(t), b(t))}{|a(t)||b(t)|}$.

e. Sur le plan n-dimensionnel P_n, les différentielles forcées font défaut, et la différentielle totale d'un vecteur $a(t)$ se réduit à sa différentielle géodésique ; si cette dernière est nulle, il en est donc de même de la différentielle totale, ce qui veut dire que le vecteur $a(t)$ ne varie pas lors d'une translation au sens classique du mot.

f. Sur un cylindre et sur un cône dans R_3, la translation peut être réalisée en développant ces surfaces sur le plan et en faisant la translation sur ce plan. Ainsi, sur un cylindre, un vecteur translaté $a(t)$ conserve l'angle qu'il fait avec chaque génératrice ou parallèle (fig. 5.6-1). Sur un cône, nous heurtons une circonstance inattendue : un vecteur, dirigé au début suivant une génératrice et translaté le long de la circonférence directrice, revient au point initial en faisant un angle avec la génératrice ce qu'on peut voir directement en examinant le développement du cône sur le plan (fig. 5.6-2).

g. Un phénomène analogue se produit en translatant un vecteur le long d'une parallèle sur la sphère. Bien qu'il soit impossible de développer la sphère sur le plan, nous sommes en mesure de construire le cône tangent à la sphère le long de la parallèle, puisque la sphère

et ce cône possèdent le même ensemble des plans tangents aux points de la parallèle, et une translation est complètement déterminée par les plans tangents le long du chemin de translation. Le résultat de la translation sera donc le même pour la sphère et pour le cône (fig. 5.6-3).

5.63. Théorème de Levi-Civita. Les exemples précédents nous montrent qu'en translatant un vecteur le long d'un contour fermé, il ne revient en général pas à sa position initiale, mais

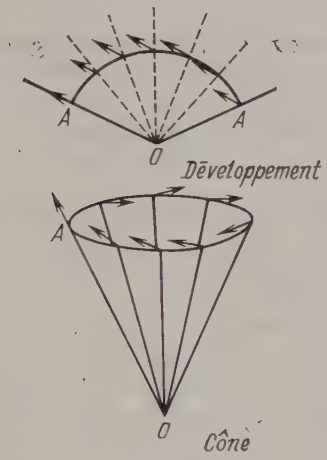


Fig. 5.6-2

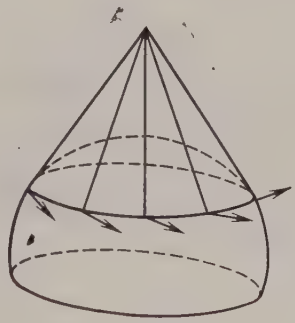


Fig. 5.6-3

se trouve tourné par rapport à cette dernière d'un certain angle. Nous tenons à calculer la valeur de cet angle dans le cas de surface bidimensionnelle.

Introduisons dans la partie considérée d'une surface $P_2 \subset R_3$ un système semi-géodésique de coordonnées w, u avec la première forme quadratique (*5.45a*)

$$ds^2 = dw^2 + G(w; u)\, du^2.$$

Soit L un contour fermé et lisse par morceaux le long duquel nous allons translater le vecteur unitaire ρ dans le sens positif (i.e. du vecteur $r_1 \equiv r_w$ au vecteur $r_2 \equiv r_u$, ou bien, ce qui est le même, de la direction positive de la ligne $u = \text{const}$ vers la direction positive de la ligne $w = \text{const}$). Désignons par S le domaine découpé sur la surface P_2 par le contour L, et par ω l'angle des vecteurs ρ et r_w compté à partir du vecteur r_w dans le sens positif (fig. 5.6-4). On s'intéresse à la quantité

$$\Delta\omega = \oint_L d\omega.$$

Comme $(\rho, r_w) = \cos\omega$, nous avons $-\sin\omega\, d\omega = (d\rho, r_w) + (\rho, dr_w) = (\rho, Dr_w)$, puisque seule la différentielle forcée du vecteur translaté ρ peut être non nulle, et pour le vecteur r_w c'est seule la différentielle géodésique qui compte. On peut calculer cette dernière d'après la formule (cf. *5.61*)

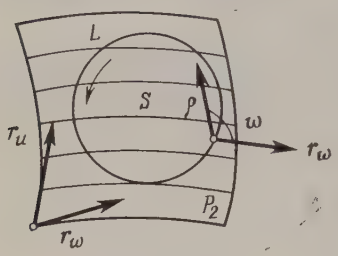


Fig. 5.6-4

$$Da = \sum_{k=1}^{2} \Big(da_k + \sum_{i,\,j=1}^{2} \Gamma_{ij}^k a_i\, du_j\Big) r_k.$$

Dans le présent cas, $a = r_w$, de sorte que $a_1 \equiv 1$, $a_2 \equiv 0$, donc

$$Dr_w = \sum_{k=1}^{2} \sum_{j=1}^{2} \Gamma_{1j}^k\, du_j \cdot r_k.$$

Ensuite, dans un système semi-géodésique (*5.45a*), $\Gamma_{11}^1 = \Gamma_{12}^1 = \Gamma_{11}^2 = 0$, $\Gamma_{12}^2 = \frac{1}{2}\frac{G_w}{G}$; définitivement, on a

$$Dr_w = \frac{1}{2}\frac{G_w}{G}\, du \cdot r_u.$$

Ainsi,

$$-\sin\omega\, d\omega = \Big(\rho,\ \frac{1}{2}\frac{G_w}{G}\, du \cdot r_u\Big) = \frac{1}{2}\frac{G_w}{G}\, du \cdot \sin\omega \cdot \sqrt{G},$$

d'où

$$d\omega = -\frac{1}{2}\frac{G_w}{\sqrt{G}}\, du$$

et

$$\Delta\omega = \oint_L d\omega = -\frac{1}{2}\oint_L \frac{G_w}{\sqrt{G}}\, du.$$

Transformons l'expression obtenue d'après la formule de Green *4.16* (3) ; en nous rappelant ensuite que l'élément dS de la surface P_2 est calculé d'après la formule $dS = \sqrt{EG - F^2}\, dw\, du = \sqrt{G}\, dw\, du$; nous avons définitivement :

$$\Delta\omega = \oint_L d\omega = -\frac{1}{2}\oint_L \frac{G_w}{\sqrt{G}}\, du = -\frac{1}{2}\iint_S \Big(\frac{G_w}{\sqrt{G}}\Big)_w dw\, du =$$

$$= -\frac{1}{2}\iint_S \Big(\frac{G_{ww}}{\sqrt{G}} - \frac{1}{2}\frac{G_w^2}{\sqrt{G^3}}\Big)\frac{dS}{\sqrt{G}} =$$

$$= -\frac{1}{2}\iint_S \Big(\frac{G_{ww}}{G} - \frac{1}{2}\frac{G_w^2}{G^2}\Big)\, dS = \iint_S K\, dS,$$

en vertu de la formule de la courbure totale dans un système semi-géodésique de coordonnées *5.45* (1). Nous aboutissons au théorème suivant :

T h é o r è m e (de Levi-Civita). *A la suite de la translation le long d'un contour fermé L (suffisamment petit), chaque vecteur subit la rotation d'un angle qui vaut l'intégrale de la courbure totale de la surface sur le domaine S limité par le contour L.*

5.64. Le contour L peut avoir des points anguleux. Considérons un contour L formé des trois lignes géodésiques faisant les angles α, β et γ aux points d'intersection respectifs A, B et C (fig. 5.6-5). En plus d'un vecteur unitaire ρ translaté le long du contour L, considérons un autre vecteur unitaire e qui se déplace d'après la règle suivante : au moment initial les vecteurs ρ et e sont au sommet A, dirigés le long du côté AB ; ensuite ils sont translatés le long de AB jusqu'au sommet B ; ici le vecteur e subit une rotation dans le sens positif d'un angle $\pi - \beta$ et devient tangent au côté BC ; puis les vecteurs e et ρ sont translatés le long de la géodésique BC jusqu'au sommet C où le vecteur e tourne de l'angle $\pi - \gamma$ et devient tangent au côté CA ; enfin, les vecteurs ρ, e sont translatés le long de la géodésique CA jusqu'au sommet A, où le vecteur e tourne de l'angle $\pi - \alpha$ en retrouvant sa position initiale. Comme une translation ne change pas l'angle de deux vecteurs (*5.62d*), l'écart définitif des vecteurs ρ et e sera dû à seules les rotations du vecteur e aux sommets A, B, C et fera la quantité $3\pi - (\alpha + \beta + \gamma)$. En somme avec la rotation $\Delta\omega$ du vecteur ρ, nous aurons l'écart du vecteur e de sa position initiale de $\Delta\omega + 3\pi - (\alpha + \beta + \gamma)$. Or, on a vu que le vecteur e revient à sa position initiale, et sa rotation totale fait 2π. Nous obtenons l'équation

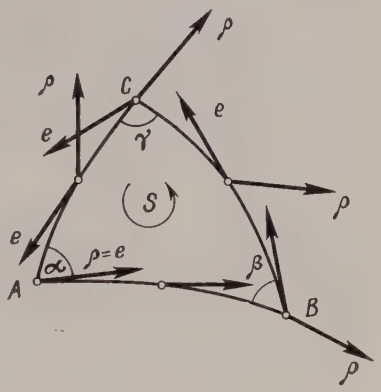


Fig. 5.6-5

$$2\pi = \Delta\omega + 3\pi - (\alpha + \beta + \gamma),$$

d'où

$$\alpha + \beta + \gamma = \pi + \Delta\omega = \pi + \iint_S K\, dS.$$

Nous avons ainsi démontré le t h é o r è m e d e G a u s s : *la somme des angles d'un triangle géodésique ABC vaut deux angles*

droits à une correction près, cette dernière étant positive sur une surface de courbure positive et négative sur une surface de courbure négative ; sur une surface de courbure constante cette correction est proportionnelle à l'aire limitée par le triangle géodésique.

Exercices

1. Ecrire l'expression de la courbure totale pour la surface $z = z(x, y)$.

2. Même question pour la surface donnée par une équation implicite $F(x, y, z) = 0$.

3. Même question pour la surface avec la forme métrique $ds^2 = \varphi(u, v) \times \times (du^2 + dv^2)$.

4. Il y a une famille de surfaces de révolution obtenue en translatant l'une d'elles le long de son axe. On construit une nouvelle surface de révolution du même axe qui soit orthogonale aux surfaces de la famille en question. Démontrer que la courbure de Gauss K de la nouvelle surface vérifie la relation $K = -\widetilde{K}$, où $\widetilde{K}$ est la courbure de Gauss de la surface de la famille passant par le même point.

5. Trouver la courbure géodésique des méridiennes et des parallèles sur une surface de révolution.

6. La ligne de striction sur une caténoïde en est une géodésique. Comment est la géodésique γ qui passe par un point A (ρ, φ) n'appartenant pas à la ligne de striction et y fait avec la méridienne un tel angle ω que $\rho \sin \omega = \rho_0$, où ρ_0 est le rayon de la ligne de striction? Tenir compte des faits que la géodésique ne peut être tangente à la ligne de striction d'après le théorème d'unicité pour les géodésiques, ni orthogonale à cette ligne d'après le théorème de Clairaut ; le dernier théorème interdit même à la géodésique d'être tangente à une autre parallèle de la caténoïde.

7. Décrire les parallèles sur une surface de révolution arbitraire au voisinage desquelles le comportement des géodésiques est analogue à celui de l'exercice 6.

8. Soit L une surface réglée $r(u, v) = R(u) + vl(u)$ (ou, ce qui est le même, une famille uniparamétrique de droites $\lambda(u)$. Pour deux droites proches $\lambda(u)$, $\lambda(u + \Delta u)$ de cette famille, trouvons sur $\lambda(u)$ un point $M(u, \Delta u)$ en lequel le minimum des distances à la droite $\lambda(u + \Delta u)$ est réalisé (nous supposons que parmi les droites $\lambda(u)$ il n'y ait pas de parallèles). Montrer que, lorsque $\Delta u \to 0$, le point $M(u, \Delta u)$ tend vers une position limite $M(u)$ (« centre » de la droite $\lambda(u)$). Le lieu géométrique des centres forme la *ligne de striction de la surface L*.

9. La ligne de striction d'un hyperboloïde à une nappe coupe ses génératrices sous un angle aigu. Montrer que, si la ligne de striction d'une surface réglée est orthogonale aux génératrices, alors toute autre ligne orthogonale aux génératrices découpe sur elles, avec la ligne de striction, les segments égaux.

10. (Suite.) Montrer que la seule surface réglée de courbure moyenne nulle est un hélicoïde.

11. Démontrer que si l'une des courbures principales $\lambda_1, \lambda_2, \ldots$ d'une surface S dans R_n est constante, alors cette surface est l'enveloppe d'une famille de sphères d'un même rayon (ou de celle de plans) ; de plus, il n'existe aucune surface dont deux courbures principales soient constantes, non nulles et différentes (A. Kostutchenko).

12. Démontrer que, pour un point ordinaire M d'une surface $S \subset R_n$, il existe un tel voisinage U, où par chaque point A il passe une et une seule géodésique issue du point M.

13. Démontrer que, pour un point ordinaire M d'une surface $S \subset R_n$, il existe un tel voisinage U que n'importe quels ses deux points A et B soient liés par une et une seule géodésique dont tout l'arc de A à B soit entièrement

contenu dans U.

14. Montrer que la surface S dans l'espace $R_4 = R_2^{(1)} + R_2^{(2)}$:

$$R(u, v) = r(u) + \rho(v), \quad r(u) \in R_2^{(1)}, \quad \rho(v) \in R_2^{(2)},$$

où $r(u)$ et $\rho(v)$ décrivent, indépendamment l'un de l'autre, les courbes fixes dans ses plans, est isométrique au plan, mais peut ne contenir aucun segment de droite.

15. Soit $\pi_p = \{ r = \{ x_1(u), \ldots, x_n(u) \} \quad u = (u_1, \ldots, u_p) \in G \subset \subset R_p \} \subset R_n$ une surface p-dimensionnelle dans R_n. On appelle *section normale élémentaire* de la surface π_p au point $M \in \pi_p$, pour un vecteur tangent $dr = \{ dx_1, \ldots, dx_n \}$ et un vecteur normal m, la courbe d'intersection de la surface π_p avec le plan bidimensionnel engendré par les vecteurs dr et m. Montrer que la surface $\pi_2 = \{ r = \{ u_1, u_2, u_1^2, u_1^3 \} \} \in R_4$ n'admet aucune section normale élémentaire au point $(0, 0, 0, 0)$ pour le vecteur tangent $\{ 1, 0, 0, 0 \}$.

16. Etant donné la même surface π_p, le point M et le vecteur tangent dr, on appelle *section normale complète* le lieu géométrique des points d'intersection de la surface π_p avec le plan $(n - p + 1)$-dimensionnel engendré par le vecteur dr et tous les $n - p$ normales linéairement indépendantes à la surface π_p au point M. Démontrer que, en un point ordinaire de la surface π_p (i.e. au point où $\operatorname{rang} \left\| \frac{\partial x}{\partial u} \right\| = p$), pour tout vecteur tangent dr, la section normale complète est une courbe lisse sur la surface π_p.

17. Pour la même surface π_p, désignons par $m_{p+1}, \ldots, m_n$ le système orthonormé des normales au point M qui dépend des paramètres u d'une façon dérivable. Montrer que les formules de dérivation de r_i et m_i peuvent être mises sous la forme:

$$\frac{\partial r_i}{\partial u_j} = \sum_{s=1}^{p} \Gamma_{ij}^s r_s + \sum_{s=p+1}^{n} b_{ij}^{(s)} m_s, \quad i = 1, \ldots, p,$$

$$\frac{\partial m_i}{\partial u_j} = \sum_{s=1}^{p} b_j^{(i)s} r_s + \sum_{s=p+1}^{n} \gamma_j^{(i,s)} m_s, \quad i = p+1, \ldots, n,$$

où

$$b_{ij}^{(s)} = (r_{ij}, m_s), \qquad b_j^{(i)s} = -\sum_{q=1}^{n} b_{qj}^{(i)} g^{qs}.$$

18. Pour la même surface π_p, montrer que la formule de Gauss *5.32* (8) reste valable si l'on sous-entend par $B_{ij,\,kl}$ la somme des mineurs du 2-ème ordre construits sur les lignes de numéros i et j et sur les colonnes de numéros k et l, cette somme étant étendue sur toutes les matrices $\| b_{ik}^{(s)} \|$, $s = p + 1, \ldots, n$.

19. Pour la même surface π_p, vérifier, compte tenu des formules de l'exercice 17, les relations suivantes généralisant les formules de Peterson *5.32* (6), de même que *5.32* (9) et (10):

$$\sum_{s=1}^{p} \Gamma_{ij}^s b_{sk}^{(\nu)} + \frac{\partial b_{ij}^{(\nu)}}{\partial u_k} + \sum_{s=p+1}^{n} b_{ij}^{(s)} \gamma_k^{(\nu,s)} = \sum_{s=1}^{p} \Gamma_{kj}^s b_{si}^{(\nu)} + \frac{\partial b_{kj}^{(\nu)}}{\partial u_i} + \sum_{s=p+1}^{n} b_{kj}^{(s)} \gamma_i^{(\nu,s)},$$

$$\frac{\partial b_j^{s(i)}}{\partial u_k} + \sum_{\nu=1}^{p} b_j^{\nu(i)} \Gamma_{\nu k}^s + \sum_{\nu=p+1}^{n} \gamma_j^{(\nu,i)} b_k^{s(\nu)} = \frac{\partial b_k^{s(i)}}{\partial u_j} + \sum_{\nu=1}^{p} b_k^{\nu(i)} \Gamma_{\nu j}^s + \sum_{\nu=p+1}^{n} \gamma_k^{(\nu,i)} b_j^{s(\nu)},$$

$$\frac{\partial \gamma_j^{(\nu, i)}}{\partial u_k} + \sum_{s=1}^{p} b_j^{s(i)} b_{sk}^{(\nu)} + \sum_{s=p+1}^{n} \gamma_j^{(s, i)} \gamma_k^{(\nu, s)} =$$

$$= \frac{\partial \gamma_k^{(\nu, i)}}{\partial u_j} + \sum_{s=1}^{p} b_k^{s(i)} b_{sj}^{(\nu)} + \sum_{s=p+1}^{n} \gamma_k^{(s, i)} \gamma_j^{(\nu, s)}.$$

20. Montrer que la surface p-dimensionnelle $\pi_p \subset R_n$ peut être construite à un déplacement près d'après les matrices données $\| g_{ij} \|$, $\| b_{ij}^{(\nu)} \|$, $\| \gamma_i^{(\nu, s)} \|$ qui satisfassent aux relations des exercices 18 et 19, et cela de façon que

$$g_{ij} = (r_i, r_j), \quad b_{ij}^{(\nu)} = (r_{ij}, m_\nu) \quad (\nu = p + 1, \ldots, n),$$

où $r = r(u)$ est le rayon vecteur de la surface π_k, $r_i = \dfrac{\partial r(u)}{\partial u_i}$, $r_{ij} = \dfrac{\partial^2 r(u)}{\partial u_i \partial u_j}$, m_ν $(\nu = p + 1, \ldots, n)$ le système orthonormé des normales.

21. Pour un cône nous avons $K = 0$, et, d'après *5.63*, un vecteur translaté le long d'un contour fermé doit revenir à sa position initiale. Or, cela n'a pas lieu (*5.61*). Pourquoi ?

Historique

En 1697, dans une lettre à Leibniz, J. Bernoulli pose un problème que l'on peut considérer comme premier problème de la géométrie différentielle : comment sont les courbes sur une surface donnée qui réalisent le munimum de la distance (sur la surface) entre deux points donnés ? J. Bernoulli appelle ces courbes lignes géodésiques. Les équations des lignes géodésiques sur n'importe quelle surface sont écrites par Euler et Lagrange durant les années 1770. En même temps, Euler donne une formule pour la distribution de la courbure des sections normales et détermine toutes les surfaces isométriques au plan. Les lignes de courbures et les lignes asymptotiques sont introduites par Monge (1795) ; Dupin et Meusnier dont les noms sont liés à la courbure des lignes sur la surface sont des disciples de Monge. Un rôle primordial dans le développement de la théorie moderne des surfaces appartient à l'ouvrage de Gauss «Allgemeine Flächentheorie » (Leipzig, 1921, Acad. Verl.Ges.). L'auteur y définit les deux formes quadratiques fondamentales, la courbure totale (avec la transformation sphérique) et démontre le théorème sur l'invariance de cette courbure par rapport à une isométrie. De plein droit, Gauss considère ce théorème tellement important qu'il l'appelle theorema egregium (« théorème extraordinaire »). Une valeur principale a la notion, introduite par Gauss, de géométrie intrinsèque d'une surface en tant que d'ensemble de ses propriétés invariantes par rapport à une isométrie. C'est toujours Gauss qui trouve une description intrinsèque de la courbure moyennant la somme des angles du triangle géodésique. La formule de Gauss pour la dérivation de la base formée des vecteurs tangents, avec la formule du mathématicien russe Peterson (1853) pour la dérivée de la normale * (et tout cela sous la forme scalaire, car les vecteurs ne sont pas encore découverts), composent le système des équations fondamentales de la théorie des surfaces. Avec ces équations, Bonnet (1867) établit le théorème d'unicité de la surface ayant les formes quadratiques fondamentales données.

Il paraît que l'isométrie d'une caténoïde et d'un hélicoïde soit remarquée pour la première fois par Dini (1865). La pseudo-sphère est construite par Beltrami en 1872. La translation d'un vecteur sur une surface est définie en 1917

* La même formule est retrouvée plus tard par les mathématiciens italiens Codazzi et Mainardi.

par Levi-Civita qui, moyennant la translation d'un vecteur le long d'un contour fermé, exprime la courbure de la surface. Le théorème de Levi-Civita représente une généralisation du théorème de Bonnet (1867), où le triangle géodésique de Gauss est remplacé par n'importe quelle courbe fermée, et la somme des angles par l'intégrale de la courbure géodésique du contour.

On peut trouver un exposé systématique de la géométrie différentielle multidimensionnelle dans les livres: Eisenhart L., Riemannian geometry, Princeton, Univ. Press, 1949; Schouten J. und Struik D. J., Einführung in die neueren Methoden der Differentialgeometrie, 2-e vollst. umgearb. Aufl., Bd. 1, Groningen, Noordhoff, 1935.

Géométrie riemannienne

A la fin du chapitre 5, en discutant la représentation canonique d'une surface de courbure constante négative, nous avons vu que la métrisation d'une surface avec la métrique de l'espace euclidien qui la contient devenait gênante et qu'il valait mieux considérer la surface, si possible, comme objet isolé, sans le lier à un espace le contenant, et définir une métrique d'une façon intrinsèque. C'est cela l'idée d'espace riemannien. Sa définition est donnée dans § 6.3, après les connaissances nécessaires de l'algèbre tensorielle (§ 6.1) et la notion de variété différentiable élémentaire (§ 6.2). Il faut dire que la variété différentiable est l'un des plus importants objets de l'analyse mathématique moderne. Pourtant, nous ne formulons pas ici la définition générale d'une variété différentiable; en adoptant le point de vue local nous définissons une variété différentiable élémentaire (difféomorphe à une boule) et, conformément à cela, un espace riemannien élémentaire que nous étudions dans la suite. La définition générale d'une variété différentiable sera formulée dans la troisième partie. Ici notre but principal est d'introduire et d'étudier le tenseur de courbure (§ 6.5), ses relations avec la courbure d'une surface classique considérée comme espace riemannien élémentaire et, en tant que généralisation des considérations du chapitre 5, la description locale des espaces riemanniens de courbure constante (§ 6.6).

§ 6. 1. Théorie algébrique des tenseurs

6.11. Nous avons déjà parlé, dans *L5.6*, des tenseurs dans l'espace n-dimensionnel. Rappelons les définitions de base.

On peut envisager de différents systèmes de vecteurs de base dans un espace n-dimensionnel vectoriel R_n; nous les désignerons par $e_1, \ldots, e_n$; $e_{1'}, \ldots, e_{n'}$; $e_{1''}, \ldots, e_{n''}$, etc. Tout vecteur x peut être développé suivant les vecteurs de base, de sorte que

$$x = \sum_{i=1}^{n} \xi^i e_i = \sum_{i'=1}^{n} \xi^{i'} e_{i'} = \sum_{i''=1}^{n} \xi^{i''} e_{i''} = \ldots \qquad (1)$$

Les coordonnées d'un vecteur seront toujours munies d'indices supérieurs, et la présence ou l'absence des primes indiquera la base par rapport à laquelle sont définies les coordonnées. Convenons encore de *ne pas écrire* (mais de sous-entendre) le symbole de sommation dans les cas où l'indice de sommation figure sous le signe somme deux fois: comme supérieur et comme inférieur. Ainsi, les égalités

(1) prennent la forme suivante:

$$x = \xi^i e_i = \xi^{i'} e_{i'} = \xi^{i''} e_{i''} = \dots$$

Les éléments de la matrice de passage d'une base $\{e_i\}$ à une autre $\{e_{i'}\}$ seront désignés par $p^i_{i'}$, c'est-à-dire que

$$e_{i'} = p^i_{i'} e_i$$

(sommation sur l'indice i, l'indice i' prend n'importe quelle valeur fixe de 1 à n).

Les éléments de la matrice de passage inverse seront désignés par $p^{i'}_i$, c'est-à-dire que

$$e_i = p^{i'}_i e_{i'}$$

(sommation sur i', l'indice i est fixe). La matrice $\| p^{i'}_i \|$ est l'inverse de la matrice $\| p^i_{i'} \|$, ce qu'on peut noter:

$$p^i_{i'} p^{i'}_k = \begin{cases} 1 & \text{pour } i = j, \\ 0 & \text{pour } i \neq j. \end{cases} \qquad (2)$$

Dans ce qui suit, nous désignerons par δ^i_j les éléments de la matrice unité: $\delta^i_i = 1$ ($i = 1, \dots, n$), $\delta^i_j = 0$ pour $i \neq j$. Alors les égalités (2) s'écrivent

$$p^i_{i'} p^{i'}_j = \delta^i_j.$$

Le produit des matrices $p^i_{i''}$ et $p^{i'}_i$,

$$p^i_{i''} = p^{i'}_{i''} p^i_{i'},$$

est la matrice de passage de la base $\{e_i\}$ à la base $\{e_{i''}\}$.

6.12. Formulons maintenant la définition d'un tenseur.

a. Chaque tenseur est un ensemble de nombres qui dépend du système de coordonnées et qui se transforme, lorsque le système de coordonnées change, d'après une règle que nous allons décrire. Commençons par un exemple pour passer à la définition générale.

Un tenseur T deux fois covariant et une fois contravariant représente un ensemble de n^3 nombres T^k_{ij} ($i, j, k = 1, \dots, n$) qui dépendent du choix de base et qui se transforment, en passant d'une base $e_1, \dots, e_n$ à une autre base $e_{1'}, \dots, e_{n'}$, d'après la règle

$$T^{k'}_{i'j'} = p^i_{i'} p^j_{j'} p^{k'}_k T^k_{ij} \qquad (1)$$

avec la sommation dans le second membre sur i, j, k. Les expressions de la forme $p^l_{l'}$, i.e. les éléments de la matrice de passage de la base $\{e_{l'}\}$ à la base $\{e_l\}$, figurent dans le second membre deux fois, les expressions $p^{k'}_k$, i.e. les éléments de la matrice de passage de la base

$\{e_k\}$ à la base $\{e_{k'}\}$, une fois. Les indices dans la désignation du tenseur sont placés d'une façon qui estime la règle de sommation ; dans le présent cas, les indices inférieurs i, j, i', j' sont dits *covariants*, les indices supérieurs k, k' *contravariants*.

La définition (1) tient debout quelle que soit une transformation admissible des coordonnées. Soit $e_{1''}, \ldots, e_{n''}$ une troisième base ; alors, on doit avoir

$$T^{k''}_{i''j''} = p^{i}_{i''} p^{j}_{j''} p^{k''}_{k} T^{k}_{ij}, \tag{2}$$

et, d'autre part,

$$T^{k''}_{i''j''} = p^{i'}_{i''} p^{j'}_{j''} p^{k''}_{k'} T^{k'}_{i'j'}. \tag{3}$$

Or, il est aisé de voir que, vu (1), l'égalité (3) implique (2) et réciproquement. En effet, il résulte de (3) et (1) :

$$T^{k''}_{i''j''} = p^{i'}_{i''} p^{j'}_{j''} p^{k''}_{k'} p^{i}_{i'} p^{j}_{j'} p^{k'}_{k} T^{k}_{ij} = p^{i}_{i''} p^{j}_{j''} p^{k''}_{k} T^{k}_{ij}$$

en vertu de l'égalité $p^{i'}_{i''} p^{i}_{i'} = p^{i}_{i''}$ et de deux égalités analogues. La définition du tenseur T^{k}_{ij} est donc correcte.

Pour donner un tenseur T, on peut procéder, par exemple, comme suit : en fixant ses composantes T^{k}_{ij} dans une base $e_1, \ldots, e_n$ d'une façon arbitraire, les définir dans toute autre base $e_{1'}, \ldots, e_{n'}$ d'après les formules (1). Les raisonnements faits plus haut montrent que ce sera correct.

b. D'une manière analogue, pour n'importe quels entiers $m \geqslant 0$ et $s \geqslant 0$, on définit un *tenseur* $T^{j_1 \ldots j_s}_{i_1 \ldots i_m}$ (ou, en abrégé, $T^{\cdots}_{\cdots}$) *m fois covariant* et *s fois contravariant*. Précisément, on appelle ainsi un système de n^{m+s} nombres qui est défini dans chaque base et qui se transforme en passant d'une base $\{e_i\}$ à une base $\{e_{i'}\}$ d'après la formule :

$$T^{j'_1 \ldots j'_s}_{i'_1 \ldots i'_m} = p^{i_1}_{i'_1} \ldots p^{i_m}_{i'_m} p^{j'_1}_{j_1} \ldots p^{j'_s}_{j_s} T^{j_1 \ldots j_s}_{i_1 \ldots i_m}.$$

Le nombre $m + s$ s'appelle *rang* du tenseur T. Si $m = s = 0$, le tenseur T est *du rang nul* ; il représente un nombre qui ne dépend pas de base.

Un tenseur $T^{j_1 \ldots j_s}_{i_1 \ldots i_m}$ se dit *égal* à un tenseur de la même structure $Q^{j_1 \ldots j_s}_{i_1 \ldots i_m}$ si, pour tout ensemble fixe d'indices $i_1, \ldots, i_m$, $j_1, \ldots, j_s$, les composantes correspondantes de ces tenseurs coïncident dans tout système de coordonnées : $T^{j_1 \ldots j_s}_{i_1 \ldots i_m} = Q^{j_1 \ldots j_s}_{i_1 \ldots i_m}$. Il suffit d'ailleurs de prouver cette coïncidence dans un seul système de coordonnées, dans tout autre système elle découlera alors de la loi de transformation des tenseurs. Notons encore que l'*ordre dans*

lequel sont mis les indices est essentiel: en général,

$$T^k_{ij} \neq T^k_{ji} \neq T^j_{ik}.$$

Le sens des mots « covariant », « contravariant » est bien simple: « covariant » veut dire se transformant *de même* que les vecteurs de base $\{e_i\}$, avec les coefficients $p^{i}_{i'}$; « contravariant » veut dire se transformant inversement, avec les coefficients $p^{i'}_{i}$.

c. Un tenseur $T^{\cdots}_{ij\ldots}$ s'appelle *symétrique par rapport aux indices inférieurs* i, j si $T^{\cdots}_{ij\ldots} = T^{\cdots}_{ji\ldots}$ et *antisymétrique par rapport à ces indices* si $T^{\cdots}_{ij\ldots} = -T^{\cdots}_{ji\ldots}$. Il suffit de vérifier la propriété de symétrie (antisymétrie) d'un tenseur dans un seul système de coordonnées; la même propriété dans tout autre système sera valable d'après les formules de transformation des tenseurs; par exemple, pour le cas de symétrie, on a:

$$T^{\cdots}_{i'j'\ldots} = p^i_{i'} p^j_{j'} \ldots T^{\cdots}_{ij\ldots} = p^i_{i'} p^j_{j'} T^{\cdots}_{ji\ldots} = T^{\cdots}_{j'i'\ldots}.$$

Une définition analogue de la symétrie (antisymétrie) peut être formulée pour un couple d'indices *supérieurs* (contravariants). Mais la symétrie par rapport à un indice supérieur et un indice inférieur, $T^{j\cdots}_{i\ldots} = T^{i\cdots}_{j\ldots}$, n'a plus de sens absolu, elle n'est pas conservée en passant à un autre système de coordonnées.

d. En tant qu'exemple de tenseur une fois contravariant, considérons l'ensemble des coordonnées d'un vecteur x. En effet, nous avons:

$$x = \xi^i e_i = \xi^i p^{i'}_i e_{i'} = \xi^{i'} e_{i'},$$

d'où

$$\xi^{i'} = p^{i'}_i \xi^i,$$

ce qui représente la loi de transformation d'un tenseur une fois contravariant.

D'une façon analogue, les coefficients l_i d'une forme linéaire $L(x) = l_i \xi^i$ représentent un tenseur une fois contravariant; les éléments a^j_i de la matrice d'un opérateur linéaire un tenseur du deuxième rang, une fois covariant et une fois contravariant (*L5.63*).

e. Le symbole δ^k_i représente lui aussi un tenseur, une fois contra et une fois covariant; en effet, l'égalité

$$\delta^{k'}_{i'} = p^i_{i'} p^{k'}_k \delta^k_i$$

a lieu par définition de δ^k_i et en vertu des propriétés des matrices $p^{i}_{i'}$.

6.13. Opérations sur les tenseurs. On définit les opérations suivantes sur les tenseurs:

a. Multiplication d'un tenseur par un nombre et addition de deux tenseurs d'une

même structure. Soient, par exemple, T_{ij}^{k} et Q_{ij}^{k} deux tenseurs deux fois covariants et une fois contravariants, et α et β deux nombres. Formons, dans chaque système de coordonnées, les nombres S_{ij}^{k} en additionnant pour i, j, k fixes les quantités correspondantes αT_{ij}^{k} et βQ_{ij}^{k}. Les valeurs S_{ij}^{k} définies ainsi dans tout système des coordonnées forment encore un tenseur, deux fois covariant et une fois contravariant, puisque

$$S_{i'j'}^{k'} = \alpha T_{i'j'}^{k'} + \beta Q_{i'j'}^{k'} = \alpha p_{i'}^{i} p_{j'}^{j} p_{k}^{k'} T_{ij}^{k} + \beta p_{i'}^{i} p_{j'}^{j} p_{k}^{k'} Q_{ij}^{k} =$$
$$= p_{i'}^{i} p_{j'}^{j} p_{k}^{k'} (\alpha T_{ij}^{k} + \beta Q_{ij}^{k}) = p_{i'}^{i} p_{j'}^{j} p_{k}^{k'} S_{ij}^{k}.$$

b. Comme l'addition des tenseurs et leur multiplication par les nombres se réduisent à l'addition des composantes et à leur multiplication par les nombres, ces opérations sont soumises aux lois commutative, associative et distributive.

En particulier, les *tenseurs d'une structure donnée forment un espace vectoriel*. Un tenseur de m indices ayant n^m composantes, la dimension de l'espace des tenseurs de m indices vaut n^m.

c. Multiplication de deux tenseurs de n'importe quelle structure. Par exemple, multiplions un tenseur T_{ij} deux fois covariant par un tenseur S_{k}^{l} une fois covariant et une fois contravariant. Pour le faire, multiplions dans tout système de coordonnées, pour i, j, k, l fixes, les composantes correspondantes T_{ij} et S_{k}^{l}. Nous obtenons les quantités $Q_{ijk}^{l} = T_{ij} S_{k}^{l}$ dépendant de quatre indices. Ces quantités définies dans chaque système de coordonnées forment un tenseur, car

$$Q_{i'j'k'}^{l'} = T_{i'j'} S_{k'}^{l'} = p_{i'}^{i} p_{j'}^{j} T_{ij} p_{k'}^{k} p_{l}^{l'} S_{k}^{l} = p_{i'}^{i} p_{j'}^{j} p_{k'}^{k} p_{l}^{l'} T_{ij} S_{k}^{l} =$$
$$= p_{i'}^{i} p_{j'}^{j} p_{k'}^{k} p_{l}^{l'} Q_{ijk}^{l},$$

ce tenseur étant trois fois covariant et une fois contravariant. On définit d'une manière analogue la multiplication de n'importe quels deux tenseurs ; le tenseur produit accumule tous les indices covariants et contravariants de deux facteurs.

Il faut dire que la multiplication des tenseurs n'est en général pas commutative. Par exemple, le produit des tenseurs S_i et T_j a pour composante des indices $i = 1$ et $j = 2$ le nombre $S_1 T_2$, le produit des tenseurs T_i et S_j ayant pour la même composante le nombre $S_2 T_1$.

d. Contraction d'un tenseur par rapport à un indice supérieur et un indice inférieur. Cette opération s'effectue sur un tenseur ayant au moins un indice contravariant et un indice covariant. Soit, par exemple, T_{ij}^{k} un tenseur. *Le contracter par rapport aux indices i et k* signifie : composer

dans chaque système de coordonnées, pour j fixe, les nombres

$$Q_j = T^i_{ij}.$$

Ici on sous-entend la sommation sur l'indice i dans le second membre. Les quantités obtenues Q_j, considérées dans chaque système de coordonnées, constituent encore un tenseur ; en effet, l'égalité

$$T^{k'}_{i'j'} = p^i_{i'} p^j_{j'} p^{k'}_k T^k_{ij}$$

implique

$$Q_{j'} = T^{i'}_{i'j'} = p^i_{i'} p^j_{j'} p^{i'}_k T^k_{ij} = \delta^i_k p^j_{j'} T^k_{ij} = p^j_{j'} T^i_{ij} = p^j_{j'} Q_j.$$

Qu'est-ce qu'on obtient en contractant un tenseur T^j_i par rapport à ses deux indices? La quantité T^i_i n'a aucun indice, donc, représente *un nombre* dans tout système de coordonnées qui ne change pas en passant d'un système à l'autre, i.e. est un invariant. Ceci est confirmé, indépendamment de ce qui précède, par le calcul direct :

$$T^{i'}_{i'} = p^i_{i'} p^{i'}_k T^k_i = \delta^i_k T^k_i = T^i_i.$$

Ainsi, l'opération de contraction d'un tenseur (par rapport à un ou à plusieurs couples d'indices) peut fournir des invariants.

e. Transposition des coefficients $p^{i'}_i$. Considérons l'égalité

$$p^{i'}_i S^i = T^{i'}, \tag{1}$$

où les quantités S^i et $T^{i'}$ peuvent avoir encore quelques indices. On suppose que l'indice i' soit *libre*, i.e. ne soit pas un indice de sommation. Nous affirmons que cette égalité est équivalente à

$$S^i = p^i_{i'} T^{i'},$$

c'est-à-dire qu'il est possible de transposer les coefficients $p^{i'}_i$ d'un membre dans l'autre en permutant au même temps les indices. En effet, en multipliant l'égalité (1) par $p^k_{i'}$ et sommant sur i', nous avons

$$p^k_{i'} p^{i'}_i S^i = \delta^k_i S^i = S^k = p^k_{i'} T^{i'},$$

avec un indice libre k que l'on peut, bien entendu, remplacer par i.

f. Simplification par les coefficients $p^{i'}_i$. Soit

$$p^{i'}_i S^i = p^{i'}_i Q^i, \tag{2}$$

où les quantités S^i et Q^i peuvent aussi bien avoir d'autres indices. L'indice i' est supposé libre.

Montrons que l'égalité (2) est équivalente à

$$S^i = Q^i,$$

c'est-à-dire que l'on peut simplifier par les coefficients $p_i^{i'}$. En effet multiplions (2) par $p_{i'}^k$ et sommons sur k; nous obtenons

$$p_{i'}^k p_i^{i'} S^i = \delta_i^k S^i = S^k = p_{i'}^k p_i^{i'} Q^i = \delta_i^k Q^i = Q^k,$$

avec un indice libre k qu'on peut remplacer par i.

6.14. Résolution des équations tensorielles.

a. Examinons un système linéaire d'équations

$$R_{ij} S^{j\cdots}_{\cdots} = T^{\cdots}_{i\cdots}, \tag{1}$$

où R_{ij} et $T^{\cdots}_{i\cdots}$ sont des tenseurs de structure indiquée, et $\det \| R_{ij} \| \neq 0$. Le système (1) permet, pour tout ensemble d'indices, de déterminer les nombres $S^{j\cdots}_{\cdots}$ d'une manière univoque. Donc, les nombres $S^{j\cdots}_{\cdots}$ sont bien déterminés *pour tout système de coordonnées*. Montrons que ces nombres *composent un tenseur*. Dans nos calculs, nous nous bornons au cas de tenseur T^q_{ir} et d'inconnues S^{jq}_r. En vertu de (1), on a

$$p_i^{i'} p_j^{j'} R_{i'j'} S_r^{iq} = p_i^{i'} p_r^{r'} p_q^q T^{q'}_{i'r'},$$

où, conformément à *6.13e, f*,

$$R_{i'j'} p^r_{r'} p_j^{j'} p_q^{q'} S_r^{jq} = T^{q'}_{i'r'};$$

l'unicité de solution du système (1) dans la base $\{e_{i'}\}$ implique

$$S_{r'}^{j'q'} = p^r_{r'} p_j^{j'} p_q^{q'} S_r^{jq},$$

ce qui signifie que les quantités S_r^{jq} forment un tenseur.

b. Un raisonnement analogue montre que les équations

$$R_i^j S^k_{j\cdots} = T^k_{i\cdots}, \qquad R^{ij} S_{jk\cdots} = T^i_{k\cdots},$$

où R et T sont des tenseurs de structure indiquée, $\det \| R_i^j \| \neq 0$, $\det \| R^{ij} \| \neq 0$, et $S^{\cdots}_{\cdots}$ des inconnues, ont pour solutions les *tenseurs* $S^{\cdots}_{\cdots}$ de structure indiquée.

c. Un autre exemple est fourni par le système d'équations de la forme

$$T^{\cdots}_{\cdots k} \underset{i}{\xi^k} = S^{\cdots}_{\cdots}, \tag{2}$$

où $\underset{1}{\xi^k}, \ldots, \underset{n}{\xi^k}$ sont n tenseurs une fois contravariants linéairement indépendants, de sorte que $\det \| \underset{i}{\xi^k} \| \neq 0$, les indices des tenseurs $\underset{i}{S^{\cdots}_{\cdots}}$ coïncidant avec les indices correspondants des quantités $T^{\cdots}_{\cdots k}$ qu'on a remplacés par des points.

Le système (2) permet, pour toutes les valeurs des indices non explicités, de déterminer d'une façon univoque les quantités $T^{\cdots}_{\cdots k}$. Montrons que $T^{\cdots}_{\cdots k}$ *est un tenseur de même structure que* $\underset{i}{S^{\cdots}_{\cdots}}$, *avec un indice covariant supplémentaire* k. Pour fixer les idées, faisons le calcul pour les tenseurs $\underset{i}{S^{m}_{rq}}$ et pour les quantités T^{m}_{rqk}. En passant à la base $\{e_{i'}\}$ on peut écrire :

$$p^{r'}_{r}\, p^{q'}_{q}\, p^{m}_{m'} \underset{i}{S^{m'}_{r'q'}} = \underset{i}{S^{m}_{rq}} = T^{m}_{rqk} \underset{i}{\xi^{k}} = T^{m}_{rqk} p^{k}_{k'} \underset{i}{\xi^{k'}},$$

$$\underset{i}{S^{m'}_{r'q'}} = p^{r}_{r'} p^{q}_{q'} p^{m'}_{m} p^{k}_{k'} \underset{i}{\xi^{k'}} T^{m}_{rqk},$$

d'où, vu l'unicité de solution du système (2) dans le nouveau système de coordonnées, on a

$$T^{m'}_{r'q'k'} = p^{r}_{r'} p^{q}_{q'} p^{m'}_{m} p^{k}_{k'} T^{m}_{rqk},$$

ce qu'il fallait.

d. Considérons un système plus compliqué

$$T^{\cdots}_{\cdots km} \underset{i}{\xi^{k}} \underset{j}{\eta^{m}} = \underset{ij}{S^{\cdots}_{\cdots}},$$

où $\underset{i}{\xi^{k}}$ est une famille de vecteurs linéairement indépendants, $\underset{j}{\eta^{m}}$ une autre famille de vecteurs linéairement indépendants, et les indices, remplacés par les points des tenseurs $\underset{ij}{S^{\cdots}_{\cdots}}$ sont les mêmes que les indices correspondants du tenseur $T^{\cdots}_{\cdots km}$. Dans ce cas, $T^{\cdots}_{\cdots km}$ s'avère *tenseur de la même structure que* $\underset{ij}{S^{\cdots}_{\cdots}}$, *avec deux indices covariants supplémentaires* k *et* m.

Pour le démontrer, posons

$$T^{\cdots}_{\cdots km} \xi^{k} = \underset{i}{R^{\cdots}_{\cdots m}}\,; \tag{3}$$

nous avons $\underset{i}{R^{\cdots}_{\cdots m}} \eta^{m} = \underset{ij}{S^{\cdots}_{\cdots}}$, et, d'après ce qu'on a démontré dans c, $\underset{i}{R^{\cdots}_{\cdots m}}$ est, pour tout i, un tenseur de la même structure que $\underset{ij}{S^{\cdots}_{\cdots}}$, avec un indice covariant supplémentaire m. En appliquant maintenant c à l'équation (3) nous arrivons à la proposition relative à $T^{\cdots}_{\cdots km}$. Il est évident qu'un résultat analogue sera valable pour certains systèmes plus compliqués, par exemple de la forme $T^{\cdots}_{\cdots kmr} \underset{i}{\xi^{k}} \underset{j}{\eta^{m}} \underset{l}{\zeta^{r}} = \underset{ijl}{S^{\cdots}_{\cdots}}$, etc.

6.15. Bivecteur.

a. Soient ξ et η deux vecteurs. Le tenseur

$$f = f^{ij}(\xi, \eta) = \xi^{i}\eta^{j} - \xi^{j}\eta^{i} \equiv [\xi, \eta]^{ij} \tag{1}$$

s'appelle *bivecteur engendré par les vecteurs* ξ *et* η. Les composantes du bivecteur f^{ij} sont les mineurs du second ordre dans la matrice

$$\left\| \begin{matrix} \xi^1 & \xi^2 & \dots & \xi^n \\ \eta^1 & \eta^2 & \dots & \eta^n \end{matrix} \right\| \tag{2}$$

composée des coordonnées des vecteurs ξ et η. Il est clair que le bivecteur f ne change pas en additionnant au vecteur η le produit du vecteur ξ par un nombre arbitraire. Ainsi, il est impossible de déterminer les vecteurs ξ et η d'une façon univoque d'après le bivecteur f, i.e. d'après les mineurs de la matrice (2).

Posons la question suivante: quelle information géométrique sur les vecteurs ξ et η peut-on tirer de la connaissance du bivecteur f?

b. Avant tout, étant donné les composantes du bivecteur f, on peut déterminer de façon univoque le *plan des vecteurs* ξ et η. En effet, l'hypothèse d'appartenance d'un vecteur $\tau = \tau^i e_i$ au plan ξ, η a la forme $\tau = C_1\xi + C_2\eta$ ou, ce qui revient au même, la dépendance linéaire des lignes de la matrice

$$\left\| \begin{matrix} \xi^1 & \dots & \xi^n \\ \eta^1 & \dots & \eta^n \\ \tau^1 & \dots & \tau^n \end{matrix} \right\|. \tag{3}$$

Autrement dit, le rang de la matrice (3) doit être inférieur à 3, de sorte que tous les mineurs du 3-ème ordre de cette matrice doivent être nuls. En développant les mineurs du 3-ème ordre

$$\begin{vmatrix} \xi^i & \xi^j & \xi^k \\ \eta^i & \eta^j & \eta^k \\ \tau^i & \tau^j & \tau^k \end{vmatrix}$$

suivant les éléments de la dernière ligne, nous arrivons à une série de conditions d'appartenance du vecteur τ au plan ξ, η exprimées par les mineurs du 2-ème ordre de la matrice (3), i.e. par les composantes du bivecteur f.

c. Remplaçons les vecteurs ξ et η par leurs combinaisons linéaires

$$\left.\begin{aligned} p &= \alpha_{11}\xi + \alpha_{12}\eta, \\ q &= \alpha_{21}\xi + \alpha_{22}\eta \end{aligned}\right\} \tag{4}$$

et trouvons le bivecteur engendré par les vecteurs p et q:

$$[p, q]^{ij} = \begin{vmatrix} p^i & p^j \\ q^i & q^j \end{vmatrix} = \begin{vmatrix} \alpha_{11}\xi^i + \alpha_{12}\eta^i & \alpha_{11}\xi^j + \alpha_{12}\eta^j \\ \alpha_{21}\xi^i + \alpha_{22}\eta^i & \alpha_{21}\xi^j + \alpha_{22}\eta^j \end{vmatrix} =$$

$$= \begin{vmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{vmatrix} \begin{vmatrix} \xi^i & \xi^j \\ \eta^i & \eta^j \end{vmatrix} = \det \| \alpha_{ij} \| \, [\xi, \eta]^{ij}.$$

Ainsi, à la suite de substitution (4), le bivecteur $[\xi, \eta]$ se multiplie par le nombre $\det \| \alpha_{ij} \|$. Appelons les couples ξ, η et p, q (4) *équivalents* si $\det \| \alpha_{ij} \| = 1$. Pour deux couples équivalents, les composantes de mêmes indices du bivecteur f sont égales: $[p, q]^{ij} = [\xi, \eta]^{ij}$. On peut classer tous les couples de vecteurs ξ, η, dans leur plan γ, en en formant les classes de couples équivalents deux à deux disjointes; on peut alors dire que les composantes d'un bivecteur $[\xi, \eta]$ déterminent d'une façon univoque le plan γ et la classe contenant le couple ξ, η. Chaque autre classe peut être obtenue en multipliant le bivecteur $[\xi, \eta]$ par un nombre convenable.

6.16. Tenseur métrique. Un produit scalaire (x, y) dans l'espace R_n peut être défini comme valeur $G(x, y)$ d'une forme bilinéaire symétrique $(G(x, y) = G(y, x))$ et définie positive $(G(x, x) > 0$ pour $x \neq 0)$. En coordonnées, la forme $G(x, y)$ s'écrit

$$G(x, y) = \sum_i \sum_j g_{ij} \xi^i \eta^j = g_{ij} \xi^i \eta^j, \tag{1}$$

où $x = \xi^i e_i$, $y = \eta^j e_j$, g_{ij} étant des nombres.

On peut dire que donner une forme $G(x, y)$ équivaut à donner un tenseur g_{ij} deux fois covariant, symétrique $(g_{ij} = g_{ji})$ et défini positif $(g_{ij} \xi^i \xi^j > 0$ pour $\xi \neq 0)$. En effet, s'il y a un tel tenseur g_{ij}, l'expression (1) représente, pour tout couple de vecteurs (de tenseurs une fois contravariants), un nombre qui satisfait aux axiomes de produit scalaire. Un tenseur g_{ij} avec les propriétés indiquées s'appelle *tenseur métrique*. Un tenseur métrique g_{ij} transforme l'espace affine R_n en un espace euclidien, où il est possible de mesurer les longueurs des vecteurs et les angles qu'ils forment, les aires des figures et les volumes des corps.

La matrice g_{ij} est non dégénérée puisque, définie positive, elle admet donc la matrice inverse g^{jk}:

$$g_{ij} g^{jk} = \delta_i^k, \quad i, k = 1, \ldots, n. \tag{2}$$

En vertu de *6.14*, les nombres g^{jk} forment un tenseur deux fois contravariant. Si l'on considère, par exemple, un tenseur quelconque T_{ij}^k (de la structure indiquée), alors les tenseurs

$$T_{ijs} = T_{ij}^k g_{ks}, \quad T_i^{ks} = T_{ij}^k g^{js}, \quad T^{rks} = T_i^{ks} g^{ir}$$

sont dits *adjoints de* T_{ij}^k par rapport au tenseur g_{ij}.

Notons que, pour un tenseur g_{ij}, il existe une base dans l'espace R_n telle que

$$g_{ij} = \begin{cases} 0 \text{ pour } i \neq j, \\ 1 \text{ pour } i = j, \end{cases}$$

c'est-à-dire une base orthonormée par rapport au produit scalaire (1). Les composantes du tenseur métrique dans une telle base sont désignées par δ_{ij}.

6.17.a. Dans l'espace euclidien R_n, l'aire du parallélogramme construit sur les vecteurs ξ et η peut être exprimée par les composantes du bivecteur $[\xi, \eta]$. En effet, ladite aire S est calculée d'après la formule *3.55* (2):

$$S^2 = \begin{vmatrix} (\xi, \xi) & (\xi, \eta) \\ (\eta, \xi) & (\eta, \eta) \end{vmatrix} = \begin{vmatrix} g_{ik}\xi^i\xi^k & g_{jl}\xi^j\eta^l \\ g_{ik}\eta^i\xi^k & g_{jl}\eta^j\eta^l \end{vmatrix} =$$

$$= g_{ik}g_{jl}\xi^k\eta^l \begin{vmatrix} \xi^i & \xi^j \\ \eta^i & \eta^j \end{vmatrix}. \qquad (1)$$

En permutant les indices i et j, de même que k et l, on a:

$$S^2 = g_{jl}g_{ik}\xi^l\eta^k \begin{vmatrix} \xi^j & \xi^i \\ \eta^j & \eta^i \end{vmatrix} = -g_{ik}g_{jl}\xi^l\eta^k \begin{vmatrix} \xi^i & \xi^j \\ \eta^i & \eta^j \end{vmatrix}. \qquad (2)$$

Il résulte de (1) et (2):

$$2S^2 = g_{ik}g_{jl} \begin{vmatrix} \xi^i & \xi^j \\ \eta^i & \eta^j \end{vmatrix} \begin{vmatrix} \xi^k & \xi^l \\ \eta^k & \eta^l \end{vmatrix}. \qquad (3)$$

En permutant ensuite dans (3) les indices k et l on obtient:

$$2S^2 = g_{il}g_{jk} \begin{vmatrix} \xi^i & \xi^j \\ \eta^i & \eta^j \end{vmatrix} \begin{vmatrix} \xi^l & \xi^k \\ \eta^l & \eta^k \end{vmatrix} = -g_{jk}g_{il} \begin{vmatrix} \xi^i & \xi^j \\ \eta^i & \eta^j \end{vmatrix} \begin{vmatrix} \xi^k & \xi^l \\ \eta^k & \eta^l \end{vmatrix}, \qquad (4)$$

donc,

$$4S^2 = \begin{vmatrix} g_{ik} & g_{il} \\ g_{jk} & g_{jl} \end{vmatrix} \begin{vmatrix} \xi^i & \xi^j \\ \eta^i & \eta^j \end{vmatrix} \begin{vmatrix} \xi^k & \xi^l \\ \eta^k & \eta^l \end{vmatrix}. \qquad (5)$$

La formule (5) fournit l'expression cherchée de S par les composantes du bivecteur $[\xi, \eta]$.

b. La quantité $G_{ij,\,kl} = g_{ik}g_{jl} - g_{jk}g_{il}$ représente le mineur de la matrice $\| g_{ij} \|$ qui occupe les lignes de numéros i et j et les colonnes de numéros k et l. Cette quantité est un tenseur quatre fois covariant.

Le tenseur $G_{ij,\,kl}$ s'appelle *tenseur métrique dérivé*. On voit de la formule (5) que l'aire du parallélogramme construit sur les vecteurs ξ et η est bien déterminée par le bivecteur $[\xi, \eta]$. Par conséquent, il est possible de définir « l'aire d'un bivecteur » en sous-entendant par là l'aire de n'importe quel parallélogramme de la classe des parallélogrammes déterminés par le bivecteur donné. Si l'aire du bivecteur $[\xi, \eta]$ vaut 1, il s'appelle *bivecteur unitaire*; tous les parallélogrammes qu'il détermine ont l'aire 1. Chaque autre bivecteur $[\zeta, \tau]$ du même plan γ, ayant S pour aire, se déduit d'un bivecteur unitaire en le multipliant par le nombre S.

c. Le tenseur $G_{ij,\,kl}$, par sa définition même, est antisymétrique par rapport aux indices i et j, antisymétrique par rapport aux indices k et l, et ne change pas en remplaçant i et j par k et l respectivement.

Il satisfait à la relation suivante appelée identité de Ricci :

$$G_{ij,\,kl} + G_{jk,\,il} + G_{ki,\,jl} = 0. \qquad (6)$$

Les termes de cette somme s'obtiennent en permettant les trois premiers indices d'une façon cyclique, le quatrième étant fixe. Comme c'est une identité tensorielle, il suffit de la démontrer dans un seul système de coordonnées. Choisissons un système où les composantes δ_{ij} du tenseur g_{ij} valent 0 pour $i \neq j$ et 1, pour $i = j$.

L'égalité (6) prend alors la forme :

$$\begin{vmatrix} \delta_{ik} & \delta_{il} \\ \delta_{jk} & \delta_{jl} \end{vmatrix} + \begin{vmatrix} \delta_{ji} & \delta_{jl} \\ \delta_{ki} & \delta_{kl} \end{vmatrix} + \begin{vmatrix} \delta_{kj} & \delta_{kl} \\ \delta_{ij} & \delta_{il} \end{vmatrix} = 0.$$

Si $l \neq i$, $l \neq j$, $l \neq k$, alors les deuxièmes colonnes dans les déterminants ci-écrits sont nulles, et l'égalité est vérifiée. Supposons que l coïncide avec l'un des indices i, j, k. En raison de symétrie par rapport à i, j, k, il suffit de poser $l = i$, $l \neq j$, $l \neq k$. Alors la deuxième colonne dans le second déterminant s'annule ; la somme du premier et du troisième devient

$$\begin{vmatrix} 0 & 1 \\ \delta_{jk} & 0 \end{vmatrix} + \begin{vmatrix} \delta_{kj} & 0 \\ 0 & 1 \end{vmatrix} = -\delta_{jk} + \delta_{kj} = 0.$$

Supposons que l coïncide avec deux quelconques des trois indices i, j, k, par exemple $l = i = j$, $l \neq k$. Alors la somme des trois déterminants prend la forme

$$\begin{vmatrix} 0 & 1 \\ 0 & 1 \end{vmatrix} + \begin{vmatrix} 1 & 1 \\ 0 & 0 \end{vmatrix} + \begin{vmatrix} 0 & 0 \\ 1 & 1 \end{vmatrix} = 0.$$

Enfin, si $l = i = j = k$, tous les déterminants s'annulent, car tous leurs éléments seront égaux à 1. Ainsi, le tenseur $G_{ij,\,kl}$ satisfait effectivement à l'identité de Ricci.

6.18. Tenseurs du type Ricci.

a. Un tenseur $T_{ij,\,kl}$ quatre fois covariant s'appelle *tenseur du type Ricci* s'il possède les propriétés suivantes :

$T_{ij,\,kl} = -T_{ji,\,kl}$ (antisymétrie par rapport à i et j) ; (1)
$T_{ij,\,kl} = T_{kl,\,ij}$ (symétrie par rapport aux couples i, j et k, l) ; (2)
$T_{ij,\,kl} + T_{jk,\,il} + T_{ki,\,jl} = 0$ (identité de Ricci). (3)

De (1) et (2) on déduit l'antisymétrie par rapport aux indices k et l :

$$T_{ij,\,lk} = T_{lk,\,ij} = -T_{kl,\,ij} = -T_{ij,\,kl}. \qquad (4)$$

Le tenseur métrique dérivé $G_{ij,\,kl}$ (*6.17c*) fournit un exemple de tenseur du type Ricci. Nous en verrons par la suite d'autres exemples importants.

Il est évident que l'on peut additionner et multiplier par les nombres les tenseurs du type Ricci en obtenant toujours les tenseurs du type Ricci.

b. Soit $z = [x, y]$, $z^{ij} = \xi^i\eta^j - \xi^j\eta^i$ un bivecteur. Considérons la contraction d'un tenseur de Ricci $T_{ij,\,kl}$ et de deux bivecteurs égaux z:

$$T(z, z) = T_{ij,\,kl}z^{ij}z^{kl}. \tag{5}$$

Dans tout système de coordonnées, cette contraction est un nombre dépendant du bivecteur z.

Il s'avère que *toutes les composantes d'un tenseur du type Ricci sont déterminées d'une manière univoque par les contractions* (5) *avec tous les bivecteurs z possibles.*

En guise de démonstration, il suffit de prouver que l'égalité $T(z, z) \equiv 0$ implique $T_{ij,\,kl} \equiv 0$ pour tous les i, j, k, l.

Supposons que

$$T(z, z) = T_{ij,\,kl}z^{ij}z^{kl} = T_{ij,\,kl}(\xi^i\eta^j - \xi^j\eta^i)\,z^{kl} = \\ = T_{ij,\,kl}\xi^i\eta^jz^{kl} - T_{ij,\,kl}\xi^j\eta^iz^{kl} \equiv 0.$$

En permutant dans le deuxième terme les indices i et j, puis remplaçant $T_{ji,\,kl}$ par $-T_{ij,\,kl}$, nous trouvons

$$T_{ij,\,kl}\xi^i\eta^jz^{kl} + T_{ij,\,kl}\xi^i\eta^jz^{kl} \equiv 0,$$

de sorte que

$$T_{ij,\,kl}\xi^i\eta^jz^{kl} \equiv 0.$$

En faisant de même avec le facteur $z^{kl} = \xi^k\eta^l - \xi^l\eta^k$, nous avons

$$T_{ij,\,kl}\xi^i\eta^j\xi^k\eta^l \equiv 0.$$

Cette relation doit avoir lieu pour toutes les valeurs de $\xi^1, \ldots, \xi^n$, $\eta^1, \ldots, \eta^n$ et pour les valeurs des i, j, k, l de 1 à n. Fixons $i = k$, $j = l$ et posons $x = (0, \ldots, \underset{i}{1}, \ldots, 0)$, $y = (0, \ldots, \underset{j}{1}, \ldots 0)$; quels que soient i, j, nous obtenons

$$T_{ij,\,ij} = 0. \tag{6}$$

Fixons $i = k$, $j \neq l$ et posons $x = (0, \ldots, \underset{i}{1}, \ldots 0)$, $y = (0, \ldots, \underset{j}{1}, \ldots, \underset{l}{1}, \ldots, 0)$; nous obtenons

$$T_{ij,\,il} + T_{il,\,ij} = 0.$$

En vertu de la propriété (2), cela amène à l'égalité

$$T_{ij,\,il} = 0, \tag{7}$$

pour tous les i et $j \neq l$. Avec (1) et (4) nous trouvons encore que

$$T_{jl,\,il} = 0, \tag{8}$$

pour tous les i et $l \neq j$. Fixons $i \neq k$, $j \neq l$ et posons

$$x = (0, \ldots, \underset{i}{1}, \ldots, \underset{k}{1}, \ldots, 0),$$

$$y = (0, \ldots, \underset{j}{1}, \ldots, \underset{l}{1}, \ldots, 0);$$

nous obtenons

$$T_{ij,\,kl} + T_{kj,\,il} + T_{il,\,kj} + T_{kl,\,ij} = 0.$$

En vertu de (2), cette relation se réduit à

$$T_{ij,\,kl} + T_{kj,\,il} = 0. \tag{9}$$

En permutant i et j et utilisant (1), nous avons

$$T_{ij,\,kl} + T_{ik,\,jl} = 0. \tag{10}$$

Ecrivons encore une égalité découlant de (1):

$$T_{ij,\,kl} + T_{ji,\,kl} = 0. \tag{11}$$

Additionnons les trois égalités (9)-(11); d'après l'identité de Ricci (3), les deuxièmes termes disparaissent et, après avoir divisé par 3, nous voyons que, pour $i \neq k$, $j \neq l$,

$$T_{ij,\,kl} = 0. \tag{12}$$

Avec (6) et (7), cela signifie que toutes les composantes du tenseur $T_{ij,\,kl}$ sont nulles, et la proposition est démontrée.

§ 6. 2. Variété différentiable élémentaire

6.21.a. Soit M un ensemble équipotent à une boule dans l'espace n-dimensionnel R_n. On peut donc définir sur M certaines coordonnées réelles $x^1, \ldots, x^n$, de sorte que le point $x = (x^1, \ldots, x^n)$ parcourt dans l'espace R_n une boule ouverte. A côté des coordonnées $x^1, \ldots, x^n$, nous considérons sur M n'importe quelles autres coordonnées $x^{1'}, \ldots, x^{n'}$, pourvu qu'elles soient liées à $x^1, \ldots, x^n$ par une correspondance biunivoque de la forme

$$\left.\begin{aligned} x^{1'} &= x^{1'}(x^1, \ldots, x^n), \\ &\cdots\cdots\cdots\cdots \\ x^{n'} &= x^{n'}(x^1, \ldots, x^n), \end{aligned}\right\} \tag{1}$$

les fonctions $x^{i'}(x^1, \ldots, x^n)$ possédant les dérivées continues de N premiers ordres et la matrice $\frac{\partial(x^{1'}, \ldots, x^{n'})}{\partial(x^1, \ldots x^n)}$ étant non dégénérée. De tels systèmes de coordonnées sont dits *admissibles*. Un ensemble $M = M_{n,\,N}$ muni d'un ensemble des systèmes de coordonnées admissibles s'appelle *variété différentiable élémentaire de dimension n* (n-dimensionnelle) *de la classe* D_N (ou N tout court).

Si l'on n'admet que les systèmes de coordonnées $x^1, \ldots, x^n$ pour lesquels les fonctions (1) soient indéfiniment dérivables, alors la variété M, par définition, appartient à la classe D_∞.

b. Deux variétés différentiables M_1 et M_2 de dimensions n_1 et n_2 et de classes N_1 et N_2 respectivement sont dits *équivalentes* si $n_1 = n_2 = n$ et $N_1 = N_2 = N$. Alors il est possible d'établir une telle correspondance biunivoque entre leurs points que les coordonnées d'un point de la variété M_2 soient fonctions de la classe N des coordonnées du point correspondant de la variété M_1. De plus, on peut choisir la correspondance d'une telle façon que ces fonctions soient du premier degré. En effet, les coordonnées sur la variété M_1 établissent une correspondance entre M_1 et une boule dans R_n, les coordonnées sur la variété M_2 en font autant ; or, n'importe quelles deux boules dans R_n peuvent être transformées l'une en l'autre moyennant une translation et une homothétie, c'est-à-dire avec les fonctions du premier degré.

6.22. Exemples.

a. *Un cercle ouvert dans le plan, est-il une variété différentiable*?

Réponse. La question est dépourvue de sens du moment que les systèmes de coordonnées admissibles ne soient pas donnés.

b. *Un cercle dans le plan est décrit par les coordonnées polaires* $\rho < 1$, $0 \leqslant \varphi \leqslant 2\pi$, *et l'on choisit pour coordonnées admissibles celles qui sont liées à* ρ et φ *par les relations du type* (1) *d'une classe fixe* D_N. *Est-ce que ce cercle représente une variété différentiable*?

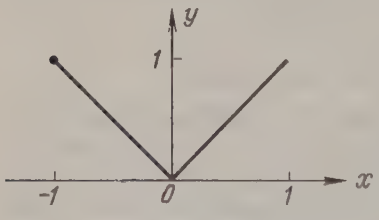


Fig. 6.2-1

Réponse. Non. L'ensemble des valeurs des paramètres $0 \leqslant \rho < 1$, $0 < \varphi \leqslant 2\pi$ n'est pas un cercle ouvert dans le plan ρ, φ, où ρ et φ sont les coordonnées cartésiennes.

c. *Un cercle dans le plan est déterminé par l'inégalité* $x^2 + y^2 < 1$ *en coordonnées cartésiennes*; *on choisit pour coordonnées admissibles les coordonnées qui sont liées à* x, y *par les relations du type* (1) *d'une classe fixe* D_N. *Est-ce que ce cercle est une variété différentiable*?

Réponse. Oui, il en est une.

d. *Un ensemble du plan est défini par les conditions* (fig. 6.2-1)

$$x \geqslant 0\colon\ y = x, \qquad 0 \leqslant x < 1,$$
$$x \leqslant 0\colon\ y = -x, \qquad -1 < x \leqslant 0$$

et muni de la coordonnée s (*longueur d'arc*) *comptée à partir du point* 0, *avec le signe* + *à droite de ce point et le signe* — *à sa gauche*; *on considère comme coordonnées admissibles toutes les fonctions* $\sigma = \varphi(s)$ *ayant*

les dérivées continues jusqu'à un ordre N, $\varphi'(s) \neq 0$. Est-ce que cet ensemble est une variété différentiable?

Réponse. Oui. Il est équivalent à la variété $M_1 = \{-\sqrt{2} < s < \sqrt{2}\}$, ou même à la variété $M_2 = \{-1 < s < 1\}$ sur l'axe réel, avec les mêmes coordonnées admissibles.

6.23. Une propriété exprimée en coordonnées $x^1, \ldots, x^n$ est appelée propriété *absolue* ou *géométrique* de la variété M, si elle a la même expression dans tout autre système admissible de coordonnées.

a. Par exemple, une suite $A_1, \ldots, A_m, \ldots$ de points d'une variété M *converge vers un point A* de cette variété (on note $A_m \to A$), si dans un système de coordonnées $x^1, \ldots, x^n$, pour $m \to \infty$, on a

$$x^1(A_m) \to x^1(A), \ \ldots, \ x^n(A_m) \to x^n(A).$$

Alors, dans tout autre système de coordonnées admissible $x^{1'}, \ldots, x^{n'}$, on a aussi

$$x^{1'}(A_m) \to x^{1'}(A), \ \ldots \ x^{n'}(A_m) \to x^{n'}(A),$$

en vertu de la continuité des fonctions dérivables $x^{i'}(x^1, \ldots, x^n)$. Ainsi, la propriété $A_m \to A$ *est une propriété absolue de la variété M*.

b. Soient $x^i = x^i(t)$ $(i = 1, \ldots, n;\ a \leqslant t \leqslant b)$ des fonctions k fois dérivables; le lieu géométrique L des points correspondants $(x^1(t), \ldots, x^n(t)) \in M$ sera appelé *courbe k fois dérivable sur la variété M*. En passant à un autre système de coordonnées, nous avons

$$x^{i'} = x^{i'}(x^1(t), \ \ldots, \ x^n(t)) \quad (i' = 1, \ \ldots, \ n);$$

ce sont toujours des fonctions k fois dérivables en t, du moment que $k \leqslant N$. Ainsi la notion « *courbe k fois dérivable sur la variété M* » *est absolue pour* $k \leqslant N$.

c. Un point $A = \{x^1(t_0), \ldots, x^n(t_0)\} \in L \subset M$ sur la courbe L est dit *ordinaire*, si $\sum_{i=1}^{n} \left[\frac{dx^i(t_0)}{dt}\right]^2 > 0$, et *singulier*, si $\sum_{i=1}^{n} \left[\frac{dx^i(t_0)}{dt}\right]^2 = 0$. Puisque dans tout autre système de coordonnées on a

$$\frac{\partial x^{i'}(t)}{dt} = \sum_{i=1}^{n} \frac{\partial x^{i'}}{\partial x^i} \frac{dx^i(t)}{dt},$$

nous voyons qu'un point A sur la courbe L, singulier dans un système de coordonnées sur M, reste singulier sur L pour tout autre système de coordonnées. Ainsi, la notion « point singulier », donc aussi la notion « point ordinaire » sont absolues.

d. Par analogie, pour $k \leqslant N$, est absolue la notion « surface *s-dimensionnelle k fois dérivable P sur la variété M* »; on appelle ainsi

le lieu géométrique des points définis en coordonnées $x^1, \ldots, x^n$ par un système de la forme suivante, avec $s \leqslant n$ paramètres :

$$x^i = x^i(t_1, \ldots, t_s), \quad i = 1, \ldots, n, \quad t = (t_1, \ldots, t_s) \in G \subset R_s,$$

les fonctions $x^i(t_1, \ldots, t_s)$ étant k fois continûment dérivables.

e. Un point $A \in P$ s'appelle *point ordinaire de la surface* $P \subset M$ si le rang de la matrice $\left\| \frac{\partial x^i(A)}{\partial t^j} \right\|$ vaut s (le nombre des paramètres), et *singulier* si le rang de cette matrice est inférieur à s. En vertu de l'égalité *1.33* (7)

$$\left\| \frac{\partial x^{i'}}{\partial t^j} \right\| = \left\| \frac{\partial x^{i'}}{\partial x^i} \right\| \left\| \frac{\partial x^i}{\partial t^j} \right\|$$

et de la relation entre les mineurs du s-ème ordre du premier membre et les mineurs du s-ème ordre du deuxième facteur à droite qui en découle, un point A, qui est singulier sur la surface P en coordonnées $\{x^i\}$, sera singulier en coordonnées $\{x^{i'}\}$. Ainsi, la notion « point singulier » sur la surface de même que la notion « point ordinaire » sont elles aussi absolues.

6.24. Espace tangent. Soit $A \in M_n$ un point fixe. Examinons une courbe dérivable $L = \{x \in M : x^i = x^i(t),\ a \leqslant t \leqslant b\}$ sur la variété M_n qui passe par ce point. L'ensemble numérique $(\xi^1, \ldots, \xi^n)$, où

$$\xi^i = \frac{dx^i(A)}{dt}, \quad i = 1, \ldots, n,$$

forme, par définition, *un vecteur tangent à la courbe* L au point A. De tels vecteurs tangents remplissent, évidemment, tout l'espace n-dimensionnel R_n, car n'importe quel vecteur $(\xi^1, \ldots, \xi^n) \in R_n$ est tangent à une courbe $L \subset M_n$, par exemple à la courbe

$$x^i(t) = x^i(A) + \xi^i t$$

qui passe par le point A pour $t = 0$. Etant donné cette correspondance, l'espace R_n est désigné par $T_n(A)$ et s'appelle *espace tangent à la variété* M_n *au point* A. Une base de cet espace est composée par les vecteurs tangents aux courbes

$$\underset{j}{x^i}(t) = x^i(A) + \delta^i_j t \qquad (i, j = 1, \ldots, n)$$

(« lignes de coordonnées menées par le point A »).

Chaque transformation *admissible* de coordonnées $x^{i'} = x^{i'}(x^1, \ldots, x^n)$ amène à une transformation *linéaire* dans l'espace tangent au point A. Précisément dans le système de coordonnées $x^{i'}$, nous avons

$$\xi^{i'} = \frac{dx^{i'}(A)}{dt} = \sum_{i=1}^{n} \frac{\partial x^{i'}(A)}{\partial x^i} \frac{dx^i(A)}{dt} = p^{i'}_i \xi^i,$$

où $p_i^{i'} = \frac{\partial x^{i'}(A)}{\partial x^i}$. Pour voir que ce résultat soit conforme aux règles tensorielles, il faut écrire les formules de transformation des vecteurs de base. Un nouveau vecteur de base $e_{i'}$ représente un vecteur tangent à la ligne de coordonnées nouvelle $x^{i'}$, i.e. le vecteur tangent à la ligne $x^{1'} = x^{1'}(A), \ldots, x^{i'} = x^{i'}(A) + t, \ldots, x^{n'} = x^{n'}(A)$. De plus,

$$\frac{\partial x^{1'}}{\partial x^i}\frac{dx^i(A)}{dt} = 0, \ldots, \frac{\partial x^{i'}}{\partial x^i}\frac{dx^i(A)}{dt} = 1, \ldots, \frac{\partial x^{n'}}{\partial x^i}\frac{dx^i(A)}{dt} = 0,$$

donc, pour cette ligne, on a

$$\frac{dx^i(A)}{dt} = p_{i'}^i(A), \tag{1}$$

où $\| p_{i'}^i(A) \|$ est la matrice inverse à la matrice $\| p_i^{i'}(A) \|$; cette matrice inverse, comme cela résulte de *1.33b*, est formée par les éléments $p_{i'}^i = \frac{\partial x^i(A)}{\partial x^{i'}}$. En multipliant (1) par e_i et sommant sur i, on trouve

$$e_{i'} = p_{i'}^i e_i.$$

Ainsi, *les quantités* ξ^i *forment des tenseurs une fois contravariants dans l'espace tangent au point A*. Pour $M = M_{n,\,N}$, les fonctions $p_i^{i'}$ $(p_{i'}^i)$ sont dérivables $N - 1$ fois au moins.

Il est ensuite possible de définir un tenseur de structure arbitraire dans l'espace tangent $T_n(A)$. Par exemple, un tenseur $T_{ij}^k(A)$ est un système de n^3 nombres qui dépend du système de coordonnées sur M_n et qui se transforme en passant des coordonnées $x^1, \ldots, x^n$ aux coordonnées $x^{1'}, \ldots, x^{n'}$ d'après les formules

$$T_{i'j'}^{k'} = p_{i'}^i p_{j'}^j p_k^{k'} T_{ij}^k,$$

où, comme plus haut,

$$p_{i'}^i = \frac{\partial x^i(A)}{\partial x^{i'}}, \quad p_{j'}^j = \frac{\partial x^j(A)}{\partial x^{j'}}, \quad p_k^{k'} = \frac{\partial x^{k'}(A)}{\partial x^k}.$$

De la sorte, dans l'espace tangent $T_n(A)$, il est possible de réaliser toute l'algèbre tensorielle décrite dans § 6.1.

6.25. Champ tensoriel.

a. Si un tenseur de structure fixe, par exemple un tenseur $T_{ij}^k(x)$, est donné en tout point $x \in M_n$, et que ses composantes ont les dérivées en coordonnées jusqu'à un ordre m y compris, alors on dit qu'*un champ tensoriel* $T_{ij}^k(x)$ *m fois dérivable est donné sur la variété* M_n. Il résulte de la relation fondamentale

$$T_{i'j'}^{k'}(x) = p_{i'}^i p_{j'}^j p_k^{k'} T_{ij}^k(x)$$

et de la $(N-1)$-dérivabilité des fonctions $p^{i}_{i'}(x)$, $p^{j}_{j'}(x)$, $p^{k'}_{k}(x)$, que, pour $m \leqslant N-1$, la propriété du champ $T^{k}_{ij}(x)$ d'avoir les dérivées de m premiers ordres est une propriété absolue.

b. Soit $f(x)$ une fonction numérique donnée sur une variété M_n de classe N et $m < N$ fois dérivable par rapport aux coordonnées $x^1, \ldots, x^n$. On peut dire que la fonction $f(x)$ définit sur la variété M_n *un champ tensoriel de rang nul.*

Considérons en tout point $x \in M_n$ les quantités

$$\frac{\partial f(x)}{\partial x^1}, \frac{\partial f(x)}{\partial x^2}, \ldots, \frac{\partial f(x)}{\partial x^n};$$

nous affirmons qu'*elles déterminent un champ tensoriel de rang un, à savoir un champ une fois covariant d'ordre de dérivabilité* $m-1$. En effet, dans un autre système de coordonnées $x^{1'}, \ldots, x^{n'}$ on a

$$\frac{\partial f(x)}{\partial x^{i'}} = \frac{\partial f(x)}{\partial x^{i}} \frac{\partial x^{i}}{\partial x^{i'}} = p^{i}_{i'} \frac{\partial f(x)}{\partial x^{i}},$$

ce qui représente bien la loi de transformation des composantes d'un tenseur une fois covariant. Quant à l'ordre de dérivabilité, il est clair que la m-dérivabilité de la fonction $f(x)$ implique la $(m-1)$-dérivabilité des fonctions $\frac{\partial f(x)}{\partial x^{i}}$.

c. On peut considérer les champs tensoriels définis non pas sur toute la variété M_n, mais sur une ligne ou une surface dans M_n; la définition de l'ordre de dérivabilité du champ ne change pas.

On peut additionner les champs tensoriels d'un ordre de dérivabilité m donnés sur tout le M_n ou sur une même ligne ou surface, les multiplier l'un par l'autre et les contracter (en tout point); cela conduira à d'autres champs tensoriels du même ordre de dérivabilité m.

6.26. Pourtant la dérivation d'un tenseur par rapport à des coordonnées (le long de la ligne ou de la surface où il est donné) n'amène plus à des quantités tensorielles.

Soit $T^{j_1 \ldots j_m}_{i_1 \ldots i_k}(x)$ un champ tensoriel dérivable.

Trouvons la formule de transformation des quantités $\frac{\partial T^{j_1 \ldots j_m}_{i_1 \ldots i_k}}{\partial x^{r}}$.

En désignant $p^{i}_{i'r'} = \frac{\partial^2 x}{\partial x^{i'} \partial x^{r'}}$ (et les autres dérivées secondes de façon analogue) nous obtenons :

$$\frac{\partial T^{j'_1 \ldots j'_m}_{i'_1 \ldots i'_k}}{\partial x^{r'}} = \frac{\partial}{\partial x^{r'}} (p^{i_1}_{i'_1} \ldots p^{i_k}_{i'_k} p^{j'_1}_{j_1} \ldots p^{j'_m}_{j_m} T^{j_1 \ldots j_m}_{i_1 \ldots i_k}) =$$

$$= p^{i_1}_{i'_1} \dots p^{i_k}_{i'_k} p^{j'_1}_{j_1} \dots p^{j'_m}_{j_m} \frac{\partial T^{j_1 \dots j_m}_{i_1 \dots i_k}}{\partial x^r} p^r_{r'} + p^{i_1}_{i'_1 r'} p^{i_2}_{i'_2} \dots p^{j'_m}_{j_m} T^{j_1 \dots j_m}_{i_1 \dots i_k} + \dots$$

$$\dots + p^{i_1}_{i'_1} \dots p^{j'_{m-1}}_{j_{m-1}} p^{j'_m}_{j_m r} T^{j_1 \dots j_m}_{i_1 \dots i_k} p^r_{r'}. \qquad (1)$$

S'il n'y avait dans cette somme que le premier terme, la transformation des quantités $\frac{\partial T^{j_1 \dots j_m}_{i_1 \dots i_k}}{\partial x^r}$ se ferait d'après la loi tensorielle (avec un indice covariant r de plus). Mais la présence des termes contenant les dérivées secondes complique la loi de transformation. Le nombre de tels termes est égal au nombre des indices du tenseur initial. Chaque indice covariant fournit un terme avec un facteur de la forme $p^i_{i'r'}$, chaque indice contravariant un terme avec un facteur de la forme $p^{j'}_{jr} p^r_{r'}$. Par exemple, pour les dérivées d'un tenseur une fois contravariant ξ^i, on a

$$\frac{\partial \xi^{i'}}{\partial x^{k'}} = p^{i'}_i p^k_{k'} \frac{\partial \xi^i}{\partial x^k} + p^{i'}_{ik} p^k_{k'} \xi^i \qquad (2)$$

ou, vu que $dx^{k'} = p^{k'}_k \, dx^k$,

$$d\xi^{i'} = \frac{\partial \xi^{i'}}{\partial x^{k'}} dx^{k'} = p^{i'}_i \, d\xi^i + p^{i'}_{ik} \xi^i \, dx^k. \qquad (3)$$

Pour un tenseur deux fois covariant $g_{ij}(x)$, on a

$$\frac{\partial g_{i'j'}}{\partial x^{k'}} = p^i_{i'} p^j_{j'} p^k_{k'} \frac{\partial g_{ij}}{\partial x^k} + p^i_{i'k'} p^j_{j'} g_{ij} + p^i_{i'} p^j_{j'k'} g_{ij}. \qquad (4)$$

§ 6. 3. Espace riemannien élémentaire

6.31.a. Une variété différentiable élémentaire M_n s'appelle *espace riemannien élémentaire*, s'il est donné sur M_n un champ tensoriel deux fois covariant $g_{ij}(x)$ qui, à tout point, est symétrique ($g_{ij}(x) = g_{ji}(x)$) et défini positif. Le dernier veut dire que, pour tout tenseur une fois contravariant non nul $\xi^j(x)$, a lieu l'inégalité

$$g_{ij}(x) \, \xi^i(x) \, \xi^j(x) > 0.$$

b. Deux espaces riemanniens élémentaires M_1 et M_2 sont dits *équivalents* ou *isométriques* si l'on peut y introduire les coordonnées admissibles d'une telle manière que les quantités $g_{ij}(x)$ s'expriment par les mêmes fonctions des coordonnées sur M_1 et sur M_2. Nous avons déjà vu cette notion dans la théorie des surfaces.

c. Pour un espace riemannien élémentaire M_n, on peut définir le produit scalaire (*6.16*) de deux vecteurs tangents (tenseurs une

fois contravariants) $\xi = \{\xi^i(x)\}$ et $\eta = \{\eta^i(x)\}$ comme suit :

$$(\xi, \eta) = g_{ij}(x)\,\xi^i(x)\,\eta^j(x);$$

ce produit scalaire munit chaque espace tangent $T_n(x)$ d'une *métrique*, i.e. définit les longueurs des vecteurs et les angles qu'ils font.

En particulier, pour les vecteurs de base d'un espace tangent $T_n(x)$, nous avons

$$(e_i, e_j) = g_{ij},$$

l'égalité qui, plus haut, servait de *définition* aux nombres g_{ij}.

d. La métrique des espaces tangents permet de doter d'une métrique la variété M_n même. En effet l'élément d'arc d'une courbe $L = \{x \in M_n\}: x^i = x^i(t),\ a \leqslant t \leqslant b\}$ à un point $P \in L$ est défini par la formule

$$ds^2 = g_{ij}(P)\,dx^i\,dx^j = |\,dx^i e_i\,|^2.$$

Alors la longueur de toute la courbe L entre deux points A et B répondant aux valeurs du paramètre $t = a$ et $t = b$ respectivement sera calculée d'après la formule

$$s = \int_A^B \sqrt{g_{ij}(x)\,dx^i\,dx^j} = \int_a^b \sqrt{g_{ij}(x(t)\,\xi^i(x)\,\xi^j(x)}\,dt,$$

où $\xi^i(x) = \dfrac{dx^i(t)}{\partial t}$. Cette expression, en raison de son caractère tensoriel, ne dépend plus du système de coordonnées. Si la courbe L n'a pas de points singuliers (le tenseur $\xi^i(t)$ ne s'annule pas), alors la longueur d'arc s comptée du point A au point courant $P = P(t)$, en tant que fonction de t, a une dérivée non nulle ; il existe donc une fonction inverse $t = t(s)$; par conséquent, la courbe peut être déterminée, comme dans la géométrie différentielle classique, par le paramètre naturel s.

e. La mesure des aires et des volumes dans un espace riemannien se fait de la même façon que sur une surface dans l'espace euclidien R_n.

Soit, par exemple, $Q = \{x \in M_n : x^i = x^i(u, v),\ (u, v) \in \Omega \subset \subset R_2\}$ une surface bidimensionnelle. Alors les différentielles du et dv définissent, dans l'espace tangent $T_n(x)$, un parallélogramme élémentaire de côtés

$$\frac{\partial x}{\partial u}\,du = \frac{\partial x^i}{\partial u}\,du \cdot e_i \quad \text{et} \quad \frac{\partial x}{\partial v}\,dv = \frac{\partial x^i}{\partial v}\,dv \cdot e_i$$

et d'aire (*6.17a*)

$$dS^2 = \begin{vmatrix} \left(\frac{\partial x}{\partial u}\,du,\ \frac{\partial x}{\partial u}\,du\right) & \left(\frac{\partial x}{\partial u}\,du,\ \frac{\partial x}{\partial v}\,dv\right) \\ \left(\frac{\partial x}{\partial v}\,dv,\ \frac{\partial x}{\partial u}\,du\right) & \left(\frac{\partial x}{\partial v}\,dv,\ \frac{\partial x}{\partial v}\,dv\right) \end{vmatrix} = (EG - F^2)\,du^2\,dv^2, \qquad (1)$$

où

$$E=\left(\frac{\partial x}{\partial u}, \frac{\partial x}{\partial u}\right), \quad F=\left(\frac{\partial x}{\partial u}, \frac{\partial x}{\partial v}\right), \quad G=\left(\frac{\partial x}{\partial v}, \frac{\partial x}{\partial v}\right).$$

En intégrant dS sur le domaine Ω, nous obtenons l'aire de ce domaine :

$$S(Q)=\int\int_{\Omega} \sqrt{EG-F^2}\, du\, dv.$$

6.32. Translation. Définissons à présent la notion de translation d'un tenseur une fois contravariant ξ^i le long d'une courbe $L=\{x \in M_n,\ x^i=x^i(t),\ a \leqslant t \leqslant b\}$. Dans la théorie des surfaces, la translation était introduite à l'aide d'une construction géométrique ayant pour résultat l'équation différentielle *5.62* (1) à coefficients Γ^k_{ij} qui s'expriment d'une certaine façon par les quantités g_{ij}. Ici nous définissons la translation comme solution de l'équation différentielle

$$d\xi^k=-\Gamma^k_{ij}(x)\,\xi^i(x)\,dx^j, \tag{1}$$

où $\Gamma^k_{ij}(x)$ sont des fonctions de x. Ces fonctions sont choisies de façon à ce que le produit scalaire des tenseurs ξ^i et η^i (comme pour la translation des vecteurs sur une surface) ne change pas en les translatant le long de n'importe quelle ligne L, autrement dit, que la quantité

$$(\xi,\ \eta)=g_{ij}\xi^i\eta^j \tag{2}$$

reste constante le long de chaque ligne L. Pour déterminer les coefficients $\Gamma^k_{ij}(x)$, différentions (2) :

$$0=d(g_{ij}\xi^i\eta^j)=dg_{ij}\xi^i\eta^j+g_{ij}\,d\xi^i\eta^j+g_{ij}\xi^i\,d\eta^j.$$

En substituant $dg_{ij}=\frac{\partial g_{ij}}{\partial x^r}\,dx^r$, $d\xi^i$ de (1) et $d\eta^j$ d'une égalité analogue, faisons un changement des indices de sorte qu'il ne reste que les quantités dx^r, ξ^p, η^q :

$$0=\frac{\partial g_{pq}}{\partial x^r}\,\xi^p\eta^q\,dx^r-g_{iq}\Gamma^i_{pr}\xi^p\eta^q\,dx^r-g_{pj}\Gamma^j_{qr}\xi^p\eta^q\,dx^r.$$

Il en résulte, les champs ξ, η et le chemin L étant arbitraires :

$$\frac{\partial g_{pq}}{\partial x^r}=g_{iq}\Gamma^i_{pr}+g_{pj}\Gamma^j_{qr}.$$

En désignant

$$\Gamma_{pr,\,q}=g_{iq}\Gamma^i_{pr},$$

mettons la même égalité sous la forme :

$$\frac{\partial g_{pq}}{\partial x^r}=\Gamma_{pr,\,q}+\Gamma_{qr,\,p}. \tag{3}$$

Imposons aux quantités Γ_{ij}^{k} (et par cela aux quantités $\Gamma_{ij,\,k}$) *la condition* $\Gamma_{ij}^{k} = \Gamma_{ji}^{k}$, i.e. *la condition de symétrie par rapport aux indices* i, j; alors, étant donné les relations

$$\frac{\partial g_{ik}}{\partial x^j} = \Gamma_{ij,\,k} + \Gamma_{kj,\,i},$$

$$\frac{\partial g_{jk}}{\partial x^i} = \Gamma_{ji,\,k} + \Gamma_{ki,\,j},$$

$$\frac{\partial g_{ij}}{\partial x^k} = \Gamma_{ik,\,j} + \Gamma_{jk,\,i},$$

on en additionne les deux premières, y retranche la troisième, et on trouve

$$\Gamma_{ij,\,k} = \frac{1}{2}\left(\frac{\partial g_{ik}}{\partial x^j} + \frac{\partial g_{jk}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^k}\right). \tag{4}$$

Les coefficients $\Gamma_{ij,\,k}$ sont ainsi déterminés d'une façon univoque. Il en est de même des coefficients

$$\Gamma_{ij}^{k} = g^{ks}\Gamma_{ij,\,s}, \tag{5}$$

où $\| g^{ks} \|$ est, comme d'habitude, la matrice inverse de $\| g_{ij} \|$. Dans ce qui suit, nous choisissons les coefficients $\Gamma_{ij,\,k}$ et Γ_{ij}^{k} juste d'après les formules (4) et (5). Notons que, d'après ces formules, les quantités $\Gamma_{ij,\,k}$ et Γ_{ij}^{k} sont symétriques en i et j. Ceci garantit la possibilité de définir une translation le long de n'importe quelle ligne L selon les formules (1), i.e. un déplacement qui ne change pas la valeur du produit scalaire (2). En particulier, elle ne change pas les longueurs des vecteurs (des tenseurs une fois contravariants) ni les angles qu'ils forment.

En outre, l'équation différentielle (1) de la translation est une équation linéaire homogène, et l'on peut additionner ses solutions et les multiplier par les nombres en obtenant d'autres solutions. Il en résulte que si un vecteur $\xi(t)$ est translaté le long d'une courbe L, alors il en est de même du vecteur $C\xi(t)$, pour toute constante C; avec deux vecteurs $\xi(t)$ et $\eta(t)$ leur somme $\xi(t) + \eta(t)$ est translatée le long d'une courbe L.

6.33. Les coefficients $\Gamma_{ij,\,k}$ s'expriment par les dérivées du tenseur g_{ij}, donc ne se transforment pas suivant la loi tensorielle. Avec la formule *6.26* (4), on peut écrire

$$\Gamma_{i'j',\,k'} = \frac{1}{2}\left(\frac{\partial g_{i'k'}}{\partial x^{j'}} + \frac{\partial g_{j'k'}}{\partial x^{i'}} - \frac{\partial g_{i'j'}}{\partial x^{k'}}\right) =$$

$$= p_{i'}^{i} p_{j'}^{j} p_{k'}^{k} \cdot \frac{1}{2}\left(\frac{\partial g_{ik}}{\partial x^j} + \frac{\partial g_{jk}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^k}\right) +$$

$$+\frac{1}{2}(p^i_{i',j'}p^k_{k'}g_{ik}+p^i_{i'}p^k_{k',j'}g_{ik})+\frac{1}{2}(p^j_{j',i'}p^k_{k'}g_{jk}+p^j_{j'}p^k_{k',i'}g_{jk})-$$

$$-\frac{1}{2}(p^i_{i',k'}p^j_{j'}g_{ij}+p^i_{i'}p^j_{j',k'}g_{ij}).$$

En changeant les indices de sommation de façon à avoir partout les quantités g_{ik}, on trouve

$$\Gamma_{i'j',k'}=p^i_{i'}p^j_{j'}p^k_{k'}\Gamma_{ij,k}+p^i_{i',j'}p^k_{k'}g_{ik}.$$

On obtient maintenant pour les coefficients Γ^k_{ij}:

$$\Gamma^{k'}_{i'j'}=g^{k's'}\Gamma_{i'j',s'}=p^{k'}_k p^{s'}_s g^{ks}(p^i_{i'}p^j_{j'}p^r_{s'}\Gamma_{ij,r}+p^i_{i',j'}p^m_{s'}g_{im})=$$

$$=\delta^r_s p^{k'}_k p^i_{i'}p^j_{j'}g^{ks}\Gamma_{ij,r}+\delta^m_s p^{k'}_k g^{ks}g_{im}p^i_{i',j'}=$$

$$=p^i_{i'}p^j_{j'}p^{k'}_k\Gamma^k_{ij}+p^{k'}_k p^k_{i',j'}. \qquad (1)$$

Nous voyons que la formule de transformation n'est pas tensorielle: c'est le dernier terme contenant les dérivées secondes des coordonnées non primées par rapport aux coordonnées primées qui empêche. D'ailleurs, ce terme disparaît en cas d'une transformation *du premier degré*; par conséquent, *pour une transformation du premier degré, les quantités* Γ^k_{ij} *se transforment de même que les composantes d'un tenseur avec les indices correspondants* (deux fois co et une fois contravariant). A côté de la formule générale (1), il nous sera utile la formule pour la dérivée deuxième:

$$p^k_{i'j'}=p^k_{k'}\Gamma^{k'}_{i'j'}-p^i_{i'}p^j_{j'}\Gamma^k_{ij}. \qquad (2)$$

6.34. Différentielle absolue d'un tenseur une fois contravariant. Soit $\xi^i(t)$ le champ d'un tenseur une fois contravariant $\xi_i(t)$ donné le long d'une ligne $L\subset M_n$. Définissons sa *différentielle absolue* $D\xi^i$ par la formule

$$D\xi^k=d\xi^k+\Gamma^k_{ij}\xi^i\,dx^j.$$

En vertu de *6.32*, *l'égalité de la différentielle absolue au zéro signifie que le tenseur* ξ^i *est translaté le long de la ligne L.*

Trouvons la formule de transformation de la différentielle absolue pour le passage à un autre système de coordonnées. D'après *6.33.*, nous avons:

$$D\xi^{k'}=d\xi^{k'}+\Gamma^{k'}_{i'j'}\xi^{i'}\,dx^{j'}=p^{k'}_k\,d\xi^k+p^{k'}_{km}\xi^k\,dx^m+$$

$$+(p^i_{i'}p^j_{j'}p^{k'}_k\Gamma^k_{ij}+p^{k'}_k p^k_{i',j'})\,p^{i'}_r\xi^r p^{j'}_s\,dx^s=$$

$$=p^{k'}_k\,d\xi^k+p^{k'}_{km}\xi^k\,dx^m+\delta^i_r\delta^j_s p^{k'}_k\Gamma^k_{ij}\xi^r\,dx^s+$$

$$+ p_k^{k'} p_r^{i'} p_s^{j'} p_{i'j'}^{k} \xi^r \, dx^s = p_k^{k'} (d\xi^k + \Gamma_{ij}^k \xi^i \, dx^j) +$$
$$+ (p_{km}^{k'} + p_k^{i'} p_m^{j'} p_l^{k'} p_{i'j'}^{l}) \xi^k \, dx^m.$$

La parenthèse dans le deuxième terme se transforme en remplaçant les dérivées secondes d'après la formule *6.33* (2):

$$p_{km}^{k'} + p_k^{i'} p_m^{j'} p_l^{k'} p_{i'j'}^{l} = p_s^{k'} \Gamma_{km}^s - p_k^{i'} p_m^{j'} \Gamma_{i'j'}^{k'} +$$
$$+ p_k^{i'} p_m^{j'} p_l^{k'} (p_{s'}^{l} \Gamma_{i'j'}^{s'} - p_{i'}^{i} p_{j'}^{j} \Gamma_{ij}^{l}) = p_s^{k'} \Gamma_{km}^s - p_k^{i'} p_m^{j'} \Gamma_{i'j'}^{k'} +$$
$$+ \delta_{s'}^{k'} p_k^{i'} p_m^{j'} \Gamma_{i'j'}^{s'} - \delta_k^i \delta_m^j p_l^{k'} \Gamma_{ij}^{l} = 0,$$

et définitivement,

$$D\xi^{k'} = p_k^{k'} D\xi^k.$$

Cette égalité veut dire que la *différentielle absolue d'un tenseur une fois contravariant* (contrairement à la différentielle ordinaire) *est aussi un tenseur une fois contravariant.*

En conséquence, nous obtenons la proposition suivante: *si un tenseur une fois contravariant est translaté le long d'une courbe dans un système de coordonnées, il l'est dans tout autre système de coordonnées.*

6.35. Lignes géodésiques.

a. Définition. Une ligne $L = \{x \in M_n : x^i = x^i(t),\ a \leqslant t \leqslant b\}$ sur une variété M_n est dite *géodésique* si le vecteur unitaire tangent $\frac{dx^i}{ds} = \tau^i(s)$ est translaté le long de cette ligne.

Après ce qu'on a dit plus haut, il est clair que la définition d'une ligne géodésique a un caractère intrinsèque, ne dépendant pas du système de coordonnées.

b. En divisant l'équation *6.32* (1) par ds et en y substituant $\xi^k = \tau^k = \frac{dx^k}{ds}$ nous arrivons à l'équation différentielle des lignes géodésiques:

$$\frac{d^2x^k}{ds^2} = -\Gamma_{ij}^k(x) \frac{dx^i}{ds} \frac{dx^j}{ds}.$$

Il en résulte, comme dans *5.42*, le théorème fondamental d'existence et d'unicité des lignes géodésiques:

Théorème. *Si les coefficients* $\Gamma_{ij}^k(x)$ *sont continus en un point* $A \in M_n$, *alors c'est une géodésique et une seule qui passe par ce point dans chaque direction* (déterminée, par exemple, par un vecteur unitaire τ).

c. Ensuite, comme dans *5.44*, on peut reproduire la construction de la surface $(n-1)$-dimensionnelle géodésiquement parallèle à une surface $(n-1)$-dimensionnelle $\Pi_{n-1} \subset M_{n-1}$, et démontrer

qu'elle est orthogonale aux géodésiques coupant Π_{n-1} orthogonalement. Ce fait, à son tour, amène *à la propriété extrémale des lignes géodésiques: parmi toutes les courbes reliant deux points suffisamment proches de la variété M_n, c'est la géodésique qui a la moindre longueur.*

d. Pour une variété riemannienne bidimensionnelle M_2, on peut répéter la définition de la courbure formelle de *5.33a*:

$$K = \frac{\sum_{h=1}^{2} \left[\frac{\partial \Gamma_{11}^h}{\partial x^2} - \frac{\partial \Gamma_{12}^h}{\partial x^1} + \sum_{s=1}^{2} (\Gamma_{11}^s \Gamma_{s2}^h - \Gamma_{12}^s \Gamma_{s1}^h) \right] g_{h2}}{g_{11} g_{22} - g_{12}^2}$$

et, de même que dans *5.63*, établir sa relation avec la translation d'un vecteur le long d'un contour fermé. Ensuite, comme dans *5.51*, on peut établir l'isométrie locale d'une variété M_2 de courbure constante et de la surface canonique correspondante (plan, sphère, pseudo-sphère). Dans ce qui suit (§§ 6.5, 6.6), nous tenons à généraliser les définitions et les constructions correspondantes au cas n-dimensionnel. Mais il sera plus commode d'étudier préalablement la géométrie d'un espace plus abstrait, à savoir d'un espace à connexion affine.

§ 6. 4. Espace à connexion affine

6.41.a. Soit M_n une variété différentiable élémentaire, de dimension n et de classe N.

Nous tenons à définir, encore une fois, la translation d'un tenseur une fois contravariant ξ le long d'une ligne dérivable $L = \{x \in M_n : x^i = x^i(t),\ a \leqslant t \leqslant b\}$ comme solution de l'équation

$$d\xi^h = -\Gamma_{ij}^h(x)\, \xi^i\, dx^j. \tag{1}$$

Mais cette fois-ci, l'espace M_n n'est pas muni de tenseur métrique g_{ij}, donc on peut soumettre les coefficients $\Gamma_{ij}^k(x)$ à une seule condition, celle de l'indépendance du résultat d'une translation du système de coordonnées. Ceci nous fait imposer à la loi de transformation des quantités Γ_{ij}^k des conditions qui garantissent le caractère tensoriel du résultat d'une translation. Pour la métrique riemannienne, lorsque Γ_{ij}^k sont définis de façon univoque par la condition que le produit scalaire des vecteurs en translation reste constant et par la symétrie en i et j, nous avons (*6.33*) la formule de transformation

$$\Gamma_{i'j'}^{k'} = p_{i'}^i p_{j'}^j p_k^{k'} \Gamma_{ij}^k + p_k^{k'} p_{i'j'}^k \tag{2}$$

qui a pour conséquence, sans plus utiliser la métrique, le caractère tensoriel de la translation. On peut s'attendre à ce que la condition (2) soit non seulement suffisante, mais aussi bien nécessaire pour le

caractère tensoriel de la translation. Prouvons-le directement. Supposons que la solution de l'équation (1), pour n'importe quelles données initiales ξ^i (A), soit d'un caractère tensoriel. Alors, dans un nouveau système de coordonnées $x^{1'}, \ldots, x^{n'}$, nous avons

$$d\xi^k = -\Gamma^k_{ij}\xi^i\,dx^j = -\Gamma^k_{ij}p^i_{i'}p^j_{j'}\xi^{i'}\,dx^{j'},$$

or,

$$d\xi^k = d\,(p^k_{k'}\xi^{k'}_i) = dp^k_{k'}\xi^k + p^k_{k'}\,d\xi^{k'} = p^k_{k'm'}\,dx^{m'}\xi^{k'} - p^k_{k'}\Gamma^{k'}_{i'j'}\xi^{i'}\,dx^{j'} =$$
$$= (p^k_{i'j'} - p^k_{k'}\Gamma^{k'}_{i'j'})\,\xi^{i'}\,dx^{j'}.$$

En comparant les résultats et vu qu'ils sont valables pour n'importe quels $\xi^{i'}$ et $dx^{j'}$, on trouve

$$p^i_{i'}p^j_{j'}\Gamma^k_{ij} = p^k_{k'}\Gamma^{k'}_{i'j'} - p^k_{i'j'}, \tag{3}$$

ou, ce qui est le même,

$$\Gamma^{k'}_{i'j'} = p^i_{i'}p^j_{j'}p^{k'}_k\Gamma^k_{ij} + p^{k'}_k p^k_{i'j'}, \tag{4}$$

ce qu'il fallait démontrer.

Les nombres $\Gamma^k_{ij}(x)$ donnés dans tout système de coordonnées et se transformant d'après la règle (2) s'appellent *coefficients de connexion affine de la variété* M_n.

b. Deux variétés M_n et $\widetilde{M}_n$, avec Γ^k_{ij} et $\widetilde{\Gamma}^k_{ij}$ respectivement pour coefficients de connexion affine, sont dites *affinement équivalentes* si l'on peut les munir de systèmes de coordonnées (admissibles) d'une telle manière que les coefficients Γ^k_{ij} et $\widetilde{\Gamma}^k_{ij}$ soient les mêmes fonctions des coordonnées sur M_n et sur $\widetilde{M}_n$ respectivement.

c. On peut définir les coefficients $\Gamma^k_{ij}(x)$ arbitrairement dans un système de coordonnées, et d'après les formules (2) dans tout autre système. Le fait que cette définition est correcte (que les formules du type (2) restent valables en effectuant deux passages successifs à des coordonnées nouvelles) résulte du caractère tensoriel d'une translation et de l'unicité déjà établie des coefficients Γ^k_{ij} pour une translation donnée.

d. Une connexion $\Gamma = \{\Gamma^k_{ij}(x)\}$ sur une variété M_n est dite *symétrique* si

$$\Gamma^k_{ij}(x) \equiv \Gamma^k_{ji}(x)$$

dans chaque système de coordonnées. Il suffit que ces égalités aient lieu dans un seul système de coordonnées, parce qu'elles le seront dans tout autre système d'après les formules (2). En général, une connexion Γ n'est pas forcément symétrique, pour la raison, par exemple, qu'on peut définir les coefficients $\Gamma^k_{ij}(x)$ arbitrairement dans un système de coordonnées, en particulier, on peut les choisir

non symétriques. Dans le cas général, la différence

$$S_{ij}^{k}(x) = \Gamma_{ij}^{k}(x) - \Gamma_{ji}^{k}(x)$$

sera appelée *torsion* de la connexion Γ_{ij}^{k}. Les quantités S_{ij}^{k} forment déjà un *tenseur*, car

$$S_{i'j'}^{k'} = \Gamma_{i'j'}^{k'} - \Gamma_{j'i'}^{k'} = p_{i'}^{i} p_{j'}^{j} p_{k}^{k'} (\Gamma_{ij}^{k} - \Gamma_{ji}^{k}) + \\ + p_{k}^{k'} (p_{i'j'}^{k} - p_{j'i'}^{k}) = p_{i'}^{i} p_{j'}^{j} p_{k}^{k'} S_{ij}^{k}.$$

Et si $S_{ij}^{k} \equiv 0$, c'est-à-dire que la connexion Γ est symétrique, on dira que la connexion Γ *n'a pas de torsion*. L'interprétation géométrique de la torsion sera discutée plus bas, dans *6.45*. Il est évident qu'une variété M_n à connexion affine sans torsion *n'est pas équivalente* à une variété $\tilde{M}_n$ dont la connexion affine possède une torsion non nulle.

e. S'il y a sur une variété M_n une métrique riemannienne $g_{ij}(x)$ et que sa connexion affine est construite d'après les formules *6.32* (4), (5) :

$$\Gamma_{ij}^{k} = g^{ks}\Gamma_{ij,\,s} = \frac{1}{2} g^{ks} \left(\frac{\partial g_{is}}{\partial x^{j}} + \frac{\partial g_{js}}{\partial x^{i}} - \frac{\partial g_{ij}}{\partial x^{s}} \right),$$

alors cette connexion est dite *riemannienne*.

Mais une connexion affine peut bien exister sans aucune métrique riemannienne. Par exemple, une connexion riemannienne est toujours symétrique par rapport aux indices inférieurs ($\Gamma_{ij}^{k}(x) = \Gamma_{ji}^{k}(x)$), ce qui n'est pas le cas pour une connexion quelconque, comme nous l'avons vu.

6.42.a. Dans un espace à connexion affine, on ne peut plus dire que le produit scalaire des vecteurs en translation reste constant. Mais les propriétés linéaires de tels vecteurs ne sont pas atteintes : s'il y a l'égalité

$$\xi(A) = \alpha\eta(A) + \beta\zeta(A),$$

elle reste valable après avoir translaté tous les trois vecteurs ξ, η, ζ le long de n'importe quelle courbe L passant par le point A :

$$\xi(t) = \alpha\eta(t) + \beta\zeta(t).$$

Ce fait résulte immédiatement de la linéarité de l'équation d'une translation *6.32* (1).

En tant que conséquence, nous avons : *les vecteurs linéairement indépendants* ξ, η, . . . *restent linéairement indépendants après une translation.*

b. Dans un espace à connexion affine, on ne peut dire non plus qu'une translation conserve la longueur du vecteur. Contrairement

à ce qu'on avait pour un espace riemannien, dans un espace à connexion affine un point A peut admettre un voisinage arbitrairement petit, dans lequel les composantes d'un vecteur ξ deviennent arbitrairement grandes à la suite d'une translation (cf. exercice 5). On ne peut qu'affirmer : *si l'on choisit pour paramètre sur une courbe L une quantité s telle que* $|ds| \geqslant C \max_j |dx^j|$ (par exemple, l'« arc formel » défini par la condition $ds^2 = \sum_{j=1}^{n} (dx^j)^2$), *alors, dans les limites de l'intervalle* $0 \leqslant s \leqslant h$, *où h est un nombre fixe, les composantes d'un vecteur ξ translaté seront majorées par une quantité qui ne dépend pas du choix de l'arc L, mais seulement de h.* En abordant la démonstration posons $\sigma = \sum_{k=1}^{n} (\xi^k)^2$, de sorte que

$$d\sigma = 2 \sum_{k=1}^{n} \xi^k \, d\xi^k = -2 \sum_{i,\, j,\, k=1}^{n} \xi^k \Gamma_{ij}^k \xi^i \frac{dx^j}{ds} \, ds.$$

Alors, si les quantités $|\Gamma_{ij}^k(x)|$ sont majorées par une constante C_1, on obtient

$$\left| \sum_{j=1}^{n} \Gamma_{ij}^k(x) \frac{dx^j}{ds} \right| \leqslant C^{-1} \sum_{j=1}^{n} |\Gamma_{ij}^k(x)| \leqslant nC_1C^{-1} = C_2,$$

$$\sum_{i,\, k=1}^{n} |\xi^i \xi^k| \leqslant \frac{1}{2} \sum_{i,\, k=1}^{n} [(\xi^i)^2 + (\xi^k)^2] \leqslant n\sigma,$$

donc

$$|d\sigma| \leqslant C_3 \sigma \, ds, \quad C_3 = nC_2,$$

ou bien

$$\left| \frac{d \ln \sigma}{ds} \right| \leqslant C_3.$$

Il en résulte

$$\sigma \leqslant \sigma_0 e^{C_3 s} \leqslant \sigma_0 e^{C_3 h}, \quad \sigma_0 = \sum_{j=1}^{n} [\xi^j(A)]^2, \tag{1}$$

d'où notre affirmation.

6.43. Considérons la connexion riemannienne de l'espace vectoriel euclidien $M_n = R_n$. Il existe dans cet espace un système de coordonnées dans lequel la forme métrique est $(dx^1)^2 + \ldots + (dx^n)^2$, de sorte que tous les coefficients $g_{ij}(x)$ sont constants (zéros et unités), donc tous les coefficients de connexion $\Gamma_{ij}^k(x)$ sont nuls. Dans les autres systèmes de coordonnées, les coefficients $\Gamma_{ij}^k(x)$ sont, en général, non nuls (rappelons que la loi de transformation de Γ_{ij}^k

n'est pas tensorielle). Tout de même, dans les autres systèmes de coordonnées de l'espace euclidien R_n, il y a aussi quelque chose de spécifique qui n'a pas lieu dans les variétés qui ne lui sont pas équivalentes. Premièrement, les coefficients $\Gamma_{ij}^k(x)$ sont symétriques par rapport aux indices inférieurs, comme pour chaque connexion riemannienne. Deuxièmement, a lieu une propriété appelée *parallélisme absolu*: le résultat de la translation d'un vecteur ξ ne dépend pas du chemin, mais seulement de son point de départ et de son point d'arrivée; autrement dit, après une translation le long d'un contour fermé, le vecteur ξ reprend sa position initiale. En effet, dans un système de coordonnées, où $\Gamma_{ij}^k(x) \equiv 0$, cela résulte de la définition d'une translation en tant que solution du système d'équations

$$d\xi^k = -\Gamma_{ij}^k(x)\,\xi^i\,dx^j,$$

qui, dans ledit système de coordonnées, se réduit à la forme

$$d\xi^k \equiv 0$$

ayant pour solution

$$\xi^k = \text{const};$$

ce qui signifie que les composantes d'un vecteur restent constantes en subissant une translation. Dans tout autre système de coordonnées, notre affirmation est valable en vertu du caractère absolu d'une translation: si un vecteur revient à sa position initiale après être translaté le long d'un contour fermé dans un système de coordonnées, il en sera de même de tout autre système de coordonnées.

Montrons que c'est seul l'espace euclidien qui possède les propriétés mentionnées de la connexion:

Théorème. *Si, pour une variété M_n de dimension n et de classe N, la connexion affine $\Gamma_{ij}^k(x)$ est symétrique, $N-2$ fois dérivable et amène au parallélisme absolu, alors la variété M_n est affinement équivalente à l'espace euclidien R_n.*

Démonstration. Nous allons trouver un nouveau système de coordonnées $x^{1'}, \ldots, x^{n'}$ de l'espace M_n dans lequel $\Gamma_{i'j'}^{k'}(x) \equiv 0$.

Choisissons une base $e_1, \ldots, e_n$ de l'espace tangent en un point fixe $A \in M_n$ et translatons tous ces vecteurs à un autre point $B \in M_n$. Les résultats seront déterminés d'une façon univoque en vertu de l'hypothèse du parallélisme absolu et formeront une base du plan tangent au point B, car une translation ne viole pas l'indépendance linéaire (*6.42*). Désignons les composantes des vecteurs obtenus (dans le système initial de coordonnées) par $\underset{m}{\xi^k}(x^1, \ldots, x^n)$, $m = 1, \ldots, n$. On a

$$\underset{m}{d\xi^k} = -\Gamma_{ij}^k(x) u^i(x)\,dx^j,$$

ou bien

$$\frac{\partial \underset{m}{\xi^k}}{\partial x^j} = -\Gamma^k_{ij}(x)\,\underset{m}{\xi^i}(x). \tag{1}$$

Pour construire les nouvelles coordonnées $x^{1'}, \ldots, x^{n'}$, considérons d'abord le système d'équations différentielles

$$\frac{\partial x^k}{\partial x^{s'}} = \underset{s}{\xi^k}(x^1, \ldots, x^n) \quad (k, s' = 1, \ldots, n), \tag{2}$$

où $x^{1'}, \ldots, x^{n'}$ sont, pour le moment, les variables indépendantes formelles. Nous montrerons que ce système, avec les données initiales

$$x^k(0, \ldots, 0) = x^k(A) \quad (k = 1, \ldots, n), \tag{3}$$

admet une unique solution, de sorte que, dans un voisinage de l'origine des coordonnées de l'espace $(x^{1'}, \ldots, x^{n'})$, soient définies les fonctions

$$x^1 = \varphi^1(x^{1'}, \ldots, x^{n'}), \ldots, x^n = \varphi^n(x^{1'}, \ldots, x^{n'}) \tag{4}$$

satisfaisant au système (2) et aux conditions (3). Nous verrons que l'inégalité $\det \left\| \frac{\partial x^k(0)}{\partial x^{j'}} \right\| \neq 0$ a lieu; on pourra alors inverser les équations (4) au voisinage du point A:

$$\begin{aligned} x^{1'} &= \psi^1(x^1, \ldots, x^n), \\ &\cdots\cdots\cdots\cdots \\ x^{n'} &= \psi^n(x^1, \ldots, x^n), \end{aligned}$$

et, de cette manière, faire correspondre à tout point $(x^1, \ldots, x^n)$ du voisinage de A les nombres $x^{1'}, \ldots, x^{n'}$; ceci dit, il est clair que ces nombres pourrons servir de nouvelles coordonnées.

En abordant la réalisation de notre plan, prouvons l'existence et l'unicité de solution du problème (2)-(3). Il suffit de vérifier l'hypothèse du théorème de Frobenius *2.55*, qui dans notre cas a la forme

$$\frac{\partial \underset{s}{\xi^k}}{\partial x^p}\,\underset{r}{\xi^p} \equiv \frac{\partial \underset{r}{\xi^k}}{\partial x^p}\,\underset{s}{\xi^p} \quad (k, r, s = 1, \ldots, n). \tag{5}$$

Or, en vertu de (1),

$$\frac{\partial \underset{s}{\xi^k}}{\partial x^p}\,\underset{r}{\xi^p} = -\Gamma^k_{ip}\,\underset{s}{\xi^i}\,\underset{r}{\xi^p}, \qquad \frac{\partial \underset{r}{\xi^k}}{\partial x^p}\,\underset{s}{\xi^p} = -\Gamma^k_{ip}\,\underset{r}{\xi^i}\,\underset{s}{\xi^p},$$

et, comme $\Gamma^k_{ip} = \Gamma^k_{pi}$, on a

$$\Gamma^k_{ip}\,\underset{s}{\xi^i}\,\underset{r}{\xi^p} = \Gamma^k_{pi}\,\underset{s}{\xi^i}\,\underset{r}{\xi^p} = \Gamma^k_{ip}\,\underset{s}{\xi^p}\,\underset{r}{\xi^i},$$

la condition (5) étant ainsi satisfaite.

Il existe donc un système de fonctions $x^h = x^h(x_k^{1'}, \ldots, x_k^{n'})$ qui vérifie les équations (2) et les conditions initiales (3). Ces fonctions ont une dérivée de plus que les fonctions $\underset{s}{\xi^h}(x^1, \ldots, x^n)$, donc deux dérivées de plus que les fonctions $\Gamma_{ij}^k(x)$; autrement dit, les fonctions $x^h(x^{1'}, \ldots, x^{n'})$ sont dérivables N fois au moins. De plus $\det \left\| \frac{\partial x^h}{\partial x^{s'}} \right\| = \det \| \underset{s}{\xi^h}(x) \| \neq 0$, ce qui démontre que les quantités $x^{1'}, \ldots, x^{n'}$ peuvent servir, localement au moins, de nouvelles coordonnées sur la variété M. Dans ces nouvelles coordonnées, on a

$$\underset{s}{\xi^{h'}} = p_h^{h'} \underset{s}{\xi^h} = \frac{\partial x^{h'}}{\partial x^h} \underset{s}{\xi^h} = \frac{\partial x^{h'}}{\partial x^h} \frac{\partial x^h}{\partial x^{s'}} = \delta_{s'}^{h'},$$

c'est-à-dire que les vecteurs que nous avons translatés sont de base. La condition de translation est

$$d\underset{s}{\xi^{h'}} = -\Gamma_{i'j'}^{h'} \underset{s}{\xi^{i'}} dx^{j'}.$$

Comme $\underset{s}{\xi^{h'}} = \delta_{s'}^{h'}$, on obtient

$$\Gamma_{s'j'}^{h'} dx^{j'} = 0,$$

d'où, $dx^{j'}$ étant arbitraires,

$$\Gamma_{s'j'}^{h'} \equiv 0,$$

ce qu'il nous fallait.

6.44. Lignes géodésiques d'une variété à connexion affine.

a. La définition d'une ligne géodésique que nous avons adoptée pour l'espace riemannien prend la forme suivante pour l'espace à connexion affine : *une ligne géodésique est une courbe dont un vecteur tangent reste tangent après l'avoir translaté à n'importe quel point de la courbe.*

Cette définition est, évidemment, intrinsèque.

Soit $L = \{x \in M_n : x^i = x^i(t), a \leqslant t \leqslant b\}$ une ligne géodésique issue d'un point A. Soient $a^i = \frac{dx^i(a)}{dt}$, et $\xi^i(t)$ le translaté de a^i le long de la courbe L au point répondant à t. Alors

$$\xi^i(t) = \lambda(t) \frac{dx^i(t)}{dt},$$

où $\lambda(t)$ est un facteur numérique. Introduisons un nouveau paramètre τ sur la ligne L de façon à avoir $dt = \lambda(t)\, d\tau$, $\tau(A) = 0$; en

passant à τ, nous avons

$$\xi^i(\tau) = \frac{dx^i(\tau)}{d\tau},$$

c'est-à-dire que, pour la courbe L paramétrisée par τ, c'est déjà le vecteur tangent $\frac{dx^i(\tau)}{d\tau}$ qui *est translaté*. Appelons τ *paramètre canonique* sur la géodésique L. En portant $\xi^i = \frac{dx^i}{d\tau}$ dans l'équation de la translation et divisant par $d\tau$, nous arrivons à l'*équation canonique d'une ligne géodésique*:

$$\frac{d^2x^k}{d\tau^2} = -\Gamma^k_{ij}(x)\frac{dx^i}{d\tau}\frac{dx^j}{d\tau} \quad (k = 1, \ldots, n). \tag{1}$$

b. Nous sommes maintenant en mesure de généraliser aux espaces à connexion affine le théorème d'existence et d'unicité des lignes géodésiques:

Théorème. *Dans un espace M_n à connexion affine dont les coefficients $\Gamma^k_{ij}(x)$ sont continus, par tout point A et dans chaque direction il passe une et une seule géodésique.*

Démonstration. Fixons un point A et une direction qui, dans un système de coordonnées donné, est déterminée par un vecteur b^i. Résolvons le système (1) avec les conditions initiales $x^i(0) = x^i(A)$, $\frac{dx^i(0)}{d\tau} = b^i$. Il est affirmé qu'aux fonctions $x^i = x^i(\tau)$ qu'on obtient en résolvant il correspond une ligne géodésique sur M_n. En effet, le vecteur $\frac{dx^i}{d\tau}$ tangent à cette ligne est translaté le long d'elle en raison des équations (1), ce qui signifie que la ligne est géodésique.

Supposons qu'il y ait une autre géodésique $\widetilde{L}$ passant par le point A dans la même direction, avec son paramètre canonique τ. Elle satisfait à l'équation (1) avec les conditions initiales

$$\widetilde{x}^i(0) = \widetilde{x}^i(A), \qquad \frac{d\widetilde{x}^i(0)}{d\tau} = \lambda b^i. \tag{2}$$

Or, en partant de la solution $x^i(\tau)$ dont nous disposons déjà, on peut obtenir d'une manière évidente une solution $\widetilde{\widetilde{x}}^i(\tau)$ qui satisfasse aux conditions (2); cela se fait d'après la formule $\widetilde{\widetilde{x}}^i(\tau) = x^i(\lambda\tau)$. En vertu du théorème d'unicité, on a $\widetilde{x}^i(\tau) \equiv \widetilde{\widetilde{x}}^i(\tau) \equiv x^i(\lambda\tau)$; ainsi, la courbe qui répond à l'équation $x^i = \widetilde{x}^i(\tau)$ est la même courbe L.

6.45. Interprétation géométrique de la torsion d'une connexion affine. Soit M_n un espace à connexion affine $\Gamma^k_{ij}(x)$ qui possède, en général, une torsion non

nulle $S_{ij}^k = \Gamma_{ij}^k - \Gamma_{ji}^k$. Considérons deux géodésiques L_ξ et L^η qui partent d'un point fixe A dans les directions déterminées par les vecteurs linéairement indépendants $\xi = \{\xi^i\}$ et $\eta = \{\eta^i\}$ (fig. 6.4-1).

Soit τ un paramètre canonique sur la géodésique L_ξ tel que $\frac{dx^i(A)}{d\tau} = \xi^i$ et θ un paramètre canonique sur la géodésique L_η tel que $\frac{dx^i(A)}{d\theta} = \eta^i$. Translatons le vecteur η le long de la géodésique L_ξ du point A au point B déterminé par une valeur du paramètre $\tau = \rho > 0$; désignons le résultat par $\widetilde{\eta}$. Considérons la géodésique $L_{\widetilde{\eta}}$ qui part du point B dans la direction de ce vecteur $\widetilde{\eta}$. D'une manière analogue, translatons le vecteur ξ le long de la géodésique L_η du point A au point D déterminé par la même valeur $\theta = \rho$ du paramètre et considérons la géodésique $L_{\widetilde{\xi}}$ passant par D dans la direction du vecteur obtenu $\widetilde{\xi}$. Sur la géodésique $L_{\widetilde{\eta}}$, introduisons un paramètre canonique $\widetilde{\theta}$ et, sur la géodésique $L_{\widetilde{\xi}}$, un paramètre canonique $\widetilde{\tau}$ de façon à avoir

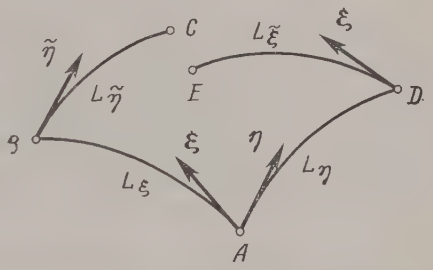


Fig. 6.4-1

$$\frac{dx^i(D)}{d\widetilde{\tau}} = \widetilde{\xi}, \quad \frac{dx^i(B)}{d\widetilde{\theta}} = \widetilde{\eta}.$$

Enfin, trouvons sur la géodésique $L_{\widetilde{\eta}}$ un point C déterminé par la valeur $\widetilde{\theta} = \rho$, et, sur la géodésique $L_{\widetilde{\xi}}$, un point E déterminé par la valeur $\widetilde{\tau} = \rho$.

Si l'on faisait cette construction dans l'espace vectoriel R_n (à connexion nulle), on arriverait à un parallélogramme, et les points C et E coïncideraient. Dans le cas général ($\Gamma_{ij}^k \neq 0$), ils s'avèrent distincts; évaluons leur écart.

L'accroissement de la coordonnée x^k le long de chaque ligne géodésique peut être écrit comme suit:

$$\Delta x^k = \frac{\partial x^k}{d\tau}\tau + \frac{1}{2}\frac{\partial^2 x^k}{\partial \tau^2}\tau^2 + o(\tau^2) = \\ = \frac{\partial x^k}{\partial \tau}\tau - \frac{1}{2}\Gamma_{ij}^k \frac{\partial x^i}{\partial \tau}\frac{\partial x^j}{\partial \tau}\tau^2 + o(\tau^2), \qquad (1)$$

τ étant le paramètre canonique, les valeurs des dérivées et des coefficients Γ_{ij}^k se rapportant au point initial.

En particulier, pour le passage du point A au point B, on a

$$\Delta_{AB}(x^k) = \xi^k \rho - \frac{1}{2}\Gamma^k_{ij}(A)\,\xi^i\xi^j\rho^2 + o(\rho^2), \tag{2}$$

et pour le passage de B à C,

$$\Delta_{BC}(x^k) = \widetilde{\eta}^k \rho - \frac{1}{2}\Gamma^k_{ij}(B)\,\widetilde{\eta}^i\widetilde{\eta}^j\rho^2 + o(\rho^2). \tag{3}$$

Ici $\widetilde{\eta}$ est le translaté du vecteur η :

$$\widetilde{\eta}^k = \eta^k - \Gamma^k_{ij}(A)\,\eta^i\,dx^j + o(dx^j). \tag{4}$$

Comme on fait le calcul aux infiniment petits du deuxième ordre près, et qu'on a encore une multiplication par ρ à effectuer en portant (4) dans (3), il est possible de se borner aux infiniment petits du premier ordre dans (4) et, en particulier, de remplacer dx^j par la quantité $\Delta_{AB}x^j$ qui, en raison de (2), vaut $\xi^j\rho$ aux infiniment petits du deuxième ordre près. En outre, on peut remplacer dans la formule (3) $\Gamma^k_{ij}(B)$ et $\widetilde{\eta}^i$ par $\Gamma^k_{ij}(A)$ et η^i respectivement. On aura :

$$\Delta_{BC}x^k = \eta^k\rho - \Gamma^k_{ij}\eta^i\xi^j\rho^2 - \frac{1}{2}\Gamma^k_{ij}\eta^i\eta^j\rho^2 + o(\rho^2). \tag{5}$$

Quant à l'accroissement total de la coordonnée x^k le long du chemin ABC, il sera de la forme

$$\Delta_{ABC}x^k = \Delta_{AB}x^k + \Delta_{BC}x^k = (\xi^k + \eta^k)\,\rho - \Gamma^k_{ij}\eta^i\xi^j\rho^2 - \\ - \frac{1}{2}\Gamma^k_{ij}\xi^i\xi^j\rho^2 - \frac{1}{2}\Gamma^k_{ij}\eta^i\eta^j\rho^2 + o(\rho^2). \tag{6}$$

Un résultat analogue, pour l'accroissement de la coordonnée x^k le long du chemin ADE, s'obtient en remplaçant les coordonnées de ξ par les coordonnées de η :

$$\Delta_{ADE}x^k = \Delta_{AD}x^k + \Delta_{DE}x^k = (\eta^k + \xi^k)\,\rho - \\ - \Gamma^k_{ij}\xi^i\eta^j\rho^2 - \frac{1}{2}\Gamma^k_{ij}\eta^i\eta^j\rho^2 - \frac{1}{2}\Gamma^k_{ij}\xi^i\xi^j\rho^2 + o(\rho^2). \tag{7}$$

On voit donc que la différence des coordonnées x^k aux points C et E vaut

$$x^k(E) - x^k(C) = \Gamma^k_{ij}(\xi^j\eta^i - \xi^i\eta^j)\,\rho^2 + o(\rho^2) = \\ = (\Gamma^k_{ij} - \Gamma^k_{ji})\,\xi^j\eta^i\rho^2 + o(\rho^2) = S^k_{ij}\xi^j\eta^i\rho^2 + o(\rho^2)$$

et est déterminée, dans sa partie principale, par le tenseur de torsion $S^k_{ij}(A)$.

6.46. Translation d'un tenseur arbitraire.

a. La translation d'un tenseur une fois contravariant $\{\xi^i\}$ d'un point A le long d'une courbe $L = \{x \in M_n : x^i = x^i(t),\ a \leqslant t \leqslant b\}$

est définie, nous l'avons vu, par l'équation

$$d\xi^k = -\Gamma^k_{ij}(x)\,\xi^k\,dx^j, \tag{1}$$

avec les valeurs $\xi^i(A)$ connues.

Soit $\{\eta_i(A)\}$ un tenseur une fois covariant. Définissons sa translation le long de la même courbe L à partir de la condition que l'invariant $\xi^k(x)\,\eta_k(x)$ reste constant, quel que soit $\{\xi^k\}$. Cela veut dire que $d(\xi^k\eta_k) = 0$, ou bien

$$\xi^k d\eta_k + d\xi^k\eta_k = 0. \tag{2}$$

En substituant à $d\xi^k$ sa valeur de l'équation (1) on trouve (on remplace dans le premier terme l'indice de sommation k par i)

$$\xi^i\,d\eta_i - \Gamma^k_{ij}\xi^i\,dx^j\eta_k = 0,$$

et, comme l'égalité a lieu pour tout vecteur ξ^i,

$$d\eta_i = \Gamma^k_{ij}\eta_k\,dx^j. \tag{3}$$

Cette équation, avec la condition initiale $\eta_i = \eta_i(A)$, permet de trouver $\eta_i(x)$ en tout point de la ligne L (au voisinage du point A).

Inversement, si les fonctions $\eta_i(x)$ sont trouvées de l'équation (3), l'équation (2) est également satisfaite, d'où la conservation de la quantité $\xi^k(x)\,\eta_k(x)$, et la condition de translation que nous avons formulée est ainsi satisfaite.

Montrons enfin que l'ensemble des quantités $\eta_i(x)$ déterminé dans chaque système de coordonnées est d'un caractère tensoriel. Par hypothèse, les quantités $\eta_i(A)$ forment un tenseur, donc la valeur $\xi^k(A)\,\eta_k(A)$ ne dépend pas du système de coordonnées. D'après ce qu'on a démontré, elle reste constante le long de la courbe L, donc la valeur $\xi^k(x)\,\eta_k(x)$ ne dépend pas du système de coordonnées aux points de la courbe L. Dans ce cas, en résolvant le système d'équations linéaires

$$\underset{1}{\xi^k}(x)\,\eta_k(x) = \underset{1}{\xi^k}(A)\,\eta_k(A),$$

$$\cdots\cdots\cdots\cdots$$

$$\underset{n}{\xi^k}(x)\,\eta_k(x) = \underset{n}{\xi^k}(A)\,\eta_k(A),$$

avec n tenseurs linéairement indépendants $\underset{1}{\xi^k}(x), \ldots, \underset{n}{\xi^k}(x)$, nous aboutissons à un tenseur une fois covariant $\{\eta_1(x), \ldots \eta_n(x)\}$ (*6.14c*), ce qu'il fallait.

b. On aurait pu commencer par la définition de la translation d'un tenseur une fois covariant η_i d'après la formule (3) et puis définir la translation d'un tenseur une fois contravariant en partant de la condition de conservation de l'invariant $\xi^i\eta_i$. On aboutirait

alors, bien entendu, à la formule (1) pour la translation du tenseur ξ^i.

c. Définissons d'une manière analogue la translation le long de la courbe L de n'importe quel tenseur T. Supposons, pour fixer les idées, que le tenseur T soit défini par les composantes T^k_{ij}, i.e. ait deux indices covariants et un contravariant. Alors, étant donné deux tenseurs une fois contravariants ξ^i, η^j et un tenseur une fois covariant ζ_k, on compose l'invariant $T^k_{ij}\xi^i\eta^j\zeta_k$. Définissons la translation du tenseur T^k_{ij} par la condition de conservation de cet invariant, ce qui amène à l'égalité $d\,(T^k_{ij}\xi^i\eta^j\zeta_k) = 0$, soit

$$dT^k_{ij}\xi^i\eta^j\zeta_k + T^k_{ij}\cdot d\xi^i\cdot\eta^j\zeta_k + T^k_{ij}\xi^i\cdot d\eta^j\cdot\zeta_k + T^k_{ij}\xi^i\eta^j\,d\zeta_k = 0.$$

En substituant aux différentielles $d\xi^i$, $d\eta^j$, $d\zeta_k$ leurs expressions des formules correspondantes de la translation des tenseurs une fois contra et covariants (1) et (3) on obtient

$$dT^k_{ij}\xi^i\eta^j\zeta_k - T^k_{ij}\Gamma^i_{pq}\xi^p\,dx^q\eta^j\zeta_k - T^k_{ij}\xi^i\Gamma^j_{pq}\eta^p\,dx^q\zeta_k + T^k_{ij}\xi^i\eta^j\Gamma^r_{kq}\zeta_r\,dx^q = 0.$$

Faisons changer les indices de sommation de façon à avoir partout les quantités $\xi^i\eta^j\zeta_k$; comme elles sont arbitraires, on obtient les relations

$$dT^k_{ij} = (\Gamma^s_{iq}T^k_{sj} + \Gamma^s_{jq}T^k_{is} - \Gamma^k_{sq}T^s_{ij})\,dx^q. \qquad (4)$$

D'une manière analogue, pour un tenseur de structure arbitraire $T^{k_1\ldots k}_{i_1\ldots i_r p}$, on a

$$dT^{k_1\ldots k_p}_{i_1\ldots i_r} = (\Gamma^s_{i_1 q}T^{k_1\ldots k_p}_{s i_2\ldots i_r} + \ldots + \Gamma^s_{i_r q}T^{k_1\ldots k_p}_{i_1\ldots i_{r-1}s} - \\ - \Gamma^{k_1}_{sq}T^{s k_2\ldots k_p}_{i_1\ldots i_r} - \ldots - \Gamma^{k_p}_{sq}T^{k_1\ldots k_{p-1}s}_{i_1\ldots i_r})\,dx^q. \qquad (5)$$

La structure des termes de l'expression obtenue est la suivante. Le nombre des termes est égal au nombre total des indices du tenseur T. Dans chaque terme, le deuxième indice inférieur de Γ est le même, q, il coïncide avec l'indice du facteur dx^q. Tous les indices du tenseur T dans le deuxième membre sont les mêmes que ses indices respectifs dans le premier membre, à l'exception d'un seul qui est remplacé par l'indice de sommation s; le symbole Γ est muni par le même indice de sommation, il est supérieur si s est l'indice inférieur de T, il est premier inférieur si s est l'indice supérieur de T. L'indice libre qui est remplacé par s occupe la seule place qui reste au symbole Γ.

Inversement, si la translation d'un tenseur $T^k_{ij}\,(A)$ est définie par la formule (4), a lieu l'égalité $d\,(T^k_{ij}\xi^i\eta^j\zeta_k) = 0$ et la quantité $T^k_{ij}\xi^i\eta^j\zeta_k$ ne varie pas sur la courbe L en translatant ξ^i, η^j, ζ_k.

Ceci amène au champ de la quantité $T^k_{ij}(x)$ dont le caractère tensoriel est prouvé, par exemple, comme suit : la quantité $T^k_{ij}\xi^i\eta^j\zeta_k$ étant constante le long de la courbe L dans tout système de coordonnées, on a

$$T^{k'}_{i'j'}(x)\,\xi^{i'}\eta^{j'}\zeta_{k'} = T^{k'}_{i'j'}(x)\,p^{i'}_i p^{j'}_j p^k_{k'}\xi^i\eta^j\zeta_k =$$

$$= T^{k'}_{i'j'}(A)\,\xi^{i'}\eta^{j'}\zeta_{k'} = T^k_{ij}(A)\,\xi^i\eta^j\zeta_k = T^k_{ij}(x)\,\xi^i\eta^j\zeta_k,$$

et, ξ^i, η^j, ζ_k étant arbitraires, on obtient

$$T^{k'}_{i'j'}(x)\,p^i_{i'}p^{j'}_j p^k_{k'} = T^k_{ij}(x),$$

ce qu'il fallait.

d. La définition précédente ne marche pas pour un tenseur du rang nul, i.e. pour un nombre T donné à un point A et ne dépendant pas du système de coordonnées. Nous définissons son translaté à n'importe quel point $B \in M_n$ comme *le même nombre T donné au point B dans tout système de coordonnées.*

e. Nous dirons qu'un champ tensoriel $T(x)$ donné sur une ligne $L \subset M_n$ est invariant par rapport à la translation le long de cette ligne si, pour tout point $A \in L$, le translaté du tenseur $T(A)$ à n'importe quel point $x \in L$ coïncide avec $T(x)$. Si un champ $T(x)$ est donné sur toute la variété M_n et invariant par rapport à la translation le long de chaque ligne L, nous dirons que *le champ $T(x)$ est invariant par rapport à la translation sur M_n.*

Le champ d'une constante T, i.e. d'un tenseur du rang nul, représente le plus simple exemple d'un champ invariant par rapport à la translation sur M_n.

Un autre exemple est fourni par le champ du tenseur mixte $\delta^j_i(x)$ dont les composantes, dans tout système de coordonnées et en tout point $x \in M_n$, sont 0 pour $i \neq j$ et 1 pour $i = j$. En effet, on a d'après *c*:

$$d\delta^j_i(x) = (\Gamma^s_{iq}\delta^j_s - \Gamma^j_{sq}\delta^s_i)\,dx^q = (\Gamma^j_{iq} - \Gamma^j_{iq})\,dx^q = 0,$$

d'où le résultat cherché.

f. Montrons à présent que, *dans un espace riemannien M_n, le tenseur métrique g_{ij} est invariant par rapport à la translation sur tout le M_n.* Désignons par $\widetilde{g}_{ij}(x)$ le translaté du tenseur $g_{ij}(A)$ le long d'une ligne L. D'après la définition *c*, quel que soit un couple de vecteurs translatés $\xi(x)$ et $\eta(x)$, on doit avoir

$$\widetilde{g}_{ij}(x)\,\xi^i(x)\,\eta^j(x) = \text{const}.$$

Or, conformément à la définition de la translation dans un espace riemannien, c'est le tenseur $g_{ij}(x)$ lui-même qui possède cette propriété (*6.32*), d'où $\widetilde{g}_{ij}(x) \equiv g_{ij}(x)$. On peut vérifier ce fait par un

calcul direct. D'après la formule (4), on a

$$d\widetilde{g}_{ij} = (\Gamma^s_{iq} g_{sj} + \Gamma^s_{jq} g_{is})\, dx^q =$$

$$= (\Gamma_{iq,\,j} + \Gamma_{jq,\,i})\, dx^q = \frac{\partial g_{ij}}{\partial x^q}\, dx^q = dg_{ij};$$

il en résulte que les quantités $\widetilde{g}_{ij}(x)$ qui coïncident avec $g_{ij}(x)$ pour $x = A$ coïncident aussi avec $g_{ij}(x)$ pour tout $x \in M_n$.

g. *Le champ du tenseur métrique deux fois contravariant* $g^{jk}(x)$, qui est défini par le système d'équations *6.16* (2)

$$g_{ij}(x)\, g^{jk}(x) = \delta^k_i, \tag{6}$$

est lui aussi *invariant par rapport à la translation sur* M_n, ce qui résulte directement de l'invariance des champs $g_{ij}(x)$ et δ^k_i et de l'unicité de solution du système (6).

6.47. Différentiation absolue. Soit $T(x)$ un champ tensoriel (dérivable) donné sur une variété M_n à connexion affine $\Gamma^k_{ij}(x)$. Supposons, par exemple, que le tenseur $T(x)$ soit une fois contra et une fois covariant; d'autres indices rendraient le calcul plus encombrant sans modifier l'essentiel.

a. Définissons la *différentielle absolue du tenseur* $T^k_i(x)$ en soustrayant de sa différentielle ordinaire sur le chemin dx^q son translaté le long de ce chemin:

$$DT^k_i(x) = dT^k_i - (\Gamma^s_{iq} T^k_s - \Gamma^k_{sq} T^s_i)\, dx^q.$$

Ainsi, l'égalité $DT^k_i(x) = 0$ le long d'une courbe L est équivalente à la condition de la translation du tenseur $T^k_i(x)$ le long de cette courbe.

Pour un tenseur $T^{k_1 \ldots k_p}_{i_1 \ldots i_r}$ de structure arbitraire, la différentielle absolue se définit de manière analogue:

$$DT^{k_1 \ldots k_p}_{i_1 \ldots i_r} = dT^{k_1 \ldots k_p}_{i_1 \ldots i_r} - (\Gamma^s_{i_1 q} T^{k_1 \ldots k_p}_{s i_2 \ldots i_r} + \ldots + \Gamma^s_{i_r q} T^{k_1 \ldots k_p}_{i_1 \ldots i_{r-1} s} -$$

$$- \Gamma^{k_1}_{sq} T^{s k_2 \ldots k_p}_{i_1 \ldots i_r} - \ldots - \Gamma^{k_p}_{sq} T^{k_1 \ldots k_{p-1}}_{i_1 \ldots i_r})\, dx^q.$$

Théorème. *La différentielle absolue d'un tenseur* $T^{k_1 \ldots k_p}_{i_1 \ldots i_r}$ *est un tenseur de la même structure.*

Pour la raison de simplicité, nous réalisons la démonstration pour un tenseur de la forme T^k_i. Transformons l'expression $DT^{k'}_{i'}$ à l'aide de la formule *6.14* (3)

$$p^k_{i' q'} - \Gamma^{k'}_{i' q'} p^k_{k'} = -p^i_{i'} p^j_{j' q'} \Gamma^k_{ij}$$

et de la formule qu'on obtient de celle-ci en y remplaçant les indices primés par non primés et vice versa:

$$p^{k'}_{iq} + p^{i'}_i p^{j'}_j \Gamma^{k'}_{i'j'} = \Gamma^k_{iq} p^{k'}_k.$$

Nous avons

$$DT^{k'}_{i'} = dT^{k'}_{i'} - (\Gamma^{s'}_{i'q'}T^{k'}_{s'} - \Gamma^{k'}_{s'q'}T^{s'}_{i'})\,dx^{q'} =$$
$$= d\,(p^{i}_{i'}p^{k'}_{k}T^{k}_{i}) - (\Gamma^{s'}_{i'q'}T^{k'}_{s'} - \Gamma^{k'}_{s'q'}T^{s'}_{i'})\,dx^{q'} = p^{i}_{i'}p^{k'}_{k}\,dT^{k}_{i} +$$
$$+ p^{i}_{i'q'}p^{q'}_{q}\,dx^{q}p^{k'}_{k}T^{k}_{i} + p^{i}_{i'}p^{k'}_{kq}\,dx^{q}T^{k}_{i} - (\Gamma^{s'}_{i'q'}p^{s}_{s'}p^{k'}_{k}T^{k}_{s} -$$
$$- \Gamma^{k'}_{s'q'}p^{i}_{i'}p^{s'}_{s}T^{s}_{i})\,p^{q'}_{q}\,dx^{q} = p^{i}_{i'}p^{k'}_{k}\,dT^{k}_{i} + (p^{i}_{i'q'} -$$
$$- \Gamma^{s'}_{i'q'}p^{i}_{s'})\,p^{q'}_{q}\,p^{k'}_{k}T^{k}_{i}\,dx^{q} + (p^{k'}_{sq} + \Gamma^{k'}_{s'q'}p^{s'}_{s}\,p^{q'}_{q})p^{i}_{i'}T^{s}_{i}\,dx^{q} =$$
$$= p^{i}_{i'}p^{k'}_{k}\,(dT^{k}_{i} - \Gamma^{s}_{iq}T^{k}_{s}\,dx^{q} + \Gamma^{k}_{sq}T^{s}_{i}\,dx^{q}) = p^{i}_{i'}p^{k'}_{k}DT^{k}_{i}$$

ce qu'il fallait.

b. On peut mettre la différentielle absolue de T^{k}_{i} sous la forme

$$DT^{k}_{i} = \nabla_{q}T^{k}_{i}\,dx^{q}, \tag{1}$$

où l'expression

$$\nabla_{q}T^{k}_{i} = \frac{\partial T^{k}_{i}}{\partial x^{q}} - (\Gamma^{s}_{iq}T^{k}_{s} - \Gamma^{k}_{sq}T^{s}_{i}) \tag{2}$$

s'appelle *dérivée absolue* ou *covariante du champ tensoriel* $T^{k}_{i}(x)$ *par rapport à la coordonnée* x^{q}. En tant que solution de l'équation tensorielle (1), les quantités $\nabla_{q}T^{k}_{i}$ forment encore un tenseur qui a un indice *covariant* (q) de plus que le tenseur T^{k}_{i}. Il vaut la peine de rappeler à ce propos que les dérivées ordinaires du champ tensoriel, i.e. les quantités $\dfrac{\partial T^{k}_{i}}{\partial x^{q}}$ ne constituent pas un tenseur (*6.26*).

c. Dans un espace riemannien, à côté de la dérivée covariante, on considère la *dérivée contravariante*:

$$\nabla^{q}T^{k}_{i} = (\nabla_{r}T^{k}_{i})\,g^{rq}. \tag{3}$$

Ici g^{rq} est le tenseur métrique deux fois contravariant (*6.16*). La dérivée contravariante d'un tenseur T est un tenseur qui a un indice *contravariant* de plus que le tenseur T lui-même.

§ 6.5. Courbure

6.51. Commutateur de deux différentiations. Supposons que, dans un espace M_n à connexion affine Γ^{k}_{ij} il y ait une surface bidimensionnelle $P = \{x \in M_n : x^{i} = x^{i}(u, v), (u, v) \in G \subset R_2\}$ qui contient un point $A = A(u_0, v_0)$. Désignons par d la différentielle (ordinaire) d'un champ tensoriel quelconque $T(x)$ dans la direction de la ligne de u (où la coordonnée v est cons-

tante):

$$dT = \frac{\partial T}{\partial u}\,du = \frac{\partial T}{\partial x^i}\,\frac{\partial x^i}{\partial u}\,du,$$

et par $\tilde{d}$ la différentielle analogue le long de la ligne de v:

$$\tilde{d}T = \frac{\partial T}{\partial v}\,dv = \frac{\partial T}{\partial x^i}\,\frac{\partial x^i}{\partial v}\,dv.$$

Les différentielles d et $\tilde{d}$ commutent, puisque

$$\tilde{d}\,(dT) = \tilde{d}\left(\frac{\partial T}{\partial u}\,du\right) = \frac{\partial^2 T}{\partial u\,\partial v}\,dv\,du,$$

$$d\,(\tilde{d}T) = d\left(\frac{\partial T}{\partial v}\,dv\right) = \frac{\partial^2 T}{\partial v\,\partial u}\,du\,dv.$$

Désignons par D et $\tilde{D}$ les différentielles absolues respectives. Il se trouve qu'en général elles ne commutent pas; calculons leur commutateur $\tilde{D}D - D\tilde{D}$ relativement au champ vectoriel $\xi = \{\xi^l(x)\}$. Conformément à la définition *6.34*, on a:

$$\begin{aligned}\tilde{D}\,(D\xi^l) &= \tilde{D}\,(d\xi^l + \Gamma^l_{hi}\xi^h\,dx^i) = \\ &= \tilde{d}\,(d\xi^l + \Gamma^l_{hi}\xi^h\,dx^i) + \Gamma^l_{pj}\,(d\xi^p + \Gamma^p_{ki}\xi^k\,dx^i)\,\tilde{d}x^j = \\ &= \tilde{d}\,(d\xi^l) + \frac{\partial\Gamma^l_{ki}}{\partial x^j}\,\tilde{d}x^j\xi^k\,dx^i + \Gamma^l_{ki}\,\tilde{d}\xi^k\,dx^i + \Gamma^l_{ki}\xi^k\,\tilde{d}\,(dx^i) + \\ &\qquad + \Gamma^l_{pj}\tilde{d}\xi^p\,dx^j + \Gamma^l_{pj}\Gamma^p_{ki}\xi^k\,dx^i\tilde{d}x^j.\end{aligned}$$

L'expression $D\,(\tilde{D}\xi^l)$ s'obtient en remplaçant ici d par $\tilde{d}$ et vice versa. Si l'on soustrait $D\,(\tilde{D}\xi^l)$ de $\tilde{D}\,(D\xi^l)$ et utilise la commutativité des symboles d et $\tilde{d}$, on obtient

$$\begin{aligned}(\tilde{D}D - D\tilde{D})\,\xi^l &= \\ &= \left(\frac{\partial\Gamma^l_{ki}}{\partial x^j} - \frac{\partial\Gamma^l_{kj}}{\partial x^i}\right)\xi^k\,dx^i\,\tilde{d}x^j + (\Gamma^p_{ki}\Gamma^l_{pj} - \Gamma^p_{kj}\Gamma^l_{pi})\,\xi^k\,dx^i\,\tilde{d}x^j = \\ &= \left(\frac{\partial\Gamma^l_{ki}}{\partial x^j} - \frac{\partial\Gamma^l_{kj}}{\partial x^i} + \Gamma^p_{ki}\Gamma^l_{pj} - \Gamma^p_{kj}\Gamma^l_{pi}\right)\xi^k\,dx^i\,\tilde{d}x^j. \qquad (1)\end{aligned}$$

Avec la désignation

$$R^l_{ij,\,k} = \frac{\partial\Gamma^l_{ki}}{\partial x^j} - \frac{\partial\Gamma^l_{kj}}{\partial x^i} + \Gamma^p_{ki}\Gamma^l_{pj} - \Gamma^p_{kj}\Gamma^l_{pi}, \qquad (2)$$

l'égalité (1) prend la forme

$$(\tilde{D}D - D\tilde{D})\,\xi^l = R^l_{ij,\,k}\xi^k\,dx^i\,\tilde{d}x^j. \qquad (3)$$

Dans le premier membre de (3), on a un tenseur une fois contravariant (comme résultat de différentiations absolues d'un tenseur une fois contravariant). Dans le second membre, les expressions ξ^h, dx^i, $\widetilde{dx}^j$ sont de caractère tensoriel (des tenseurs une fois contravariants). Elles sont arbitraires, puisque le choix du vecteur $\xi^h \in \in T_n(x)$ et des coordonnées u, v est à notre gré. Par conséquent, en vertu de *6.14d*, *la quantité $R^l_{ij,k}$ représente un tenseur trois fois covariant et une fois contravariant.* On voit des formules (2) que ce tenseur est antisymétrique par rapport aux indices i et j:

$$R^l_{ij,\,k} = -R^l_{ji,\,k}.$$

Le tenseur $R = R^l_{ij,k}$ s'appelle *tenseur de courbure de l'espace M_n à connexion affine* Γ^k_{ij}.

6.52. Courbure et parallélisme absolu. Supposons que, dans l'espace M_n, la connexion Γ engendre le parallélisme absolu (*6.43*). Alors, ayant un vecteur ξ en un point donné A, on peut construire au voisinage de A le champ vectoriel $\xi(x)$ déterminé d'une façon univoque par la translation du vecteur ξ. Si, en construisant le tenseur R, on choisit pour champ $\xi(x)$ juste ce champ obtenu en translatant le vecteur ξ, alors, la différentielle absolue d'un vecteur translaté étant nulle, on aura $D\xi = 0$, $\widetilde{D}\xi = 0$, par conséquent, le premier membre de l'égalité *6.51* (3) sera aussi nul:

$$R^l_{ij,\,k}\xi^k\, dx^i\, \widetilde{dx}^j = 0. \tag{1}$$

Comme ξ^k, dx^i, $\widetilde{dx}^j$ sont arbitraires, cela veut dire que

$$R^l_{ij,\,k}(x) \equiv 0.$$

Définitivement, *la courbure d'un espace M_n avec le parallélisme absolu est nulle.*

Etablissons la réciproque. Supposons que la courbure d'un espace M_n à une connexion affine Γ^k_{ij} soit identiquement nulle. Soit une famille de courbes lisses joignant deux points donnés A et B; chaque courbe de la famille est déterminée par une valeur fixe d'un paramètre τ variant de 0 à 1, ses points correspondent aux valeurs de t qui varie aussi de 0 à 1:

$$x^i = x^i(t, \tau),$$
$$x(0, \tau) = A,\ x(1, \tau) = B.$$

Considérons au point A un vecteur ξ et montrons que son translaté au point B le long d'une courbe $x = x(t, \tau)$ $(0 \leqslant t \leqslant 1)$ ne dépend pas de la valeur $\tau \in [0, 1]$. Désignons par D la différentielle absolue le long des courbes de la famille (i.e. par rapport à t, pour τ fixe), et par $\widetilde{D}$ la différentielle absolue par rapport à τ, pour t fixe. Comme

le vecteur ξ est translaté par hypothèse, on a $D\xi = 0$. Il résulte alors de $R \equiv 0$ que

$$D\widetilde{D}\xi = 0,$$

donc le vecteur $\widetilde{D}\xi$ est translaté lui aussi le long de chaque courbe de la famille. En réalité, pour $t = 0$, le vecteur $\xi = \xi(0, \tau)$ et le point $x = x(0, \tau) = A$ ne dépendent pas de τ; donc, pour $t = 0$ on a:

$$\widetilde{D}\xi^h = d\xi^h - \Gamma^h_{ij}\xi^i(0, \tau)\frac{dx^j}{d\tau}d\tau = 0.$$

Par conséquent, le vecteur $\widetilde{D}\xi(1, \tau)$ est lui aussi nul en tant que translaté. D'autre part,

$$\widetilde{D}\xi^h(1, \tau) = \widetilde{d}\xi^h(1, \tau) + \Gamma^h_{ij}(1, \tau)\xi^i dx^j(1, \tau),$$

et comme dans notre cas $\widetilde{d}x^j = 0$, il résulte de $\widetilde{D}\xi^h(1, \tau) = 0$ que $\widetilde{d}\xi(1, \tau) = 0$, i.e. le vecteur $\xi(1, \tau)$ n'est pas fonction de τ. Nous avons ainsi démontré le théorème suivant:

Théorème. *Un espace M_n à connexion affine a le parallélisme absolu si, et seulement si, sa courbure est nulle.*

En utilisant *6.43* nous pouvons formuler encore un critère de l'équivalence de la connexion affine d'un espace M_n et de la connexion riemannienne de l'espace euclidien:

Théorème. *Un espace M_n à connexion affine est affinement équivalent à l'espace euclidien n-dimensionnel si et seulement si sa courbure et sa torsion sont identiquement nulles.*

Il existe des espaces à connexion affine qui ont une torsion nulle et une courbure non nulle et vice versa.

N'importe quel espace (ou surface) riemannien qui n'est pas isométrique à l'espace euclidien sert d'exemple d'espace ayant la torsion nulle et la courbure non nulle. Il est plus difficile de construire un exemple d'espace ayant la courbure nulle et la torsion non nulle (cf. exercice 4).

6.53. Variation des coordonnées d'un vecteur en translation le long d'un contour fermé.

a. Supposons que, dans un espace M_n à une connexion affine deux fois dérivable $\Gamma(x)$, il y ait une surface bidimensionnelle P sans point singulier:

$$x = x(u, v) \in M_n, \qquad (u, v) \in G \subset R_2.$$

Allons translater un vecteur ξ le long des courbes L passant sur la surface P par un point fixe A, pour lesquelles la longueur formelle obtenue en intégrant l'expression $ds = \sqrt{du^2 + dv^2}$ le long de L reste inférieure à une constante fixe h (la longueur formelle n'a pas de sens absolu, mais nous procédons pour le moment dans un

système fixe de coordonnées). Toutes ces courbes se trouvent sur la surface P, dans le voisinage U du point A qui est déterminé par les inégalités $|u - u(A)| < h$, $|v - v(A)| < h$. Sur chaque tel chemin L, les coordonnées $x^i(s)$ satisfont aux relations

$$|dx^i| = \left|\frac{\partial x^i}{\partial u}du + \frac{\partial x^i}{\partial v}dv\right| \leqslant 2C_1 ds, \tag{1}$$

où $C = \max_U \left(\left|\frac{\partial x^i}{\partial u}\right|, \left|\frac{\partial x^i}{\partial v}\right|\right)$. On en tire les inégalités suivantes pour les accroissements des coordonnées x^i (en comptant du point A où $s = 0$):

$$|\Delta x^i(s)| = |x^i(s) - x^i(0)| \leqslant 2C_1 s \leqslant 2C_1 h. \tag{2}$$

Les coefficients $\Gamma^k_{ij}(x)$ satisfont sur le même chemin L aux relations

$$|\Gamma^k_{ij}(x)| \leqslant C_2, \tag{3}$$

ou aux relations plus précises

$$|\Gamma^k_{ij}(s) - \Gamma^k_{ij}(0)| \leqslant C_3 s \leqslant C_3 h, \tag{4}$$

ou aux relations encore plus précises

$$\left|\Gamma^k_{ij}(s) - \Gamma^k_{ij}(0) - \frac{\partial \Gamma^k_{ij}(0)}{\partial x^m}\Delta x^m\right| \leqslant C_4 s^2, \tag{5}$$

où la constante C_3 peut être estimée à l'aide des dérivées premières, et C_4 à l'aide des dérivées secondes des fonctions $\Gamma^k_{ij}(x)$ dans le voisinage U. Les composantes du vecteur $\xi(s)$ qui est le translaté du vecteur ξ vérifient l'inégalité découlant de *6.42* (1):

$$|\xi^i(s)| \leqslant C_5 |\xi| \equiv C_5 \sqrt{\sum_{j=1}^{n} |\xi^j|^2}, \tag{6}$$

d'où, en utilisant (3), on trouve

$$|d\xi^k| = |\Gamma^k_{ij}\xi^i dx^j| \leqslant C_6 |\xi| ds, \tag{7}$$

donc aussi

$$|\nabla\xi^k| = |\xi^k(s) - \xi^k| \leqslant C_6 |\xi| s \leqslant C_6 |\xi| h. \tag{8}$$

Nous affirmons enfin qu'a lieu l'inégalité

$$I(s) \equiv \left|\int_0^s \Gamma^k_{ij}(x)\,\xi^i(x)\frac{dx^j}{ds}\,ds - \Gamma^k_{ij}(0)\,\xi^i\Delta x^j\right| \leqslant C_7 s^2 |\xi|. \tag{9}$$

En effet, $\Delta x^j = \int_0^s dx^j$, par conséquent:

$$I(s) \equiv \left|\int_0^s \Gamma^k_{ij}(x)\,\xi^i(x)\,dx^j - \int_0^s \Gamma^k_{ij}(0)\,\xi^i\,dx^j\right| =$$

$$= \left| \int_0^s [\Gamma_{ij}^k(x)\,\xi^i(x) - \Gamma_{ij}^k(0)\,\xi^i]\,dx^j \right| =$$

$$= \left| \int_0^s \{[\Gamma_{ij}^k(x) - \Gamma_{ij}^k(0)]\,\xi^i(x) + \Gamma_{ij}^k(0)\,[\xi^i(x) - \xi^i]\}\,dx^j \right|.$$

Avec les estimations (4), (6), (3) et (8) on a la relation cherchée :

$$I(s) \leqslant \int_0^s (C_3 s C_5 |\xi| + C_2 C_6 |\xi| s)\,|dx^j| \leqslant C|\xi| s^2.$$

b. Soit maintenant $L \in P$ un chemin fermé, de longueur formelle $\leqslant h$, qui revient au point A. Translatons un vecteur $\xi = \{\xi^i\}$, choisi arbitrairement au point A, le long de L. Dans un espace M_n sans parallélisme absolu, le vecteur ξ aura, à la suite de la translation, un accroissement $\Delta\xi$ que nous tenons à évaluer. L'accroissement total de la coordonnée ξ^i s'exprimera par la formule

$$\Delta\xi^l = \oint_L d\xi^l = -\oint_L \Gamma_{pj}^l(x)\,\xi^p(x)\,dx^j. \qquad (10)$$

En vertu de (5), on a

$$\Gamma_{pj}^l(x) = \Gamma_{pj}^l(0) + \frac{\partial \Gamma_{pj}^l(0)}{\partial x^i}\,\Delta x^i + O(s^2),$$

et, d'après (9),

$$\xi^p(x) = \xi^p - \int_0^s \Gamma_{hi}^p \xi^k(x)\,dx^i = \xi^p - \Gamma_{hi}^p(0)\,\xi^h \Delta x^i + O(s^2).$$

On en tire

$$\Delta\xi^l = -\oint_L \left[\Gamma_{pj}^l(0) + \frac{\partial \Gamma_{pj}^l(0)}{\partial x^i}\,\Delta x^i + O(s^2)\right] \times$$

$$\times [\xi^p - \Gamma_{hi}^p(0)\,\xi^h \Delta x^i + O(s^2)]\,dx^j =$$

$$= -\Gamma_{pj}^l(0)\,\xi^p \oint_L dx^j - \frac{\partial \Gamma_{pj}^l}{\partial x^i}\,\xi^p \oint_L \Delta x^i\,dx^j +$$

$$+ \Gamma_{hi}^p(0)\,\Gamma_{pj}^l(0)\,\xi^k \oint_L \Delta x^i\,dx^j + O(h^3).$$

Le premier terme est, évidemment, nul. On obtient

$$\Delta\xi^l = \left(\Gamma_{hi}^p(0)\,\Gamma_{pj}^l(0) - \frac{\partial \Gamma_{hj}^l(0)}{\partial x^i}\right) \xi^k \oint_L \Delta x^i\,dx^j + O(h^3). \qquad (11$$

Passons aux paramètres u et v dans l'intégrale intérieure. Avec la formule de Green *4.16* (3), nous avons:

$$\oint_L \Delta x^i\, dx^j = \oint_L \Delta x^i \left(\frac{\partial x^j}{\partial u}\, du + \frac{\partial x^j}{\partial v}\, dv\right) =$$

$$= \iint_G \left[\frac{\partial}{\partial u}\left(\Delta x \frac{\partial x^j}{\partial v}\right) - \frac{\partial}{\partial v}\left(\Delta x^i \frac{\partial x^j}{\partial u}\right)\right] du\, dv =$$

$$= \iint_G \left(\frac{\partial x^i}{\partial u}\frac{\partial x^j}{\partial v} - \frac{\partial x^j}{\partial u}\frac{\partial x^i}{\partial v}\right) du\, dv =$$

$$= \left[\frac{\partial x^i(0)}{\partial u}\frac{\partial x^j(0)}{\partial v} - \frac{\partial x^j(0)}{\partial u}\frac{\partial x^i(0)}{\partial v}\right] \iint_G du\, dv + O(h) \iint_G du\, dv. \qquad (12)$$

Le facteur dans le crochet est le bivecteur (*6.15*) construit sur les vecteurs $\frac{\partial x(0)}{\partial u}$ et $\frac{\partial x(0)}{\partial v}$; nous le désignons par x^{ij}. Posons

$$\sigma = \iint_G du\, dv;$$

comme chacune des coordonnées u, v ne varie pas, dans les limites du domaine G, plus que de h, cette quantité représente un infiniment petit du deuxième ordre (ou plus) par rapport à h.

L'égalité (11) prend maintenant la forme

$$\Delta\xi^l = \left(\Gamma^p_{ki}\Gamma^l_{pj} - \frac{\partial \Gamma^l_{kj}}{\partial x^i}\right) \xi^k x^{ij}\sigma + O(h^3), \qquad (13)$$

où les valeurs des fonctions Γ et de leurs dérivées se rapportent au point A. Si l'on permute les indices i et j dans le second membre et effectue la multiplication par -1, alors, vu l'antisymétrie du tenseur x^{ij}, le second membre ne change pas. En composant la demi-somme de l'égalité (13) et de celle qui en résulte à la suite de la transformation indiquée, on trouve

$$\Delta\xi^l = \frac{1}{2}\left(\frac{\partial \Gamma^l_{ki}}{\partial x^j} - \frac{\partial \Gamma^l_{kj}}{\partial x^i} + \Gamma^p_{ki}\Gamma^l_{pj} - \Gamma^p_{kj}\Gamma^l_{pi}\right) \xi^k x^{ij}\sigma + O(h^3),$$

ou bien

$$\Delta\xi^l = \frac{1}{2} R^l_{ij,\,k}\xi^k x^{ij}\sigma + O(h^3). \qquad (14)$$

Ainsi, *aux infiniment petits du troisième ordre près, la rotation du vecteur* ξ *s'exprime par la formule* (14) *où le rôle principal appartient au tenseur de courbure* $R^l_{ij,k}$.

6.54. Tenseur de courbure dans un espace riemannien. Dans un espace riemannien où les coefficients Γ^k_{ij}

sont définis par la formule *6.32* (5)

$$\Gamma^k_{ij} = g^{ks}\Gamma_{ij,\,s} = \frac{1}{2}\,g^{ks}\left(\frac{\partial g_{is}}{\partial x^j} + \frac{\partial g_{js}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^s}\right) g^{ks}$$

et, en particulier, symétriques par rapport aux indices i, j, le tenseur de courbure

$$R^l_{ii,\,k} = \frac{\partial \Gamma^l_{ki}}{\partial x^j} - \frac{\partial \Gamma^l_{kj}}{\partial x^i} + \Gamma^p_{ki}\Gamma^l_{pj} - \Gamma^p_{kj}\Gamma^l_{pi}$$

montre des particularités supplémentaires de la symétrie.

Considérons le tenseur quatre fois covariant

$$R_{ij,\,kl} = R^s_{ij,\,k}g_{sl} = \left(\frac{\partial \Gamma^s_{ik}}{\partial x^j} - \frac{\partial \Gamma^s_{jk}}{\partial x^i} + \Gamma^p_{ik}\Gamma^s_{jp} - \Gamma^p_{jk}\Gamma^s_{ip}\right) g_{sl}.$$

Notons que, pour une surface n-dimensionnelle dans l'espace $(n+1)$-dimensionnel euclidien, les composantes du tenseur $R_{ij,\,kl}$ sont des mineurs de la matrice de la deuxième forme quadratique (*5.32* (8)). Nous affirmons que, dans le cas général, *le tenseur $R_{ij,\,kl}$ est un tenseur du type Ricci* (*6.18*). En effet, le tenseur $R_{ij,\,kl}$, de même que le tenseur $R^s_{ij,\,s}$, est antisymétrique par rapport aux indices i et j. Ajoutons qu'on peut permuter les couples d'indices i, j, et k, l, $R_{ij,\,kl} = R_{kl,\,ij}$. On a

$$\left(\frac{\partial \Gamma^s_{ki}}{\partial x^j} + \Gamma^p_{ki}\Gamma^s_{pj}\right) g_{sl} = \frac{\partial}{\partial x^j}\left(\Gamma^s_{ki}g_{sl}\right) - \Gamma^s_{ki}\frac{\partial g_{sl}}{\partial x^j} +$$

$$+ \Gamma^p_{ki}\Gamma_{pj,\,l} = \frac{1}{2}\frac{\partial}{\partial x^j}\left(\frac{\partial g_{il}}{\partial x^k} + \frac{\partial g_{kl}}{\partial x^i} - \frac{\partial g_{ik}}{\partial x^l}\right) - \Gamma^s_{ki}\left(\Gamma_{lj,\,s} + \Gamma_{sj,\,l}\right) +$$

$$+ \Gamma^s_{ki}\Gamma_{sj,\,l} = \frac{1}{2}\left(\frac{\partial^2 g_{il}}{\partial x^j\,\partial x^k} + \frac{\partial^2 g_{kl}}{\partial x^i\,\partial x^j} - \frac{\partial^2 g_{ik}}{\partial x^j\,\partial x^l}\right) - \Gamma^s_{ki}\Gamma_{lj,\,s}.$$

En alternant l'expression obtenue par rapport à i et j on obtient

$$R_{ij,\,kl} = \frac{1}{2}\left(\frac{\partial^2 g_{il}}{\partial x^j\,\partial x^k} - \frac{\partial^2 g_{jl}}{\partial x^i\,\partial x^k} - \frac{\partial^2 g_{ik}}{\partial x^j\,\partial x^l} + \frac{\partial^2 g_{ik}}{\partial x^i\,\partial x^l}\right) +$$

$$+ g_{sp}\left(\Gamma^s_{ki}\Gamma^p_{lj} - \Gamma^s_{kj}\Gamma^p_{li}\right). \qquad (1)$$

Cette expression admet de façon évidente la permutation des couples i, j et k, l en vertu de la symétrie du tenseur g_{sp} et de la connexion riemannienne Γ^l_{ik}.

Prouvons, enfin, l'identité de Ricci. Nous avons

$$R^s_{ij,\,k} + R^s_{jk,\,i} + R^s_{ki,\,j} =$$

$$= \frac{\partial \Gamma^s_{ki}}{\partial x^j} - \frac{\partial \Gamma^s_{kj}}{\partial x^i} + \frac{\partial \Gamma^s_{ij}}{\partial x^k} - \frac{\partial \Gamma^s_{ik}}{\partial x^j} + \frac{\partial \Gamma^s_{jk}}{\partial x^i} - \frac{\partial \Gamma^s_{ji}}{\partial x^k} +$$

$$+ \Gamma^p_{ki}\Gamma^s_{pj} - \Gamma^p_{kj}\Gamma^s_{pi} + \Gamma^p_{ij}\Gamma^s_{pk} - \Gamma^p_{ik}\Gamma^s_{pj} + \Gamma^p_{jk}\Gamma^s_{pi} - \Gamma^p_{ji}\Gamma^s_{jk} = 0$$

d'après la symétrie des symboles Γ^k_{ij} par rapport aux indices inférieurs. En contractant cette identité avec le tenseur g_{sl} nous aboutissons à l'identité de Ricci pour le tenseur $R_{ij,\,kl}$.

6.55. Courbure et angle de rotation d'un vecteur translaté le long d'un contour fermé. Pour un espace riemannien, la formule *6.53* (14) peut être précisée.

a. Dans un espace riemannien, on peut définir l'aire des parties des surfaces bidimensionnelles données, par exemple, par les équations

$$x^i = x^i(u, v),$$
$$(u, v) \in \Omega \subset R_2,$$

en utilisant la formule *6.31* (1)

$$dS = \sqrt{EG - F^2}\, du\, dv, \tag{1}$$

où

$$E = \left(\frac{\partial x}{\partial u}, \frac{\partial x}{\partial u}\right), \quad F = \left(\frac{\partial x}{\partial u}, \frac{\partial x}{\partial v}\right), \quad G = \left(\frac{\partial x}{\partial v}, \frac{\partial x}{\partial v}\right).$$

Si l'on passe des coordonnées u, v aux nouvelles coordonnées $\tilde{u}$, $\tilde{v}$, l'élément d'aire sera de la forme (cf. *3.61c*)

$$dS = \sqrt{EG - F^2}\, du\, dv = \sqrt{EG - F^2} \left|\frac{\partial(u, v)}{\partial(\tilde{u}, \tilde{v})}\right| d\tilde{u}\, d\tilde{v}. \tag{2}$$

Si les nouvelles coordonnées sont choisies de façon que

$$\left|\frac{\partial(u, v)}{\partial(\tilde{u}, \tilde{v})}\right| = \frac{1}{\sqrt{EG - F^2}} \quad \text{ou} \quad \left|\frac{\partial(\tilde{u}, \tilde{v})}{\partial(u, v)}\right| = \sqrt{EG - F^2}, \tag{3}$$

alors, en coordonnées $\tilde{u}$, $\tilde{v}$, on aura

$$dS = d\tilde{u}\, d\tilde{v}. \tag{4}$$

Pour que la condition (3) soit satisfaite, posons $\tilde{u} = \varphi(u, v)$, $\tilde{v} = v$; alors

$$\frac{\partial(\tilde{u}, \tilde{v})}{\partial(u, v)} = \begin{vmatrix} \frac{\partial \varphi}{\partial u} & \frac{\partial \varphi}{\partial v} \\ 0 & 1 \end{vmatrix} = \frac{\partial \varphi}{\partial u},$$

et il suffit de choisir pour φ la fonction

$$\varphi(u, v) = \int \sqrt{EG - F^2}\, du.$$

C'est dans ce système des coordonnées $\tilde{u}$, $\tilde{v}$, (nous les redésignons par u, v) que nous considérons la formule *6.53* (14). Dans ces coordonnées, σ est l'aire riemannienne du domaine limité par le con-

tour L. Dans les mêmes coordonnées, nous avons $EG - F^2 = 1$, de sorte que le bivecteur $x^{ij}(A) = x^{ij}$ construit sur les vecteurs $\frac{\partial x(A)}{\partial u}$ et $\frac{\partial x(A)}{\partial v}$ et figurant dans *6.53* (14) devient bivecteur unitaire.

b. Ensuite, en partant de l'évaluation de l'accroissement des coordonnées qu'on a réalisée dans *6.53*, on peut passer au calcul de l'angle de rotation d'un vecteur que l'on translate le long d'un contour fermé L.

Sur le plan Π déterminé par les vecteurs $\frac{\partial x}{\partial u}$ et $\frac{\partial x}{\partial v}$ (i.e. sur le plan du bivecteur x^{ij}), nous choisissons pour sens positif de variation des angles celui qui va de la direction d'accroissement du paramètre u vers la direction d'accroissement du paramètre v. Fixons sur ce plan un vecteur unitaire ξ et le vecteur η déduit de ξ par une rotation de 90° dans le sens positif.

Après avoir translaté le vecteur ξ le long du contour, on aboutit à un vecteur que nous désignons par $\xi + \Delta\xi$. Décomposons le vecteur $\Delta\xi$ en somme de trois composantes de la forme

$$\Delta\xi = \Delta_1 \cdot \xi + \Delta_2 \cdot \eta + \Delta_3 \cdot \zeta$$

où les coefficients Δ_1, Δ_2, Δ_3 sont des nombres réels, le vecteur ζ étant orthogonal au plan Π. La formule *6.53* (14) fournira alors

$$\Delta_1 = (\xi, \Delta\xi) = (\xi, \Delta_1 \cdot \xi) = g_{lp}\xi^p \Delta\xi^l = \frac{1}{2} g_{lp}\xi^p (R^l_{ij,\,k}\xi^k x^{ij}\sigma + O(h^3)) =$$

$$= \frac{1}{2}\xi^p R_{ij,\,kp}\xi^k x^{ij}\sigma + O(h^3) = O(h^3),$$

car le premier terme s'annule en raison de l'antisymétrie du tenseur $R_{ij,\,kp}$ par rapport aux indices k et p.

Désignons par φ l'angle compté du vecteur ξ jusqu'au vecteur $\xi + \Delta_1\xi + \Delta_2\eta$ (i.e. à la projection du vecteur $\xi + \Delta\xi$ sur le plan Π) dans le sens positif. Nous avons (fig. 6.5-1) :

$$\operatorname{tg}\varphi = \frac{\Delta_2}{1+\Delta_1} = \Delta_2 (1 + O(h^3)) =$$

$$= (\eta, \Delta_2\eta)(1 + O(h^3)) = (\eta, \Delta\xi)(1 + O(h^3)) =$$

$$= g_{lp}\eta^p \Delta\xi^l (1 + O(h^3)) = \frac{1}{2} g_{lp}\eta^p (R^l_{ij,\,k}\xi^k x^{ij}\sigma + O(h^3))(1 + O(h^3)) =$$

$$= \frac{1}{2}\xi^k \eta^p R_{ij,\,kp} x^{ij}\sigma + O(h^3). \quad (5)$$

Le bivecteur $\xi^h\eta^p - \xi^p\eta^h$ détermine le même plan Π et l'aire égale à 1 (car ξ et η sont orthonormés), donc coïncide avec le bivecteur x^{ij}. Pour passer à ce dernier, permutons dans (5) les indices k et p, et formons la demi-somme de deux égalités ; vu l'antisymétrie du

tenseur $R_{ij,\,kp}$ par rapport aux indices k et p, nous trouvons

$$\operatorname{tg}\varphi=\frac{1}{4}x^{ij}x^{kp}R_{ij,\,kp}\sigma+O(h^3).$$

Pour l'angle $\varphi = \operatorname{arctg}(\operatorname{tg}\varphi) = \operatorname{tg}\varphi + O(\operatorname{tg}^3\varphi)$, nous obtenons une expression analogue:

$$\varphi=\frac{1}{4}x^{ij}x^{kp}R_{ij,\,kp}\sigma+O(h^3). \tag{6}$$

c. Si l'on divise la dernière égalité par σ (en supposant que σ soit un infiniment petit du deuxième ordre exactement par rapport à h) et passe à la limite lorsque le contour L se contracte vers le point A, on aura

$$\lim_{L\to A}\frac{\varphi}{\sigma}=\frac{1}{4}x^{ij}x^{kp}R_{ij,\,kp}. \tag{7}$$

La quantité scalaire ainsi obtenue s'appelle *courbure de l'espace riemannien M_n au point A, dans la direction bidimensionnelle déterminée par le bivecteur unitaire x^{ij}*; on la désigne par K.

d. Enfin, on peut choisir pour bivecteur unitaire x^{ij} n'importe quel bivecteur x^{ij} dans le même plan Π divisé par son aire. Avec la formule *6.17* (5), nous trouvons pour un bivecteur quelconque x^{ij}:

$$K=\frac{R_{ij,\,kp}x^{ij}x^{kp}}{G_{ij,\,kp}x^{ij}x^{kp}} \tag{8}$$

où $G_{ij,\,kp}=g_{ik}g_{jp}-g_{jk}g_{ip}$ est le tenseur dérivé métrique (*6.17b*) au point A.

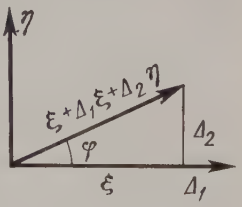


Fig. 6.5-1

6.56. Relation entre la courbure dans une direction bidimensionnelle et les courbures des surfaces bidimensionnelles correspondantes. Calculons d'après la formule *6.55* (8) la courbure de l'espace riemannien dans la direction bidimensionnelle déterminée par la surface de coordonnées bidimensionnelle $P_{12}=\{x\in M_n:\ x^1=u,\ x^2=v,\ x^3=\ldots=x^n=0\}$. Le bivecteur x^{ij} sera déterminé par les mineurs de la matrice

$$\begin{Vmatrix}1&0&0\ldots0\\0&1&0\ldots0\end{Vmatrix}$$

et n'aura les composantes non nulles que pour les valeurs $i=1$, $j=2$, où $x^{12}=1$, et $i=2$, $j=1$, où $x^{21}=-1$. La courbure K a donc la forme

$$K=\frac{R_{12,\,12}+R_{21,\,21}-R_{12,\,21}-R_{21,\,12}}{G_{12,\,12}+G_{21,\,21}-G_{12,\,21}-G_{21,\,12}}.$$

Or, les deux tenseurs R et G sont antisymétriques par rapport au premier et au deuxième couples d'indices, et ne varient pas en permutant les couples; nous avons donc

$$K = \frac{R_{12,\,12}}{G_{12,\,12}}. \tag{1}$$

En particulier, si toute la variété riemannienne M_n est bidimensionnelle, $n = 2$, cette formule fournit une expression de la courbure de Gauss de la variété M_2 (*6.35d*). Ceci est facile à concevoir géométriquement: d'après la formule *6.55* (8), sur la variété M_2, la quantité K est la limite du rapport de l'angle de rotation d'un vecteur translaté le long d'un contour fermé à l'aire limitée par ce contour, ce qui coïncide avec la courbure de Gauss de la surface M_2 (*5.63*).

Dans le cas général ($n > 2$), *la courbure de Gauss de la surface de coordonnées* P_{12} ne correspond pas à la formule (1). Ceci est dû au fait que la translation d'un vecteur le long d'un contour L a sur la variété M_n un autre sens que sur la surface P_{12} considérée comme variété bidimensionnelle: dans le premier cas, la translation est déterminée par les valeurs de tous les Γ^k_{ij}, ce qui implique que le vecteur en translation sort du plan tangent à la surface P_{12}; dans le deuxième cas, la translation est déterminée par seules les valeurs de Γ^k_{ij} avec $i, j, k = 1, 2$, et le vecteur en translation reste dans le plan tangent à la surface P_{12}. Par exemple, lorsque la variété M_n est l'espace euclidien R_n, et la surface P_{12} représente une partie d'une sphère bidimensionnelle dans R_n, nous nous trouvons exactement dans la situation ci-décrite: la translation le long de chaque contour fermé dans l'espace euclidien, en particulier, d'un contour sur la sphère, fait revenir le vecteur à sa position initiale, tandis que la translation le long d'un contour fermé sur la sphère en tant que surface riemannienne ne le fait en général pas (*5.62g*).

Démontrons la même proposition analytiquement. Conformément à *6.54* (1), le numérateur dans la formule (1) pour la courbure K d'une variété M_n dans la direction bidimensionnelle P_{12} vaut

$$R_{12,\,12} = \frac{1}{2}\left(\frac{\partial^2 g_{12}}{\partial x^1\,\partial x^2} - \frac{\partial^2 g_{22}}{\partial x^1\,\partial x^1} - \frac{\partial^2 g_{11}}{\partial x^2\,\partial x^2} + \frac{\partial^2 g_{12}}{\partial x^1\,\partial x^2}\right) +$$

$$+ g_{sp}\left(\Gamma^s_{11}\Gamma^p_{22} - \Gamma^s_{12}\Gamma^p_{21}\right), \tag{2}$$

avec la sommation sur tous les s et p de 1 à n. Or, dans l'expression

$$\widetilde{K} = \frac{\widetilde{R}_{12,\,12}}{\widetilde{G}_{12,\,12}}, \quad \widetilde{G}_{12,\,12} = G_{12,\,12}, \tag{3}$$

pour la courbure de Gauss $\widetilde{K}$ de la surface P_{12}, le numérateur $\widetilde{R}_{12,\,12}$ est égal à

$$\frac{1}{2}\left(\frac{\partial^2 g_{12}}{\partial x^1\,\partial x^2}-\frac{\partial^2 g_{22}}{\partial x^1\,\partial x^1}-\frac{\partial^2 g_{11}}{\partial x^2\,\partial x^2}+\frac{\partial^2 g_{12}}{\partial x^1\,\partial x^2}\right)+$$

$$+g_{\alpha\beta}\left(\Gamma^{\alpha}_{11}\Gamma^{\beta}_{22}-\Gamma^{\alpha}_{12}\Gamma^{\beta}_{21}\right), \qquad (4)$$

avec la sommation sur seules les valeurs 1 et 2 des indices α et β.

On peut poser la question suivante: est-ce qu'il existe dans une variété M_n, pour tout bivecteur x^{ij} (A), une surface bidimensionnelle P telle que son plan tangent au point A soit déterminé par le bivecteur x^{ij} (A) et sa courbure de Gauss soit la courbure de la variété M_n dans la direction bidimensionnelle correspondante? (Ainsi, dans l'exemple précédent pour l'espace euclidien R_n, on peut choisir en tant que telle surface le plan déterminé par le bivecteur x^{ij} (A)). Nous allons y formuler la réponse affirmative.

Supposons que, en un point donné A, soit satisfaite la condition $\Gamma^k_{ij}(A)=0$ $(i,\,j,\,k=1,\,\ldots,\,n)$. Il est alors évident que les égalités (1)-(3), vu (4), nous donnent $\widetilde{K}=K$, i.e. la courbure de la variété M_n dans la direction bidimensionnelle x^1, x^2 coïncide avec la courbure de Gauss de la surface de coordonnées P_{12}. Même plus, dans notre cas on peut indiquer, *pour tout bivecteur x^{ij}, une surface bidimensionnelle P satisfaisant à la condition imposée.* Pour la raison de simplicité posons $x^i(A)=0$ $(i=1,\,\ldots,\,n)$ et, pour un bivecteur donné x^{ij} (A), effectuons une transformation linéaire des coordonnées $x^i=a^i_{i'}x^{i'}$ qui fait transformer le plan du bivecteur x^{ij} (A) en le plan de coordonnées $x^{1'}$, $x^{2'}$. Vu que, pour une transformation linéaire des coordonnées, les quantités Γ^k_{ij} se transforment suivant une loi tensorielle (*6.33*), on a $\Gamma^{k'}_{i'j'}(A)=0$ dans le système des coordonnées $\{x^{i'}\}$. D'après ce qu'on a démontré, la surface de coordonnées $x^{1'}$, $x^{2'}$ a la courbure égale à celle de la variété M_n dans la direction bidimensionnelle déterminée par le bivecteur x^{ij}.

Il reste à montrer que si les égalités $\Gamma^k_{ij}(A)=0$ ne sont pas remplies dans le système de coordonnées donné $x^1,\,\ldots,\,x^n$, on peut toujours trouver un autre système $x^{1'},\,\ldots,\,x^{n'}$ dans lequel elles seront valables. En supposant comme plus haut que $x^1(A)=\ldots$ $\ldots=x^n(A)=0$, on peut poser, par exemple,

$$x^k=a^k_{k'}x^{k'}+\frac{1}{2}\,b^k_{i'j'}x^{i'j'},$$

avec les coefficients indéterminés pour le moment $a^k_{k'}$ et $b^k_{i'j'}$. Alors

$$p^k_{k'}=\frac{\partial x^k(A)}{\partial x^{k'}}=a^k_{k'},\quad p^k_{i'j'}=\frac{\partial^2 x^k(A)}{\partial x^{i'}\,\partial x^{j'}}=b^k_{i'j'},$$

et, en posant

$$\Gamma^{k'}_{i'j'}(A) = p^i_{i'} p^j_{j'} p^k_{k'} \Gamma^k_{ij}(A) - p^k_{k'} p^k_{i'j'} = 0,$$

on obtient, en simplifiant par $p^k_{k'}$, l'équation

$$p^k_{i'j'} = p^k_{i'j'} = a^i_{i'} a^j_{j'} \Gamma^k_{ij}(A). \tag{5}$$

En choisissant arbitrairement une matrice non dégénérée $a^k_{k'}$ (par exemple, $a^k_{k'} = \delta^k_{k'}$) déterminant les coefficients $b^k_{i'j'}$ selon les formules (5), nous aboutissons au système de coordonnées cherché dans lequel $\Gamma^{k'}_{i'j'}(A) = 0$.

§ 6. 6. Espaces riemanniens de courbure constante

6.61. Soit L une surface n-dimensionnelle dans l'espace euclidien $(n+1)$-dimensionnel R_{n+1} donnée par les équations

$$x^i = x^i(u_1, \ldots, u_n), \quad u = (u_1, \ldots, u_n) \in G \subset R_n.$$

En considérant la surface L comme variété riemannienne, avec la métrique empruntée à l'espace R_{n+1}, trouvons sa courbure dans la direction bidimensionnelle déterminée par un bivecteur x^{ij}. D'après la formule *6.55* (8), on a

$$K = \frac{R_{ij,\,kl} x^{ij} x^{kl}}{G_{ij,\,kl} x^{ij} x^{kl}}. \tag{1}$$

Dans le présent cas, comme nous avons déjà dit, les composantes du tenseur de courbure $R_{ij,\,kl}$ coïncident avec les mineurs $B_{ij,\,kl}$ de la deuxième forme quadratique de la surface L, donc

$$K = \frac{B_{ij,\,kl} x^{ij} x^{kl}}{G_{ij,\,kl} x^{ij} x^{kl}}. \tag{2}$$

6.62. Calculons la courbure K dans le cas où L est la sphère d'un rayon r centrée à l'origine des coordonnées:

$$L = \{x \in R_{n+1} : |x| = r\}.$$

Le rayon vecteur est alors proportionnel à la normale:

$$x = rm. \tag{1}$$

D'après la formule de Weingarten *5.31* (2), on a

$$\frac{\partial m}{\partial u_\alpha} = b^\sigma_\alpha \frac{\partial x}{\partial u_\sigma} \qquad (\alpha = 1, \ldots, n);$$

en substituant à x sa valeur de (1) on trouve

$$m_\alpha = b^\sigma_\alpha r m_\sigma \qquad (\alpha = 1, \ldots, n),$$

d'où $b^\sigma_\alpha r = \delta^\sigma_\alpha$. Il en découle que les coefficients b_{ij} de la deuxième forme quadratique qui sont liés à b^σ_α par la relation *5.31* (3)

$$b_{ij} = -b^\sigma_i g_{\sigma j}$$

prennent la forme

$$b_{ij} = -\frac{1}{r} g_{ij},$$

i.e. se trouvent proportionnels aux coefficients de la première forme quadratique. Par conséquent,

$$B_{ij,\,kl} = \frac{1}{r^2} G_{ij,\,kl},$$

et la formule *6.61* (2) donne

$$K = \frac{1}{r^2}.$$

Ainsi, *la sphère d'un rayon r dans l'espace* $(n+1)$*-dimensionnel a une même courbure* $K = 1/r^2$ *en chaque son point et dans toute direction bidimensionnelle.*

6.63. Un espace riemannien M_n est appelé *espace de courbure constante* si sa courbure en tout point et dans toute direction bidimensionnelle x^{ij},

$$K = \frac{R_{ij,\,kl} x^{ij} x^{kl}}{G_{ij,\,kl} x^{ij} x^{kl}}, \qquad (1)$$

a une même valeur. Nous affirmons, avant tout, que dans cette situation a lieu la relation

$$R_{ij,\,kl} = K G_{ij,\,kl}. \qquad (2)$$

En effet, posons $T_{ij,\,kl} = R_{ij,\,kl} - K G_{ij,\,kl}$. Avec $R_{ij,\,kl}$ et $G_{ij,\,kl}$, le tenseur $T_{ij,\,kl}$ est un tenseur du type Ricci. Conformément à (1), pour tout bivecteur x^{ij}, on a

$$T_{ij,\,kl} x^{ij} x^{kl} = 0.$$

En appliquant le théorème *6.18b* nous voyons *qu'en tout point* $x \in M_n$ *le tenseur* $T_{ij,\,kl}(x)$ *a toutes ses composantes nulles*, d'où (2).

6.64. Supposons ensuite que l'espace M_n ait une courbure constante positive $K = 1/r^2$, $r > 0$. Nous affirmons alors que *l'espace* M_n *est localement isométrique à une sphère de rayon r dans l'espace euclidien* $(n+1)$-dimensionnel R_{n+1}. Pour le démontrer, définissons un tenseur $b_{ij}(x)$ par les formules

$$b_{ij}(x) = \frac{1}{r} g_{ij}(x). \qquad (1)$$

Montrons que, pour les matrices $\| g_{ij}(x) \|$ et $\| b_{ij}(x) \|$ sont satisfaites l'équation de Gauss *5.32* (8) et celle de Peterson — *5.32* (6), i.e. est remplie l'hypothèse du théorème de Bonnet *5.34*. En effet, les mineurs $B_{ij,\,kl}$ de la matrice $\| b_{ij} \|$, en vertu de la définition de b_{ij} (1) et de l'égalité *6.63* (2), vérifient la relation

$$B_{ij,\,kl} = \frac{1}{r^2} G_{ij,\,kl} = R_{ij,\,kl};$$

cela signifie que la relation de Gauss est satisfaite. La formule de Peterson *5.32* (6) qui est à prouver a la forme

$$\frac{\partial b_{ij}}{\partial x^k} - \frac{\partial b_{ik}}{\partial x^j} = \Gamma^s_{ik} b_{sj} - \Gamma^s_{ij} b_{sk}.$$

En y portant les valeurs de b_{ij} de (1) nous sommes amenés à l'égalité

$$\frac{\partial g_{ij}}{\partial x^k} - \frac{\partial g_{ik}}{\partial x^j} = \Gamma_{ik,\,j} - \Gamma_{ij,\,k}. \tag{2}$$

Or,

$$\Gamma_{ik,\,j} = \frac{1}{2}\left(\frac{\partial g_{ij}}{\partial x^k} + \frac{\partial g_{kj}}{\partial x^i} - \frac{\partial g_{ik}}{\partial x^j} \right),$$

$$\Gamma_{ij,\,k} = \frac{1}{2}\left(\frac{\partial g_{ik}}{\partial x^j} + \frac{\partial g_{jk}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^k} \right),$$

et (2) est démontrée en soustrayant l'une des dernières égalités de l'autre.

Selon le théorème de Bonnet *5.34*, il existe dans l'espace R_{n+1} une surface $L = \{x \in R_{n+1} : x = x(u_1, \ldots, u_n)\}$ pour laquelle l'expression $\sum_{i,j} g_{ij}\, du_i du_j$ représente la première, et l'expression $\sum_{i,j} b_{ij}\, du_i\, du_j$, la deuxième forme quadratique. Pour la surface L, nous avons $b^h_j = -g^{ik} b_{ij} = -g^{ik} \frac{1}{r} g_{ij} = -\frac{1}{r} \delta^k_j$, d'où $\frac{\partial m}{\partial u_j} = b^k_j \frac{\partial x}{\partial u_k} = -\frac{1}{r} \frac{\partial x}{\partial u_j}$; ensuite, $\frac{\partial}{\partial u_j}\left(m + \frac{1}{r} x \right) = 0$ implique $m = -\frac{1}{r}(x - x_0)$, d'où $| x - x_0 | = r$. Définitivement, *la surface L qui est isométrique à la variété M_n représente une sphère de rayon r.*

6.65. Passons à la construction d'un espace riemannien M_n d'une courbure constante négative $K = -q^2$, $q > 0$. Etendons au cas $(n+1)$-dimensionnel la construction exposée dans *5.54* où on a réalisé la courbure constante négative sur un hyperboloïde dans l'espace tridimensionnel, avec la métrique induite par la forme $\langle x, x \rangle = x_1^2 + x_2^2 - x_3^2$.

Désignons par H l'hyperboloïde $(x^1)^2 + \ldots + (x^n)^2 - (x^{n+1})^2 = -\rho^2$ dans l'espace euclidien $(n+1)$-dimensionnel R_{n+1}. Dotons-le de la métrique induite par le produit pseudo-scalaire $\langle x, y\rangle = x^1y^1 + \ldots + x^ny^n - x^{n+1}y^{n+1}$. Nous affirmons que cette métrique est définie positive sur H, i.e. que la forme

$$dx^1\,dx^1 + \ldots + dx^n\,dx^n - dx^{n+1}\,dx^{n+1}$$

ne prend que les valeurs positives sur les vecteurs tangents à la surface H. En effet, nous avons sur H:

$$x^1\,dx^1 + \ldots + x^n\,dx^n - x^{n+1}\,dx^{n+1} = 0.$$

Il en résulte selon l'inégalité de Cauchy-Bouniakovski:

$$\begin{aligned}(x^{n+1})^2\,(dx^{n+1})^2 &= (x^1\,dx^1 + \ldots + x^n\,dx^n)^2 \leqslant \\ &\leqslant [(x^1)^2 + \ldots + (x^n)^2]\,[(dx^1)^2 + \ldots + (dx^n)^2] = \\ &= [(x^{n+1})^2 - \rho^2]\,[(dx^1)^2 + \ldots + (dx^n)^2] \leqslant \\ &\leqslant (x^{n+1})^2\,[(dx^1)^2 + \ldots + (dx^n)^2],\end{aligned}$$

donc

$$(dx^1)^2 + \ldots + (dx^n)^2 - (dx^{n+1})^2 \geqslant 0 \qquad (1)$$

ce qu'on affirmait.

6.66. Soit L une surface n-dimensionnelle dans le cône $\langle x, x\rangle < 0$, pour laquelle l'inégalité *6.65* (1) ait lieu, et la forme $\langle x, y\rangle$ induise une métrique définie positive. Pour de telles surfaces, le vecteur de la pseudo-normale m vérifie l'inégalité $\langle m, m\rangle < 0$, sinon, la forme $\langle x, x\rangle$ étant positive dans le plan n-dimensionnel tangent et non négative sur le vecteur m pseudo-orthogonal à ce plan, elle serait non négative sur tout le R_{n+1}, ce qui n'est pas.

Trouvons la courbure de la surface L en tant qu'espace riemannien dans les directions bidimensionnelles $[dx, dy]^{ij}$. A cette fin, écrivons d'abord les formules de Gauss et de Weingarten pour le rayon vecteur de la surface $x = x(u_1, \ldots, u_n)$ (paramétrisée d'une façon arbitraire):

$$x_{ij} \equiv \frac{\partial^2 x}{\partial u_i\,\partial u_j} = \Gamma^k_{ij}x_k + \beta_{ij}m,$$

$$m_j \equiv \frac{\partial m}{\partial u_j} = b^k_j x_k.$$

Désignons $b_{ij} = \langle x_{ij}, m\rangle$. De même que dans le cas de surface dans l'espace euclidien, nous avons $\langle m, x_i\rangle = 0$, $\langle m_j, x_k\rangle + \langle m, x_{jk}\rangle = 0$, d'où

$$b_{jk} = -\langle m_j, x_k\rangle = -\langle b^s_j x_s, x_k\rangle = -b^s_j g_{sk}. \qquad (1)$$

Ensuite, comme dans *5.32*, nous obtenons:

$$\beta_{ik}b_{jl} - \beta_{jk}b_{il} = \sum_{p=1}^{n-1}\left[\frac{\partial\Gamma^p_{ik}}{\partial u_j} - \frac{\partial\Gamma^p_{jk}}{\partial u_i} + \sum_{s=1}^{n-1}(\Gamma^s_{ik}\Gamma^p_{js} - \Gamma^s_{jk}\Gamma^p_{is})\right]g_{pl}. \qquad (2)$$

Le second membre dans (2) représente le tenseur de courbure $R_{ij,\,kl}$. Les coefficients β_{ij} dans le premier membre, contrairement au cas euclidien où $\beta_{ij} = (x_{ij}, m) = b_{ij}$, se déterminent de la relation

$$b_{ij} = \langle x_{ij}, m\rangle = \beta_{ij}\,\langle m, m\rangle = -\beta_{ij}.$$

A la place d'un mineur de la deuxième forme quadratique occupant les lignes i, k et les colonnes j, l, ce qu'on avait dans le cas euclidien, nous avons maintenant le même mineur mais précédé du signe moins, i.e.

$$R_{ij,\,kl} = -B_{ij,\,kl}.$$

La quantité cherchée, c'est-à-dire la courbure de l'espace riemannien L dans la direction bidimensionnelle x^{ij}, s'avère égale à

$$K = -\frac{B_{ij,\,kl}\,[dx,\,dy]^{ij}\,[dx,\,dy]^{kl}}{G_{ij,\,kl}\,[dx,\,dy]^{ij}\,[dx,\,dy]^{kl}}. \tag{3}$$

6.67. Dans le cas spécial d'une pseudo-sphère

$$L = \{x \in R_{n+1}\colon\ \langle x, x\rangle = -\rho^2\},$$

nous avons $x = \rho m$, d'où

$$x_j = \rho m_j = \rho b_j^k x_k,$$

de sorte que $b\,\frac{k}{j} = \frac{1}{\rho}\,\delta_j^k$; il en découle $b_{ij} = \langle x_{ij}, m\rangle = -b_{ig}^k g_{jk} = -\frac{1}{\rho}\,g_{ij}$, et, par conséquent,

$$B_{ij,\,kl} = \frac{1}{\rho^2}\,G_{ij,\,kl}.$$

Ceci fait, (3) devient

$$K = -\frac{1}{\rho^2}.$$

Ainsi, la pseudo-sphère L nous fournit un exemple de variété riemannienne n-dimensionnelle de courbure négative ayant la valeur constante $-1/\rho^2$ dans toutes les directions bidimensionnelles.

6.68.a. Inversement, soit M_n un espace riemannien ayant, en tout point et dans toute direction bidimensionnelle, une même valeur négative $K = -q^2$, $q > 0$. Montrons qu'*un tel espace M_n est isométrique à une partie d'une pseudo-sphère $L \subset R_{n+1}$*.

b. Nous aurons besoin du théorème de Bonnet pour une surface dans l'espace R_{n+1} dont la métrique riemannienne est induite par la forme $\langle x, x\rangle$.

Voici l'énoncé de ce théorème:

Théorème. *Supposons qu'il y ait deux $n \times n$-matrices $G = \| g_{ij}(u) \|$ et $B = \| b_{ij}(u) \|$, $u = (u_1, \ldots, u_n) \in G \subset R_n$ satis-*

faisant à la condition

$$b_{jk}b_{il}-b_{ik}b_{jl}=\sum_{p=1}^{n-1}\left[\frac{\partial\Gamma_{ik}^{p}}{\partial u_j}-\frac{\partial\Gamma_{jk}^{p}}{\partial u_i}+\sum_{s=1}^{n-1}\left(\Gamma_{ik}^{s}\Gamma_{js}^{p}-\Gamma_{jk}^{s}\Gamma_{is}^{p}\right)\right]g_{pl}, \qquad (1)$$

où

$$\Gamma_{ij}^{p}g_{pk}=\frac{1}{2}\left(\frac{\partial g_{ik}}{\partial u_j}+\frac{\partial g_{jk}}{\partial u_k}-\frac{\partial g_{ij}}{\partial u_k}\right),$$

et à la condition

$$\sum_{p}\Gamma_{ij}^{p}b_{pk}+\frac{\partial b_{ij}}{\partial u_k}=\sum_{p}\Gamma_{ik}^{p}b_{pj}+\frac{\partial b_{ik}}{\partial u_j}. \qquad (2)$$

Alors il existe dans l'espace R_{n+1} une surface n-dimensionnelle L munie de la métrique (définie positive) qui est induite par le produit pseudo-scalaire $\langle x, y\rangle$ et pour laquelle la matrice G est celle de la première forme quadratique, et la matrice B est celle de la deuxième.

La démonstration suit la même voie que dans l'espace euclidien (*5.34*). On écrit le système d'équations

$$\frac{\partial x_i}{\partial u_j}=\Gamma_{ij}^{k}x_k-b_{ij}m,$$

$$\frac{\partial m}{\partial u_j}=b_j^{k}x_k$$

où $b_j^k g_{kp}=-b_{jp}$, avec les fonctions vectorielles inconnues $x_i(u)$ et $m(u)$. Les conditions de compatibilité de ce système coïncident avec (1), (2). Il existe donc une solution $x_i(u)$, $m(u)$, elle est unique si l'on a les conditions initiales $x_i(u_0)$ et $m(u_0)$ qui vérifient les relations

$$\langle x_i(u_0), x_j(u_0)\rangle=g_{ij}(u_0),\ m(u_0)=(0,\ldots,0,1).$$

On détermine ensuite la fonction $x(u)$ à partir des conditions

$$\frac{\partial x(u)}{\partial u_i}=x_i(u),\ x(0)=0.$$

C'est la surface cherchée, ce qu'on prouve par le même raisonnement que dans *5.34*.

c. Passons à la démonstration de la proposition *a*. Dans l'espace M_n de courbure constante dans toutes les directions bidimensionnelles

$$K=\frac{R_{ij,\,kl}x^{ij}x^{kl}}{G_{ij,\,kl}x^{ij}x^{kl}}=-q^2,\ q>0,$$

on a, conformément à *6.63*:

$$R_{ij,\,kl}=-q^2G_{ij,\,kl}.$$

Posons $b_{ij} = -qg_{ij}$ et montrons que, pour les formes b_{ij} et g_{ij}, l'hypothèse du théorème de Bonnet dans l'espace pseudo-euclidien (b) est remplie. En effet, on a par construction :

$$b_{ij}b_{kl} - b_{ik}b_{jk} = q^2 (g_{ij}g_{kl} - g_{ik}b_{jl}) = q^2 G_{ij,\,kl} = -R_{ij,\,kl},$$

d'où la condition (1).

On vérifie la condition (2) tout comme dans *6.64*.

L'hypothèse du théorème de Bonnet étant remplie, il existe une surface $L \subset R_{n+1}$ dont la métrique est induite par la forme $\langle x, x\rangle$ et pour laquelle $\| g_{ij} \|$ représente la matrice de la première forme quadratique et $\| b_{ij} \|$, de la deuxième. Pour cette surface L, nous avons $b_j^k g_{kp} = -b_{jp} = qg_{jp}$, d'où $b_j^k = q\delta_j^k$; par conséquent,

$$m_j = b_j^k x_k = qx_j, \qquad (m - qx)_j = 0,$$

donc

$$m = q\,(x - x_0).$$

Il en résulte

$$\langle x - x_0,\ x - x_0\rangle = \left\langle \frac{m}{q},\ \frac{m}{q} \right\rangle = -\frac{1}{q^2},$$

de sorte que la surface L se trouve sur la pseudo-sphère $\langle x, x\rangle = -1/q^2$ translatée d'un vecteur x_0.

Ainsi, *l'hyperboloïde* $\langle x, x\rangle = -1/q^2$, *avec la métrique induite par la forme* $\langle x, x\rangle$, *représente un modèle canonique de l'espace riemannien de courbure constante négative* $-q^2$.

EXERCICES

1. Montrer que deux sphères S_1 et S_2 de rayons différents dans R_3 ne sont pas isométriques mais affinement équivalentes (c'est-à-dire que, dans certains systèmes de coordonnées sur S_1 et S_2, les coefficients de la connexion riemannienne s'expriment par les mêmes fonctions de coordonnées).

2. Prouver la formule de dérivation covariante d'un produit de tenseurs :

$$\nabla_q (TS) = \nabla_q T \cdot S + T \cdot \nabla_q S.$$

3. Prouver la formule de dérivation covariante d'une contraction de tenseurs :

$$\nabla_q (T^i S_i) = \nabla_q T^i \cdot S_i + T^i \cdot \nabla_q S_i.$$

4. Deux champs vectoriels orthogonaux l'un à l'autre sont fixés dans le plan. En tout point, deux vecteurs des champs déterminent une base locale. Définissons la translation de n'importe quel vecteur par la condition que ses coordonnées locales soient constantes. Montrer que la connexion correspondante a une courbure nulle mais, en général, une torsion non nulle.

5. Dans un domaine G sur le plan, posons $\Gamma_{11}^1(x) = f(x)$, tous les autres Γ_{ij}^k étant nuls. Définir $f(x)$ d'une telle manière que, lorsqu'un vecteur est translaté (au sens de la connexion Γ_{ij}^k) plusieurs fois le long d'un contour fermé et arbitrairement petit qui entoure un point donné $A \subset G$, les coordonnées de ce vecteur croissent indéfiniment.

6. En un point donné A d'un espace M_n à connexion affine, choisissons arbitrairement n vecteurs linéairement indépendants $e_1, \ldots, e_n$. Soit U un voisinage du point A dont tout point B puisse être joint au point A par une et une seule géodésique $\gamma(A, B)$ passant dans U (exercice 12, chap. 5). La géodésique $\gamma(A, B)$ est déterminée par son vecteur tangent au point A, par exemple par le vecteur $\xi = \sum_{i+1}^{n} \xi^i e_i$. Désignons par τ_B la valeur au point B du paramètre canonique sur $\gamma(A, B)$ qui est déterminé par l'égalité $\frac{dx(A)}{d\tau} = \xi$. Choisissons pour nouvelles coordonnées du point B les nombres $x^{i'} = \tau_B \cdot \xi^i$. Montrer que, dans ces coordonnées, $\Gamma^k_{ij}(A) = 0$.

Historique

La géométrie riemannienne a pour point de départ la conférence de Riemann « Über die Hypothesen, welche der Geometrie zu Grund liegen » (1854, publiée en 1867). Riemann y réunit l'idée d'espace n-dimensionnel avec l'idée de Gauss de métrisation d'une surface à l'aide d'une forme quadratique des différentielles de coordonnées. Dans la même conférence, Riemann propose une définition de la courbure qui est ensuite développée, sous forme analytique, par Christoffel (1869). L'analyse tensorielle créée par Ricci dans les années 1880 (« Méthodes de calcul différentiel absolu et leurs applications », l'ouvrage commun de Ricci et de son disciple Levi-Civita, 1901) se trouve le mieux faite pour servir de formalisme analytique à la géométrie riemannienne. En particulier, Ricci et Levi-Civita indiquent les formes canoniques des espaces riemanniens de courbure constante. Schur remarque (1903) que si, en tout point d'un espace riemannien M_n, la courbure a une même valeur pour toutes les directions bidimensionnelles, alors elle ne dépend pas non plus du point. Le langage tensoriel s'avère très commode dans la théorie relativiste générale d'Einstein (1915), mais on y est forcé de passer de la forme métrique définie positive ds^2 à une forme indéfinie (avec le signe moins devant l'un des carrés dans la représentation canonique). En 1917, Levi-Civita et Schouten introduisent dans la géométrie riemannienne la notion de translation, ce qui leur permet de trouver une nouvelle expression du tenseur de courbure. Les espaces à connexion affine sont introduits par Schouten et H. Weyl en 1918. Les variétés différentiables apparaissent pour la première fois chez Whitney (1936).

Différentiation et intégration sur les variétés

L'analyse mathématique sur les variétés différentiables représente actuellement un domaine très vaste des mathématiques, un domaine où les idées et les méthodes de différentes branches de la science s'intercroisent. Nous nous bornons ici à un seul, mais important, chapitre de la théorie: il s'agit de l'extension au cas de variété différentiable élémentaire de la relation entre la dérivation et l'intégration qui est donnée dans l'espace R_3 par la formule de Stokes (chap. 4), ainsi que de la position et de la résolution des problèmes directs et inverses correspondants. L'analogue multidimensionnel de l'analyse vectorielle classique est bien l'analyse tensorielle, avec des tenseurs de n'importe quel rang (nous aurons surtout besoin de tenseurs covariants pour décrire les formes multilinéaires). Il s'avère que les opérations vectorielles différentielles, i.e. gradient, divergence, rotationnel, trouvent leur généralisation en une seule opération différentielle sur les champs de formes multilinéaires, à savoir en la différentiation antisymétrique (§ 7.2). La formule de Stokes prend une forme générale et en même temps parfaitement simple: l'intégrale sur un domaine de la différentielle d'une forme est égale à l'intégrale sur la frontière du domaine de cette forme elle-même (§ 7.3). Le chapitre est conclu par une généralisation du problème inverse de l'analyse vectorielle, i.e. du problème de restitution d'un champ vectoriel d'après sa divergence et son rotationnel: ce problème devient celui de restitution (locale) d'une forme d'après sa différentielle et sa codifférentielle; la résolution de ce dernier est assez simple, en utilisant, en fin des comptes, des méthodes de l'analyse vectorielle (§ 7.4).

§ 7. 1. Formes antisymétriques

7.11.a. Multinumération. Appelons *numéro* n'importe quel nombre naturel 1, 2, ... Appelons *multinuméro*, plus exactement, (k-n)-numéro chaque suite $(i) = (i_1, \ldots, i_k)$ de k numéros qui ne dépassent pas le nombre n. Les numéros $i_1, \ldots, i_k$ s'appellent *composantes* du (k-n)-numéro $(i) = (i_1, \ldots, i_k)$. Deux (k-n)-numéros

$$(i) = (i_1, \ldots, i_k) \quad \text{et} \quad (j) = (j_1, \ldots, j_k)$$

sont dits égaux si $i_1 = j_1, \ldots, i_k = j_k$, et différents si l'inégalité $i_p \neq j_p$ a lieu pour un numéro $p \leqslant k$ au moins. Un multinuméro $(i) = (i_1, \ldots, i_k)$ est dit *ordonné* si $i_1 \leqslant i_2 \leqslant \ldots \leqslant i_k$, *strictement* si tous les numéros $i_1, \ldots, i_k$ sont différents, et *strictement ordonné* si $i_1 < i_2 < \ldots < i_k$. On calculera aisément que, k et n étant

donnés, il y a n^k (k-n)-numéros différents, C_{n+k-1}^k ordonnés, $n(n-1)\ldots(n-k+1)$ stricts et C_n^k strictement ordonnés. En particulier, pour $k=1$, il existe n (1-n)-numéros qui représentent tout simplement les numéros $1, \ldots, n$; bien que les définitions précédentes n'y sont pas appliquables directement, on les considère comme strictement ordonnés. Pour $k=n$, il existe n^n (n-n)-numéros différents, C_{2n-1}^n ordonnés, $n!$ stricts et un seul ($C_n^n=1$) strictement ordonné, c'est le (n-n)-numéro $(1, 2, \ldots, n)$. Signalons encore n (($n-1$)-n)-numéros strictement ordonnés:

$$(\hat{1}, 2, 3, \ldots, n), (1, \hat{2}, 3, \ldots, n), \ldots, (1, 2, 3, \ldots, \hat{n})$$

où le symbole $\hat{}$ veut dire que la composante correspondante est à enlever à la suite dans la parenthèse. Pour $k>n$, il ne reste plus de (k-n)-numéros stricts, ni, à fortiori, strictement ordonnés.

b. Chaque (k-n)-numéro $(i)=(i_1, \ldots, i_k)$ peut être mis sous forme d'un (k-n)-numéro ordonné $(\alpha)=(\alpha_1, \ldots, \alpha_k)$ en permutant convenablement ses composantes; une telle permutation s'appelle *ordination* et se note $O(i)$. Avec une ordination, chaque (k-n)-numéro (i) strict se transforme en un (k-n)-numéro (α) strictement ordonné, la permutation $O(i)$ étant dans ce cas définie d'une façon univoque. Le nombre $\varepsilon_{i_1 \ldots i_k}^{\alpha_1 \ldots \alpha_k} \equiv \varepsilon_{\alpha_1 \ldots \alpha_k}^{i_1 \ldots i_k}$ est, par définition, le signe de cette permutation. S'il y a une $k \times k$-matrice $\| a_{\alpha_p}^j \|$ dont les numéros des lignes sont $j=1, \ldots, k$, et ceux des colonnes α_p, $p=1, \ldots, k$, $\alpha_1 < \ldots < \alpha_k$, son déterminant est calculé d'après la formule

$$\det \| a_{\alpha_p}^j \| = \sum_{O(i)=(\alpha)} \varepsilon_{i_1 \ldots i_k}^{\alpha_1 \ldots \alpha_k} a_{i_1}^1 \ldots a_{i_k}^k \qquad (1)$$

(sommation sur tous les multinuméros (i) pour lesquels $O(i)=(\alpha)$).

c. Un multinuméro $(j)=(j_1, \ldots, j_{n-k})$ dont les composantes complètent celles d'un (k-n)-numéro strict $(i)=(i_1, \ldots, i_k)$ jusqu'au n-tuple complet $1, 2, \ldots, n$ s'appelle *complémentaire* du (k-n)-numéro (i). Ce multinuméro (j) est défini d'une façon univoque si l'on exige qu'il soit ordonné.

d. En tant qu'exemple (dont nous aurons besoin par la suite), calculons le signe de la permutation qui ordonne un (n-n)-numéro de la forme $(\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_{n-k})$, où $(\beta)=(\beta_1, \ldots, \beta_{n-k})$ est le numéro strictement ordonné, complémentaire d'un numéro strictement ordonné $(\alpha)=(\alpha_1, \ldots, \alpha_k)$. L'ordination du ($n$-$n$)-numéro indiqué peut être réalisée comme suit. Tout d'abord, mettons le numéro α_k qui occupe la k-ème place à sa propre α_k-ème place. Le nombre α_k est au moins égal à k parce que le multinuméro (α) est strictement ordonné. Cette opération nous prend $\alpha_k - k$ inversions (permutations de deux éléments voisins) qui ne modifient pas les positions relatives des numéros β_j. Ensuite, transférons le numéro α_{k-1} de la $(k-1)$-ème place à celle de numéro

$\alpha_{k-1} \geqslant k-1$, ce qui prend $\alpha_{k-1} - (k-1)$ inversions; en continuant de la même manière nous finissons par transférer le numéro α_1 de la première place à la place de numéro α_1 en y mettant $\alpha_1 - 1$ inversions. Lorsque tous les numéros $\alpha_1, \ldots, \alpha_k$ prennent leurs places, il en est automatiquement de même des numéros $\beta_1, \ldots, \ldots, \beta_{n-k}$, car ils conservent leurs positions relatives et occupent toutes les places qui restent. Définitivement, pour ordonner le multinuméro $(\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_{n-k})$, nous avons mis

$$\alpha_1 + \ldots + \alpha_k - (1 + \ldots + k) = \sum_{j=1}^{k} \alpha_j - \frac{k(k+1)}{2}$$

inversions, d'où

$$\varepsilon^{1, \ldots, k, k+1, \ldots, n}_{\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_{n-1}} = (-1)^{\sum_{j=1}^{k} \alpha_j - \frac{k(k+1)}{2}}. \qquad (2)$$

7.12. Formes multilinéaires.

a. Une fonction $A(\underset{1}{x}, \ldots, \underset{k}{x})$ de k vecteurs $\underset{1}{x}, \ldots, \underset{k}{x}$ d'un espace vectoriel R est dite *multilinéaire*, plus exactement *k-linéaire*, si elle est linéaire par rapport à chaque variable, les autres étant fixes.

b. Dans l'espace R_n, toute fonction k-linéaire s'appelle *k-forme*, le nombre k s'appelle *degré* de la forme. Trouvons l'expression générale d'une k-forme $A(\underset{1}{x}, \ldots, \underset{k}{x})$ dans l'espace n-dimensionnel R_n. A cette fin, choisissons arbitrairement une base $e_1, \ldots, e_n$, et désignons les coordonnées d'un vecteur $\underset{i}{x}$ dans cette base par $\underset{i}{\xi^1}, \ldots, \underset{i}{\xi^n}$. En vertu de la propriété de multilinéarité, nous avons

$$A(\underset{1}{x}, \ldots, \underset{k}{x}) = A\left(\sum_{i_1} \underset{1}{\xi^{i_1}} e_{i_1}, \ldots, \sum_{i_k} \underset{k}{\xi^{i_k}} e_{i_k}\right) =$$

$$= \sum_{(i)} \underset{1}{\xi^{i_1}} \ldots \underset{k}{\xi^{i_k}} A(e_{i_1}, \ldots, e_{i_k}) = \sum_{(i)} a_{(i)} \underset{1}{\xi^{i_1}} \ldots \underset{k}{\xi^{i_k}}, \qquad (1)$$

où $a_{(i)} = A(e_{i_1}, \ldots, e_{i_k})$. La réciproque est évidente: chaque fonction de k vecteurs $\underset{1}{x}, \ldots, \underset{k}{x}$ de l'espace R_n qui s'exprime par les coordonnées de ces vecteurs comme dans le dernier membre de (1), avec n'importe quels coefficients $a_{(i)}$, représente une k-forme.

c. Evidemment, on peut additionner les k-formes dans R_n et les multiplier par les nombres en obtenant toujours des k-formes. Ainsi, les k-formes dans R_n composent un nouvel espace vectoriel. Sa base est composée par les k-formes $\underset{1}{\xi^i_k} \ldots \underset{k}{\xi^{i_k}}$ dont le nombre est égal à celui de $(k\text{-}n)$-numéros (i), i.e. à n^k (*7.11, a*). L'indépendance

linéaire de ces k-formes résulte de l'unicité de la représentation (1): en y posant $x = e_{j_1}, \ldots, x = e_{j_k}$, nous arrivons nécessairement à l'expression $a_{(j)} = A(e_{j_1}, \ldots, e_{j_k})$. Ainsi, *la dimension de l'espace de toutes les k-formes dans R_n vaut n^k.*

d. Voyons comment se transforment les coefficients $a_{(i)}$ en passant à une nouvelle base $e_{i'} = p^i_{i'} e_i$ (*6.11*). Nous avons

$$a_{(i')} = A(e_{i'_1}, \ldots, e_{i'_k}) = p^{i_1}_{i'_1} \ldots p^{i_k}_{i'_k} A(e_{i_1}, \ldots, e_{i_k}) = p^{i_1}_{i'_1}, \ldots, p^{i_k}_{i'_k} a_{(i)}.$$

Ainsi, *les coefficients $a_{(i)}$ forment un tenseur k fois covariant.*

e. Introduisons encore la notion de 0-forme: par définition, une 0-forme est une fonction de $x \in R_n$ qui est constante. Toutes les 0-formes composent donc un espace de dimension 1.

7.13. k-**formes antisymétriques.**

a. Une k-forme $A(x, \ldots, x)$ qui change de signe en permutant ses deux variables indépendantes est dite *antisymétrique.*

Il est évident qu'addition et multiplication par un nombre des k-formes antisymétriques amènent à des k-formes antisymétriques. Ainsi, les k-formes antisymétriques représentent un sous-espace vectoriel dans l'espace de toutes les k-formes.

b. Soit

$$A(x, \ldots, x) = \sum_{(i)} a_{(i)} \xi^{i_1} \ldots \xi^{i_k}$$

une k-forme antisymétrique. Comme $a_{(i)} = A(e_{i_1}, \ldots, e_{i_k})$, les coefficients $a_{(i)}$ sont antisymétriques par rapport aux composantes $i_1, \ldots, i_k$ du $(k\text{-}n)$-numéro (i), i.e. changent de signe en permutant deux quelconques composantes de ce $(k\text{-}n)$-numéro; en particulier, si ce dernier n'est pas strict, c'est-à-dire a deux composantes égales, alors le coefficient $a_{(i)}$ est nul. Puisque, pour $k > n$, il n'y a point de $(k\text{-}n)$-numéros stricts, *toute k-forme antisymétrique avec $k > n$ est identiquement nulle.* Si $k \leqslant n$, alors $a_{(i)} = \varepsilon^{\alpha_1 \ldots \alpha_k}_{i_1 \ldots i_k} a_{(\alpha)}$, où $(\alpha) = O(i)$ et l'on peut mettre $A(x, \ldots, x)$ sous la forme:

$$A(x, \ldots, x) = \sum_{(i)} a_{(i)} \xi^{i_1} \ldots \xi^{i_k} =$$

$$= \sum_{(\alpha)} a_{(\alpha)} \sum_{O(i)=(\alpha)} \varepsilon^{\alpha_1 \ldots \alpha_k}_{i_1 \ldots i_k} \xi^{i_1} \ldots \xi^{i_k} = \sum_{(\alpha)} a_{(\alpha)} D_{(\alpha)}(x, \ldots, x), \qquad (1)$$

où

$$D_{(\alpha)}(x, \ldots, x) = \begin{vmatrix} \xi^{\alpha_1} & \ldots & \xi^{\alpha_1} \\ \cdot & \cdots & \cdot \\ \xi^{\alpha_k} & \ldots & \xi^{\alpha_k} \end{vmatrix}. \qquad (2)$$

Les formes $D_{(\alpha)}(\underset{1}{x}, \ldots, \underset{k}{x})$ (2) qui sont évidemment antisymétriques, sont appelées k-formes *antisymétriques canoniques*; leur nombre est égal à celui de $(k\text{-}n)$-numéros strictement ordonnés $(\alpha) = (\alpha_1, \ldots, \alpha_k)$, c'est-à-dire à C_n^k (*7.11a*). Comme les coefficients $a_{(\alpha)}$ dans (1) sont définis d'une façon univoque, $a_{(\alpha)} = A(e_{\alpha_1}, \ldots, e_{\alpha_n})$, *les formes* $D_{(\alpha)}$ *sont linéairement indépendantes, donc, représentent une base dans le sous-espace de toutes les k-formes antisymétriques.* Il en découle que *la dimension de ce sous-espace vaut* C_n^k.

c. La seule (à la dépendance linéaire près) n-forme antisymétrique ($C_n^n = 1$) est

$$D_{(1, \ldots, n)}(\underset{1}{x}, \ldots, \underset{n}{x}) = \begin{vmatrix} \xi^1_1 & \ldots & \xi^1_n \\ \ldots & \ldots & \ldots \\ \xi^n_1 & \ldots & \xi^n_n \end{vmatrix}.$$

Si l'espace R_n est euclidien, et que les coordonnées ξ^i se rapportent à une base orthonormale, alors la valeur absolue de la forme $D_{(1, \ldots n)}(\underset{1}{x}, \ldots, \underset{n}{x})$ s'interprète géométriquement comme volume du parallélépipède construit sur les vecteurs $\underset{1}{x}, \ldots, \underset{n}{x}$.

d. La définition d'antisymétrie n'est pas directement applicable aux 0-formes et 1-formes. Tout de même, nous les considérons, par définition, comme antisymétriques.

7.14. Alternation d'une k-forme.

a. Soit $A(\underset{1}{x}, \ldots, \underset{k}{x})$ une k-forme arbitraire; par définition,

$$\operatorname{Alt} A(\underset{1}{x}, \ldots, \underset{k}{x}) = \frac{1}{k!} \sum_{(i)} \varepsilon^{1 \ldots k}_{i_1 \ldots i_k} A(\underset{i_1}{x}, \ldots, \underset{i_k}{x}), \qquad (1)$$

avec la sommation sur tous les $(k\text{-}n)$-numéros (i) stricts. Le second membre représente la moyenne arithmétique des valeurs $A(\underset{i_1}{x}, \ldots, \underset{i_k}{x})$, pour toutes les permutations du $(k\text{-}k)$-numéro $(1, \ldots, k)$, pourvues de signes des permutations correspondantes. C'est encore une k-forme appelée *alternation* de la k-forme initiale $A(\underset{1}{x}, \ldots, \underset{k}{x})$. Cette forme est déjà antisymétrique, quelle que soit la forme initiale $A(\underset{1}{x}, \ldots, \underset{k}{x})$. Considérons, par exemple, la permutation de $\underset{1}{x}$ et $\underset{2}{x}$. Examinons isolément la somme de deux termes suivants du second membre de (1):

$$s = \varepsilon^{1 \ldots \ldots \ldots k}_{i_1 \ldots 1 \ldots 2 \ldots i_k} A(\underset{i_1}{x}, \ldots, \underset{1}{x}, \ldots, \underset{2}{x}, \ldots, \underset{i_k}{x}) +$$

$$+ \varepsilon^{1 \ldots \ldots \ldots k}_{i_1 \ldots 2 \ldots 1 \ldots i_k} A(\underset{i_1}{x}, \ldots, \underset{2}{x}, \ldots, \underset{1}{x}, \ldots, \underset{i_k}{x}),$$

les positions des autres numéros 3, . . ., k étant fixes. En permutant x_1 et x_k cette somme deviendra

$$\widetilde{s} = \varepsilon^{1\ldots k}_{i_1\ldots 1\ldots 2\ldots i_k} A(x_{i_1}, \ldots, x_2, \ldots, x_1, \ldots, x_{i_k}) +$$

$$+ \varepsilon^{1\ldots k}_{i_1\ldots 2\ldots 1\ldots i_k} A(x_{i_1}, \ldots, x_1, \ldots, x_2, \ldots, x_{i_k}) =$$

$$= -\varepsilon^{1\ldots k}_{i_1\ldots 2\ldots 1\ldots i_k} A(x_{i_1}, \ldots, x_2, \ldots, x_1, \ldots, x_{i_k}) -$$

$$- \varepsilon^{1\ldots k}_{i_1\ldots 1\ldots 2\ldots i_k} A(x_{i_1}, \ldots, x_1, \ldots, x_2, \ldots, x_{i_k}) = -s.$$

Il en résulte que toute l'expression $\operatorname{Alt} A(x_1, \ldots, x_k)$ change de signe en permutant x_1 et x_2.

Si une forme $A(x_1, \ldots, x_k)$ est déjà antisymétrique, son alternation ne la change pas, car tous les termes de la somme (1) coïncident:

$$\varepsilon^{1\ldots k}_{i_1\ldots i_k} A(x_{i_1}, \ldots, x_{i_k}) = A(x_1, \ldots, x_k).$$

b. Trouvons, par exemple, $\operatorname{Alt} \xi^{j_1}_1 \ldots \xi^{j_k}_k$, où $(j) = (j_1, \ldots, j_k)$ est un $(k\text{-}n)$-numéro fixe. Conformément à la définition, on a

$$\operatorname{Alt} \xi^{j_1}_1 \ldots \xi^{j_k}_k = \frac{1}{k!} \sum_{(i)} \varepsilon^{1\ldots k}_{i_1\ldots i_k} \xi^{j_1}_{i_1} \ldots \xi^{j_k}_{i_k} =$$

$$= \frac{1}{k!} \det \| \xi^{j_p}_r \| = \frac{1}{k!} \begin{vmatrix} \xi^{j_1}_1 & \ldots & \xi^{j_1}_k \\ \ldots & \ldots & \ldots \\ \xi^{j_k}_1 & \ldots & \xi^{j_k}_k \end{vmatrix} =$$

$$= \frac{1}{k!} \sum_{(\nu)} \varepsilon^{j_1\ldots j_k}_{\nu_1\ldots\nu_k} \xi^{\nu_1}_1 \ldots \xi^{\nu_k}_k. \qquad (2)$$

Si le multinuméro (j) est strictement ordonné, par exemple $(j) = (\alpha) = (\alpha_1, \ldots, \alpha_k)$, on a définitivement

$$\operatorname{Alt} \xi^{\alpha_1}_1 \ldots \xi^{\alpha_k}_k = \frac{1}{k!} \begin{vmatrix} \xi^{\alpha_1}_1 & \ldots & \xi^{\alpha_1}_k \\ \ldots & \ldots & \ldots \\ \xi^{\alpha_k}_1 & \ldots & \xi^{\alpha_k}_k \end{vmatrix} = \frac{1}{k!} D_{(\alpha)}(x_1, \ldots, x_k), \qquad (3)$$

où $D_{(\alpha)}(x_1, \ldots, x_k)$ est une forme antisymétrique canonique (7.13).

c. Signalons encore la propriété linéaire évidente de l'alternation: pour deux k-formes $A_1(x_1, \ldots, x_k)$ et $A_2(x_1, \ldots, x_k)$ et pour

deux nombres a_1 et a_2, on a

$$\operatorname{Alt}(a_1A_1 + a_2A_2) = a_1 \operatorname{Alt} A_1 + a_2 \operatorname{Alt} A_2. \tag{4}$$

7.15. Produits tensoriels des formes multilinéaires.

a. Soient $A(\underset{1}{x}, \ldots, \underset{k}{x})$ une k-forme et $B(\underset{1}{x}, \ldots, \underset{m}{x})$ une m-forme. Leur *produit tensoriel* est la $(k+m)$-forme $C(\underset{1}{x}, \ldots, \underset{k+m}{x})$ définie par l'égalité

$$C(\underset{1}{x}, \ldots, \underset{k+m}{x}) = A(\underset{1}{x}, \ldots, \underset{k}{x})\, B(\underset{k+1}{x}, \ldots, \underset{k+m}{x}). \tag{1}$$

La relation (1) se note en abrégé: $C = A \times B$.

b. Comme le produit tensoriel s'exprime par les produits ordinaires numériques des valeurs des formes, il possède la propriété linéaire: si $A = a_1A_1 + a_2A_2$, alors

$$A \times B = (a_1A_1 + a_2A_2) \times B = a_1A_1 \times B + a_2A_2 \times B.$$

c. Même si A et B sont des formes antisymétriques, la forme $A \times B$ n'est pas forcément antisymétrique. Si l'on veut en avoir une, on alterne le produit tensoriel; on aboutit à une nouvelle forme

$$D(\underset{1}{x}, \ldots, \underset{k+m}{x}) = \operatorname{Alt}[A(\underset{1}{x}, \ldots, \underset{k}{x})\, B(\underset{k+1}{x}, \ldots, \underset{k+m}{x})], \tag{2}$$

ou, en abrégé, $D = A \wedge B$, qui s'appelle *produit tensoriel alterné des formes A et B* (soulignons qu'il est défini pour n'importe quelles formes A et B, non seulement pour les antisymétriques !).

d. Calculons, par exemple, $A \wedge B$ où

$$A = \underset{1}{\xi^{i_1}} \ldots \underset{k}{\xi^{i_k}}, \quad B = \underset{1}{\xi^{j_1}} \ldots \underset{m}{\xi^{j_m}}.$$

D'après la définition, on a:

$$A \wedge B = \operatorname{Alt}(A \times B) = \operatorname{Alt} \underset{1}{\xi^{i_1}} \ldots \underset{k}{\xi^{i_k}} \underset{k+1}{\xi^{j_1}} \ldots \underset{k+m}{\xi^{j_m}}.$$

En utilisant *7.14* (3) on obtient

$$A \wedge B = \frac{1}{(k+m)!} \begin{vmatrix} \underset{1}{\xi^{i_1}} \ldots \underset{k}{\xi^{i_1}} \underset{k+1}{\xi^{i_1}} \ldots \underset{k+m}{\xi^{i_1}} \\ \cdots\cdots\cdots\cdots \\ \underset{1}{\xi^{i_k}} \ldots \underset{k}{\xi^{i_k}} \underset{k+1}{\xi^{i_k}} \ldots \underset{k+m}{\xi^{i_k}} \\ \cdots\cdots\cdots\cdots \\ \underset{1}{\xi^{j_m}} \ldots \underset{k}{\xi^{j_m}} \underset{k+1}{\xi^{j_m}} \ldots \underset{k+m}{\xi^{j_m}} \end{vmatrix} =$$

$$= \frac{1}{(k+m)!} \sum_{(\nu)} \varepsilon^{i_1 \ldots i_k j_1 \ldots j_m}_{\nu_1 \ldots \nu_k \nu_{k+1} \ldots \nu_{k+m}} \underset{1}{\xi^{\nu_1}} \ldots \underset{k}{\xi^{\nu_k}} \underset{k+1}{\xi^{\nu_{k+1}}} \ldots \underset{k+m}{\xi^{\nu_{k+m}}}. \tag{3}$$

e. *La multiplication tensorielle alternée est une opération linéaire*, avec la multiplication tensorielle et l'alternation: si $A = a_1A_1 + a_2A_2$, alors

$$(a_1A_1 + a_2A_2) \wedge B = \mathrm{Alt}\,\{(a_1A_1 + a_2A_2) \times B\} =$$
$$= \mathrm{Alt}\,(a_1A_1 \times B + a_2A_2 \times B) = a_1A_1 \wedge B + a_2A_2 \wedge B,$$

et de même pour le cas $B = b_1B_1 + b_2B_2$.

f. *Etant données deux formes linéaires* $A = \sum_i a_i\xi^i$, $B = \sum_j b_j\xi^j$, *la loi anticommutative est remplie*:

$$A \wedge B = -B \wedge A. \tag{4}$$

Une fois la propriété linéaire démontrée, il suffit de prouver (4) pour les formes monômes $A = \xi^i$, $B = \xi^j$. On a:

$$\xi^i \wedge \xi^j = \mathrm{Alt}\,\underset{1\,2}{\xi^i\xi^j} = \frac{1}{2}\,(\underset{1\,2}{\xi^i\xi^j} - \underset{1\,2}{\xi^j\xi^i}) = -\mathrm{Alt}\,\underset{1\,2}{\xi^j\xi^i} = -\xi^j \wedge \xi^i.$$

Pour les formes de degrés supérieurs, la relation (4) sera remplacée par une relation plus compliquée (cf. ex. 10).

g. Montrons que *la multiplication tensorielle alternée est associative*, c'est-à-dire que, *pour n'importe quelles trois formes* A, B *et* C, *a lieu la relation*:

$$(A \wedge B) \wedge C = A \wedge (B \wedge C). \tag{5}$$

Il suffit de le montrer pour les formes monômes de base

$$A = \underset{1}{\xi^{i_1}} \ldots \underset{k}{\xi^{i_k}},\quad B = \underset{1}{\xi^{j_1}} \ldots \underset{l}{\xi^{j_l}},\quad C = \underset{1}{\xi^{s_1}} \ldots \underset{m}{\xi^{s_m}},$$

puisque le passage au cas général se réalise aisément à l'aide de la propriété linéaire déjà démontrée.

La forme $A \wedge B$ est trouvée plus haut (3). Formons le produit $(A \wedge B) \wedge C$. Nous avons

$$(A \wedge B) \wedge C = \mathrm{Alt}\,[(A \wedge B) \times C] =$$
$$= \mathrm{Alt}\left(\frac{1}{(k+l)!} \sum_{(\nu)} \varepsilon^{i_1 \ldots i_k j_1 \ldots j_l}_{\nu_1 \ldots \ldots \nu_{k+l}} \underset{k+l}{\xi^{\nu_1}} \ldots \underset{k+l}{\xi^{\nu_{k+l}}}\ \underset{k+l+1}{\xi^{s_1}} \ldots \underset{k+l+m}{\xi^{s_m}}\right) =$$
$$= \frac{1}{(k+l)!} \sum_{(\nu)} \varepsilon^{i_1 \ldots i_k j_1 \ldots j_l}_{\nu_1 \ldots \ldots \nu_{k+l}} \mathrm{Alt}\,\underset{1}{\xi^{\nu_1}} \ldots \underset{k+l}{\xi^{\nu_{k+l}}}\ \underset{k+l+1}{\xi^{s_1}} \ldots \underset{k+l+m}{\xi^{s_m}}.$$

En appliquant encore une fois *7.14* (3) on obtient

$$(A \wedge B) \wedge C =$$

$$= \frac{1}{(k+l)!} \sum_{(\nu)} \varepsilon^{i_1 \dots i_k j_1 \dots j_l}_{\nu_1 \dots \dots \nu_{k+l}} \frac{1}{(k+l+m)!} \begin{vmatrix} \xi^{\nu_1}_1 & \dots & \xi^{\nu_1}_{k+l+m} \\ \dots & \dots & \dots \\ \xi^{\nu_{k+l}}_1 & \dots & \xi^{\nu_{k+l}}_{k+l+m} \\ \xi^{s_1}_1 & \dots & \xi^{s_1}_{k+l+m} \\ \xi^{s_m}_1 & \dots & \xi^{s_m}_{k+l+m} \end{vmatrix}.$$

Permutons les lignes du déterminant obtenu de façon qu'elles soient dans l'ordre initial $i_1, \dots, i_k, j_1, \dots, j_l$; cela fournit

$$(A \wedge B) \wedge C = \frac{1}{(k+l)!} \frac{1}{(k+l+m)!} \sum_{(\nu)} \begin{vmatrix} \xi^{i_1}_1 & \dots & \xi^{i_1}_{k+l+m} \\ \dots & \dots & \dots \\ \xi^{j_l}_1 & \dots & \xi^{j_l}_{k+l+m} \\ \xi^{s_1}_1 & \dots & \xi^{s_1}_{k+l+m} \\ \dots & \dots & \dots \\ \xi^{s_m}_1 & \dots & \xi^{s_m}_{k+l+m} \end{vmatrix}.$$

Le déterminant obtenu ne dépend plus du multinuméro (ν); tous les termes sous le signe de sommation sont devenus égaux. Leur nombre, égal à celui de tous les numéros stricts (ν), vaut $(k+l)!$, définitivement:

$$(A \wedge B) \wedge C = \frac{1}{(k+l+m)!} \begin{vmatrix} \xi^{i_1}_1 & \dots & \xi^{i_1}_{k+l+m} \\ \dots & \dots & \dots \\ \xi^{j_l}_1 & \dots & \xi^{j_l}_{k+l+m} \\ \dots & \dots & \dots \\ \xi^{s_m}_1 & \dots & \xi^{s_m}_{k+l+m} \end{vmatrix}. \qquad (6)$$

Il est clair qu'en calculant $A \wedge (B \wedge C)$, on aboutira au même résultat, et la loi associative de la multiplication tensorielle alternée est établie. Ainsi, l'expression $A \wedge B \wedge C$ et ses analogues, avec

un plus grand nombre de facteurs, sont définies d'une façon univoque.

7.16. Notation canonique des formes antisymétriques.

a. Pour trois monômes $\xi^{i_1}, \xi^{i_2}, \xi^{i_3}$ on a, conformément à *7.14* (3),

$$\xi^{i_1} \wedge \xi^{i_2} \wedge \xi^{i_3} = \frac{1}{3!} \begin{vmatrix} \xi^{i_1}_1 & \xi^{i_1}_2 & \xi^{i_1}_3 \\ \xi^{i_2}_1 & \xi^{i_2}_2 & \xi^{i_2}_3 \\ \xi^{i_3}_1 & \xi^{i_3}_2 & \xi^{i_3}_3 \end{vmatrix} = \operatorname{Alt}_{1\,2\,3} \xi^{i_1} \xi^{i_2} \xi^{i_3}.$$

En continuant de manière analogue nous trouvons pour n'importe quels $i_1, \ldots, i_k$:

$$\xi^{i_1} \wedge \ldots \wedge \xi^{i_k} = \operatorname{Alt}_{1} \xi^{i_1} \ldots \xi^{i_k}_{k}. \tag{1}$$

Si le multinuméro $(i) = (\alpha) = (\alpha_1, \ldots, \alpha_k)$ est strictement ordonné, alors la forme $\operatorname{Alt}_{1} \xi^{i_1} \ldots \xi^{i_k}_{k}$, à un facteur près, coïncide avec la forme antisymétrique de base $D_{(\alpha)}(\underset{1}{x}, \ldots, \underset{k}{x})$ (*7.14b*). Par conséquent, *toute k-forme antisymétrique* $A(\underset{1}{x}, \ldots, \underset{k}{x})$ *s'écrit comme suit* :

$$A(\underset{1}{x}, \ldots, \underset{k}{x}) = \sum_{(\alpha)} \widetilde{a}_{(\alpha)} \xi^{\alpha_1} \wedge \ldots \wedge \xi^{\alpha_k}, \tag{2}$$

avec les coefficients $\widetilde{a}_{(\alpha)}$ *définis d'une façon univoque.* L'égalité (2) s'appelle *première notation canonique de la forme antisymétrique* $A(\underset{1}{x}, \ldots, \underset{k}{x})$.

b. Bien entendu, on peut écrire autrement la forme antisymétrique

$$A(\underset{1}{x}, \ldots, \underset{k}{x}) = \sum_{(i)} a_{(i)} \xi^{i_1} \wedge \ldots \wedge \xi^{i_k} \tag{3}$$

en sommant sur tous les $(k\text{-}n)$-numéros stricts $(i) = (i_1, \ldots, i_k)$, mais les coefficients $a_{(i)}$ de cette notation ne sont en général pas définis d'une façon univoque. Si l'on exige que, dans la notation (3), les coefficients $a_{(i)}$ soient des fonctions antisymétriques du multinuméro (i), i.e. changent de signe en permutant deux composantes de ce numéro, la notation (3) devient univoque, et les coefficients $a_{(i)}$ sont faciles à exprimer par les coefficients $\widetilde{a}_{(\alpha)}$ de la notation (2). En effet, dans la supposition signalée, on a

$$A(\underset{1}{x}, \ldots, \underset{k}{x}) = \sum_{(i)} a_{(i)} \xi^{i_1} \wedge \ldots \wedge \xi^{i_k} = \sum_{(\alpha)} \sum_{O(i)=(\alpha)} a_{(i)} \xi^{i_1} \wedge \ldots \wedge \xi^{i_k};$$

en utilisant l'anticommutativité de l'opération $\wedge$ pour les formes linéaires et l'antisymétrie des coefficients $a_{(i)}$, on trouve

$$A\underset{1}{(x}, \ldots, \underset{k}{x}) = \sum_{(\alpha)} \sum_{O(i)=(\alpha)} a_{(\alpha)} \xi^{\alpha_1} \wedge \ldots \wedge \xi^{\alpha_k} =$$

$$= k! \sum_{(\alpha)} a_{(\alpha)} \xi^{\alpha_1} \wedge \ldots \wedge \xi^{\alpha_k},$$

d'où

$$\widetilde{a}_{(\alpha)} = k!\, a_{(\alpha)}. \tag{4}$$

Réciproquement, ayant (2) on peut passer à la représentation (3) en posant

$$a_{(i)} = \frac{1}{k!} \varepsilon^{\alpha_1 \ldots \alpha_k}_{i_1 \ldots i_k} \widetilde{a}_{(\alpha)}. \tag{5}$$

L'égalité (3) avec les coefficients antisymétriques $a_{(i)}$ s'appelle *deuxième notation canonique de la forme antisymétrique* $A\underset{1}{(x}, \ldots, \underset{k}{x})$.

c. La notation (3) avec $a_{(i)}$ antisymétriques a un avantage devant la notation (2) : *ses coefficients* $a_{(i)}$ (définis de façon univoque par la condition d'antisymétrie dans tout système de coordonnées) *sont des tenseurs k fois covariants.* En effet, dans les nouvelles coordonnées $\xi^{i'} = p^{i'}_{i} \xi^{i}$, on a :

$$\sum_{(i)} a_{(i)} \xi^{i_1} \wedge \ldots \wedge \xi^{i_k} =$$

$$= \sum_{(i)} \sum_{(i')} a_{(i)} p^{i_1}_{i'_1} \ldots p^{i_k}_{i'_k} \xi^{i'_1} \wedge \ldots \wedge \xi^{i'_k} = \sum_{(i')} a_{(i')} \xi^{i'_1} \wedge \ldots \wedge \xi^{i'_k},$$

et l'on peut poser

$$a_{(i')} = \sum_{(i)} a_{(i)} p^{i_1}_{i'_1} \ldots p^{i_k}_{i'_k}.$$

Ici les coefficients $a'_{(i')}$ sont antisymétriques par rapport aux composantes du multinuméro (i'), c'est-à-dire qu'ils sont justement les coefficients cherchés de la forme transformée. Ceci démontre le caractère tensoriel des quantités $a_{(i)}$.

Donc, *si dans tout système de coordonnées sont données les quantités antisymétriques* (par rapport aux composantes du multinuméro (i)) $a_{(i)}$ *et si l'expression*

$$\sum_{(i)} a_{(i)} \xi^{i_1} \wedge \ldots \wedge \xi^{i_k}$$

ne dépend pas du système de coordonnées, alors $a_{(i)}$ *composent un tenseur k fois covariant.* En ce qui concerne les coefficients $\widetilde{a}_{(\alpha)}$ dans la formule (2), la règle de leur transformation en passant aux nouvelles coordonnées est d'un tout autre caractère (cf. ex. 12).

d. Etant donné que, pour $(i) = (i_1, \ldots, i_k)$ fixe, les formes

$$A(\underset{1}{x}, \ldots, \underset{k}{x}) = \xi^{i_1} \wedge \ldots \wedge \xi^{i_k}$$

sont antisymétriques (*7.15f*), alors toute leur combinaison linéaire

$$\omega = \sum_{(i)} c_{(i)} \xi^{i_1} \wedge \ldots \wedge \xi^{i_k} \tag{6}$$

représente une forme antisymétrique. Si, de plus, les coefficients $c_{(i)}$ sont antisymétriques par rapport aux composantes de (k-n)-numéro (i), la forme ω est donnée immédiatement dans sa notation canonique (3). *Et si les coefficients $c_{(i)}$ sont symétriques par rapport à un couple de composantes au moins*, par exemple, si $c_{i_1 i_2 \ldots i_k} = c_{i_2 i_1 \ldots i_k}$, *alors la forme* (6) *est identiquement nulle.* En effet, dans ce cas la somme de chaque couple de termes

$$s = c_{i_1 i_2 \ldots i_k} \xi^{i_1} \wedge \xi^{i_2} \wedge \ldots \wedge \xi^{i_k} + c_{i_2 i_1 \ldots i_k} \xi^{i_2} \wedge \xi^{i_1} \wedge \ldots \wedge \xi^{i_k}$$

aux indices $i_3, \ldots, i_k$ fixes, s'annule:

$$s = c_{i_1 i_2 \ldots i_k} (\xi^{i_1} \wedge \xi^{i_2} + \xi^{i_2} \wedge \xi^{i_1}) \wedge \xi^{i_3} \wedge \ldots \wedge \xi^{i_k} = 0.$$

e. Dans le cas général, on peut passer de la notation (6) à la notation canonique (2) comme suit:

$$\sum_{(i)} c_{(i)} \xi^{i_1} \wedge \ldots \wedge \xi^{i_k} = \sum_{(\alpha)} \sum_{O(i)=(\alpha)} c_{(i)} \xi^{i_1} \wedge \ldots \wedge \xi^{i_k} =$$

$$= \sum_{(\alpha)} \Big(\sum_{O(i)=(\alpha)} \varepsilon^{\alpha_1 \ldots \alpha_k}_{i_1 \ldots i_k} c_{(i)} \Big) \xi^{\alpha_1} \wedge \ldots \wedge \xi^{\alpha_k} = \sum_{(\alpha)} \widetilde{a}_{(\alpha)} \xi^{\alpha_1} \wedge \ldots \wedge \xi^{\alpha_k},$$

où

$$\widetilde{a}_{(\alpha)} = \sum_{O(i)=(\alpha)} \varepsilon^{\alpha_1 \ldots \alpha_k}_{i_1 \ldots i_k} c_{(i)}. \tag{7}$$

On peut ensuite trouver la notation canonique (3) de la forme (6) d'après la formule (5):

$$\sum_{(i)} c_{(i)} \xi^{i_1} \wedge \ldots \wedge \xi^{i_k} = \sum_{(j)} a_{(j)} \xi^{j_1} \wedge \ldots \wedge \xi^{j_k},$$

où $O(i) = O(j) = (\alpha)$ et

$$a_{(j)} = \frac{1}{k!} \varepsilon^{\alpha_1 \ldots \alpha_k}_{j_1 \ldots j_k} \widetilde{a}_{(\alpha)} = \frac{1}{k!} \sum_{O(i)=(\alpha)} \varepsilon^{\alpha_1 \ldots \alpha_k}_{j_1 \ldots j_k} \varepsilon^{i_1 \ldots i_k}_{\alpha_1 \ldots \alpha_k} c_{(i)} =$$

$$= \frac{1}{k!} \sum_{O(i)=O(j)} \varepsilon^{i_1 \ldots i_k}_{j_1 \ldots j_k} c_{(i)}. \tag{8}$$

§ 7.2. Formes différentielles

7.21. Définition des formes différentielles.

a. Soit M_n une variété élémentaire différentiable n-dimensionnelle de classe $p > 1$ (*6.21*). En tout point $x \in M_n$ il y a un

espace tangent $T_n(x)$ i.e. l'espace n-dimensionnel vectoriel formé par les vecteurs tangents aux courbes qui passent par le point x. Une transformation des coordonnées $x^{i'} = x^{i'}(x^1, \ldots, x^n)$ dans l'espace M_n induit sur $T_n(x)$ une transformation linéaire des vecteurs d'après la formule $\xi^{i'} = p_i^{i'}\xi^i$ où $p_i^{i'} = \frac{\partial x^{i'}}{\partial x^i}$. Dans tout espace $T_n(x)$, on peut construire des formes multilinéaires des vecteurs $\xi^i = dx^i$, en particulier des k-formes antisymétriques

$$A(x;\ \underset{1}{dx},\ \ldots,\ \underset{k}{dx}) = \sum_{(i)} a_{(i)}(x)\, \underset{1}{dx^{i_1}} \ldots \underset{k}{dx^{i_k}}. \tag{1}$$

Vu les résultats *7.16a, b*, toute k-forme antisymétrique sur l'espace $T_n(x)$ s'écrit

$$A(x;\ \underset{1}{dx},\ \ldots,\ \underset{k}{dx}) = \sum_{(i)} a_{(i)}(x)\, dx^{i_1} \wedge \ldots \wedge dx^{i_k}, \tag{2}$$

avec les coefficients $a_{(i)}(x)$ antisymétriques par rapport aux composantes $i_1, \ldots, i_k$ du $(k\text{-}n)$-numéro (i), ou bien

$$A(x;\ \underset{1}{dx},\ \ldots,\ \underset{k}{dx}) = \sum_{(\alpha)} \widetilde{a}_{(\alpha)}(x)\, dx^{\alpha_1} \wedge \ldots \wedge dx^{\alpha_k}, \tag{3}$$

où $\alpha_1 < \ldots < \alpha_k$. N'importe quelle k-forme antisymétrique

$$\omega(x;\ \underset{1}{dx},\ \ldots,\ \underset{k}{dx}) = \sum_{(i)} b_{(i)}(x)\, dx^{i_1} \wedge \ldots \wedge dx^{i_k}, \tag{4}$$

dont les coefficients sont des fonctions r fois continûment dérivables, qui est donnée dans chaque système de coordonnées et qui a un sens invariant (ne dépendant pas du système de coordonnées) s'appelle *k-forme différentielle antisymétrique r-lisse sur la variété M_n.*

b. Pour parler de l'invariance de l'expression (4), il faut bien avoir une loi de transformation des coefficients $b_{(i)}(x)$ en passant à un nouveau système de coordonnées. Supposons d'abord que la quantité $b_{(i)}(x)$ soit un tenseur k fois covariant, de sorte que

$$b_{(i')}(x) \equiv b_{i'_1 \ldots i'_k}(x) = p_{i'_1}^{i_1} \ldots p_{i'_k}^{i_k} b_{i_1 \ldots i_k}(x).$$

Alors, en vertu des propriétés fondamentales des k-formes antisymétriques dans R_n (*7.15*), on a

$$\sum_{(i')} b_{(i')}(x)\, dx^{i'_1} \wedge \ldots \wedge dx^{i'_k} =$$

$$= \sum_{(i')} \sum_{(i)} p_{i'_1}^{i_1} \ldots p_{i'_k}^{i_k} b_{(i)}(x) \sum_{(j)} p_{j_1}^{i'_1} \ldots p_{j_k}^{i'_k}\, dx^{j_1} \wedge \ldots \wedge dx^{j} =$$

$$= \sum_{(i)} \sum_{(j)} \delta_{j_1}^{i_1} \ldots \delta_{j_k}^{i_k} b_{(i)}(x)\, dx^{j_1} \wedge \ldots \wedge dx^{j_k} =$$

$$= \sum_{(i)} b_{(i)}(x)\, dx^{i_1} \wedge \ldots \wedge dx^{i_k},$$

et la forme (4) a un sens invariant. Notons que la forme (4) peut bien avoir un sens invariant avec une loi plus compliquée, non tensorielle, de transformation des coefficients $b_{(i)}(x)$, nous en verrons des exemples par la suite. Le nombre r, ordre de dérivabilité des coefficients $b_{(i)}(x)$, est supposé suffisant pour tous les calculs qu'on effectuera plus bas.

c. La déduction qu'on vient de réaliser montre comment sont construites des formes différentielles sur une variété. On peut considérer l'expression

$$\omega = \sum_{(i)} b_{(i)}(x)\, dx^{i_1} \wedge \ldots \wedge dx^{i_k}$$

avec les coefficients antisymétriques suffisamment dérivables $b_{(i)}(x)$ qu'on choisit de manière arbitraire dans un système de coordonnées x^i, puis, dans tout autre système de coordonnées $x^{i'}$, poser

$$\omega = \sum_{(i')} b_{(i')}(x)\, dx^{i'_1} \wedge \ldots \wedge dx^{i'_k}$$

où $b_{(i')}(x)$ s'obtiennent de $b_{(i)}(x)$ d'après la loi de transformation d'un tenseur k fois covariant.

d. Les k-formes différentielles antisymétriques sur une variété M_n ($n \geqslant 1$) constituent un espace vectoriel de dimension infinie.

7.22. Différentiation des formes différentielles.

a. Soit

$$\omega = \sum_{(i)} b_{(i)}(x)\, dx^{i_1} \wedge \ldots \wedge dx^{i_k} \tag{1}$$

une k-forme différentielle antisymétrique sur une variété M_n. Examinons l'expression

$$\partial\omega = \sum_{(i)} \sum_{j} \frac{\partial b_{(i)}(x)}{\partial x^j}\, dx^j \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_k}. \tag{2}$$

Bien que les coefficients $\frac{\partial b_i(x)}{\partial x^j}$ n'ont en général plus de caractère tensoriel (*6.26*), l'expression, comme nous allons voir, a un sens invariant de même que la forme ω, donc représente encore une forme différentielle antisymétrique, déjà de $k+1$ variables. Cette $(k+1)$-forme s'appelle *différentielle de la forme* ω.

b. Pour démontrer l'invariance de l'expression $\partial\omega$, supposons d'abord que $b_{(i)}(x)$ soit un tenseur k fois covariant. Alors, d'après *6.26* (1), on a

$$\frac{\partial b_{(i')}(x)}{\partial x^{j'}} = p^{j}_{j'} p^{i_1}_{i'_1} \ldots p^{i_k}_{i'_k} \frac{\partial b_{(i)}(x)}{\partial x^j} +$$

$$+ p^{i_1}_{j'i'_1} p^{i_2}_{i'_2} \ldots p^{i_k}_{i'_k} b_{(i)}(x) + \ldots + p^{i_1}_{i'_1} \ldots p^{i_{k-1}}_{i'_{k-1}} p^{i_k}_{j'i'_k} b_{(i)}(x),$$

où $p^{i_s}_{j'i'_s} = \dfrac{\partial^2 x^{i_s}}{\partial x^{i'_s}\,\partial x^{j'}}$. Il en résulte

$$\frac{\partial b_{(i')}(x)}{\partial x^{j'}}\,dx^{j'} \wedge dx^{i'_1} \wedge \ldots \wedge dx^{i'_k} =$$

$$= \frac{\partial b_{(i)}(x)}{\partial x^{j}}\,dx^{j} \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_k} +$$

$$+ p^{i_1}_{j'i'_1} p^{i_2}_{i'_2} \ldots p^{i_k}_{i'_k} b_{(i)}(x)\, p^{j'}_{j} p^{i'_1}_{j_1} \ldots p^{i'_k}_{j_k}\, dx^{j} \wedge dx^{j_1} \wedge \ldots$$

$$\ldots \wedge dx^{j_k} + \ldots + p^{i_1}_{i'_1} \ldots p^{i_k}_{i'_{k-1} j'i'_k} b_{(i)}(x)\, p^{j'}_{j} p^{i'_1}_{j_1} \ldots$$

$$\ldots p^{i'_k}_{j_k}\, dx^{j} \wedge dx^{j_1} \wedge \ldots \wedge dx^{j_k} =$$

$$= \frac{\partial b_{(i)}(x)}{\partial x^{j}}\,dx^{j} \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_k} +$$

$$+ p^{i_1}_{j'i'_1} p^{j'}_{j} p^{i'_1}_{j_1} b_i(x)\, dx^{j} \wedge dx^{j_1} \wedge dx^{i_2} \wedge \ldots \wedge dx^{i_k} + \ldots$$

$$\ldots + p^{i_k}_{j'i'_k} p^{j'}_{j} p^{i'_k}_{j_k} b_i(x)\, dx^{j} \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_{k-1}} \ldots \wedge dx^{j_k}. \quad (3)$$

Notons ensuite que les coefficients $p^{i_1}_{j'i'_1} p^{j'}_{j} p^{i'_1}_{j_1} b_{(i)}(x)$ sont symétriques par rapport à j et j_1 ; en effet, en les permutant, puis remplaçant j' par i'_1 et i'_1 par j', on a

$$p^{i_1}_{j'i'_1} p^{j'}_{j_1} p^{i'_1}_{j} = p^{i_1}_{i'_1 j'} p^{i'_1}_{j} p^{j'}_{j} = p^{i_1}_{j'i'_1} p^{i'_1}_{j} p^{j'}_{j_1}.$$

Par conséquent, d'après *7.16d*, le deuxième terme de la somme (3) s'annule. Il en est de même de tous les termes suivants de cette somme, et l'on a :

$$\sum_{(i')} \sum_{j'} \frac{\partial b_{(i')}(x)}{\partial x^{j'}}\,dx^{j'} \wedge dx^{i'_1} \wedge \ldots \wedge dx^{i'_k} =$$

$$= \sum_{(i)} \sum_{j} \frac{\partial b_{(i)}(x)}{\partial x^{j}}\,dx^{j} \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_k},$$

ce qui démontre le sens invariant de l'expression (2).

c. Soient maintenant $b_{(i)}(x)$ des coefficients arbitraires, donnés pour tout système de coordonnées, de façon que la forme (1) ait un sens invariant. Avec *7.16e*, écrivons la forme (1) comme suit :

$$\omega_1 = \sum_{(j)} a_{(j)}(x)\, dx^{j_1} \wedge \ldots \wedge dx^{j_k},$$

où les coefficients

$$a_{(j)}(x) = \frac{1}{k!} \sum_{O(i)=O(j)} \varepsilon^{i_1 \ldots i_k}_{j_1 \ldots j_k} b_{(i)}(x),$$

qui sont les composantes d'un tenseur k fois covariant (*7.16c*) sont déjà antisymétriques. D'après ce qu'on a démontré, la forme

$$\partial\omega_1 = \sum_{(j)} \sum_{(s)} \frac{\partial a_{(j)}(x)}{\partial x^s}\, dx^s \wedge dx^{j_1} \wedge \ldots \wedge dx^{j_k}$$

a un sens invariant. Or,

$$\frac{\partial b_{(j)}(x)}{\partial x^s} = \frac{1}{k!} \sum_{O(i)=O(j)} \varepsilon^{i_1 \ldots i_k}_{j_1 \ldots j_k} \frac{\partial b_{(i)}(x)}{\partial x^s},$$

et c'est la forme suivante qui a un sens invariant:

$$\partial\omega_1 = \sum_{(j)} \sum_{s} \frac{1}{k!} \Big(\sum_{O(i)=O(j)} \varepsilon^{i_1 \ldots i_k}_{j_1 \ldots j_k} \frac{\partial b_{(i)}(x)}{\partial x^s} \Big)\, dx^s \wedge dx^{j_1} \wedge \ldots \wedge dx^{j_k} =$$

$$= \sum_{(j)} \Big[\sum_{s} \frac{1}{k!} \sum_{O(i)=O(j)} \frac{\partial b_{(i)}(x)}{\partial x^s}\, dx^s \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_k} \Big] =$$

$$= \sum_{(\alpha)} \sum_{O(j)=(\alpha)} \Big[\sum_{s} \frac{1}{k!} \sum_{O(i)=O(j)} \frac{\partial b_{(i)}(x)}{\partial x^s}\, dx^s \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_k} \Big].$$

Le crochet ne dépendant pas du choix de (j), avec $O(j) = (\alpha)$, la forme $\partial\omega_1$ peut être notée

$$\sum_{(\alpha)} \sum_{s} \sum_{O(i)=(\alpha)} \frac{\partial b_{(i)}(x)}{\partial x^s}\, dx^s \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_k} =$$

$$= \sum_{(i)} \sum_{s} \frac{\partial b_{(i)}(x)}{\partial x^s}\, dx^s \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_k} = d\omega,$$

ce qu'il fallait démontrer.

d. Il est aisé de voir que *l'opération de différentiation est linéaire*: quels que soient les k-formes ω_1 et ω_2 et les nombres a_1 et a_2,

$$\partial\,(a_1\omega_1 + a_2\omega_2) = a_1\,\partial\omega_1 + a_2\,\partial\omega_2.$$

Ceci résulte immédiatement de la linéarité de la dérivation $\frac{\partial}{\partial x^i}$ dans l'ensemble des fonctions.

e. La *différentiation du produit tensoriel antisymétrique* est effectuée d'après la formule:

$$\partial\,(\omega_1 \wedge \omega_2) = \partial\omega_1 \wedge \omega_2 + (-1)^k\, \omega_1 \wedge \partial\omega_2, \qquad (4)$$

où k est le degré de la forme ω_1. En effet, soit

$$\omega_1 = \sum_{(\alpha)} f_{(\alpha)}(x)\, dx^{\alpha_1} \wedge \ldots \wedge dx^{\alpha_k},$$

$$\omega_2 = \sum_{(\beta)} g_{(\beta)}(x)\, dx^{\beta_1} \wedge \ldots \wedge dx^{\beta_m},$$

$$\omega_1 \wedge \omega_2 = \sum_{(\alpha)} \sum_{(\beta)} f_{(\alpha)}(x)\, g_{(\beta)}(x)\, dx^{\alpha_1} \wedge \ldots \wedge dx^{\alpha_k} \wedge dx^{\beta_1} \wedge \ldots \wedge dx^{\beta_m}.$$

Alors, par définition, on a

$$\partial(\omega_1 \wedge \omega_2) = \sum_{(\alpha)}\sum_{(\beta)}\sum_{i} \frac{\partial(f_\alpha g_\beta)}{\partial x^i} dx^i \wedge dx^{\alpha_1} \wedge \ldots \wedge dx^{\beta_m} =$$

$$= \sum_{(\alpha)}\sum_{(\beta)}\sum_{i} \frac{\partial f_\alpha}{\partial x^i} g_\beta\, dx^i \wedge dx^{\alpha_1} \wedge \ldots \wedge dx^{\beta_m} +$$

$$+ \sum_{(\alpha)}\sum_{(\beta)}\sum_{i} f_\alpha \frac{\partial g_\beta}{\partial x^i} dx^i \wedge dx^{\alpha_1} \wedge \ldots \wedge dx^{\beta_m} =$$

$$= \partial\omega_1 \wedge \omega_2 + \omega_1 \wedge (-1)^k \partial\omega_2,$$

puisque, pour obtenir $\partial\omega_2$ il faut permuter successivement dx^i avec $dx^{\alpha_1}, \ldots, dx^{\alpha_k}$.

f. Il nous sera utile une formule qui représente un cas spécial de la formule (4):

$$\partial(\omega \wedge dx^m) = \partial\omega \wedge dx^m. \tag{5}$$

Ici le deuxième terme de la formule (4) est absent en vertu de l'égalité

$$\partial(dx^m) = \partial(1 \cdot dx^m) = \sum_i \frac{\partial 1}{\partial x^i} dx^i \wedge dx^m = 0.$$

7.23. Exemples.

a. Pour une 0-forme $\omega = f(x)$ nous avons

$$\partial\omega = \sum_i \frac{\partial f(x)}{\partial x^i} dx^i.$$

L'expression obtenue est invariante car $\frac{\partial f(x)}{\partial x^i}$ est un tenseur une fois covariant (*6.24b*).

b. Pour une 1-forme $\omega = \sum_i f_i(x)\, dx^i$ l'invariance signifie que $f_i(x)$ est un tenseur une fois covariant. Nous avons

$$\partial\omega = \sum_i \sum_j \frac{\partial f_i(x)}{\partial x^j} dx^j \wedge dx^i.$$

Construisons la première notation canonique *7.16* (2) de la forme $\partial\omega$. Soit $(\alpha) = (\alpha_1, \alpha_2)$ où $\alpha_1 < \alpha_2$. Alors, d'après *7.16e*, on a

$$\tilde{a}_{\alpha_1\alpha_2} = \sum_{O(i)=(\alpha)} \varepsilon^{i_1 i_2}_{\alpha_1\alpha_2} c_{i_1 i_2} = c_{\alpha_1\alpha_2} - c_{\alpha_2\alpha_1}.$$

Dans le présent cas, $c_{ji} = \frac{\partial f_i(x)}{\partial x^j}$, donc

$$\tilde{a}_{\alpha_1\alpha_2} = \frac{\partial f_{\alpha_2}(x)}{\partial x^{\alpha_1}} - \frac{\partial f_{\alpha_1}(x)}{\partial x^{\alpha_2}}$$

et, par conséquent,

$$\partial\omega = \sum_{\alpha_1<\alpha_2} \left(\frac{\partial f_{\alpha_2}(x)}{\partial x^{\alpha_1}} - \frac{\partial f_{\alpha_1}(x)}{\partial x^{\alpha_2}} \right) dx^{\alpha_1} \wedge dx^{\alpha_2}.$$

c. Pour une $(n-1)$-forme ω en notation *7.16* (2)

$$\omega = a_1(x)\, dx^2 \wedge \ldots \wedge dx^n - a_2(x)\, dx^1 \wedge dx^3 \wedge \ldots$$
$$\ldots \wedge dx^n + \ldots + (-1)^{n-1} a_n(x)\, dx^1 \wedge \ldots \wedge dx^{n-1},$$

nous avons

$$\partial\omega = \frac{\partial a_1}{\partial x^1} dx^1 \wedge dx^2 \wedge \ldots \wedge dx^n - \frac{\partial a_2}{\partial x^2} dx^2 \wedge dx^1 \wedge dx^3 \wedge \ldots$$
$$\ldots \wedge dx^n + \ldots + (-1)^n \frac{\partial a_n}{\partial x^n} dx^n \wedge dx^1 \wedge \ldots \wedge dx^{n-1};$$

l'absence des autres termes est expliquée par le fait que $dx^i \wedge dx^i = 0$. En mettant le premier facteur dx^i à la i-ème place, on trouve

$$\partial\omega = \left(\frac{\partial a_1}{\partial x^1} + \ldots + \frac{\partial a_n}{\partial x^n} \right) dx^1 \wedge \ldots \wedge dx^n.$$

d. Pour une n-forme ω, la forme $\partial\omega$ est nulle en tant que $(n+1)$-forme antisymétrique.

7.24. Théorème de Poincaré.

a. Lemme. *Pour toute forme différentielle antisymétrique ω sur une variété M_n, on a*

$$\partial^2\omega \equiv \partial(\partial\omega) \equiv 0.$$

Démonstration. Soit

$$\omega = \sum_{(\alpha)} f_{(\alpha)}(x)\, dx^{\alpha_1} \wedge \ldots \wedge dx^{\alpha_k}.$$

Alors, d'après la définition *7.22a*, on a

$$\partial\omega = \sum_{(\alpha)} \sum_i \frac{\partial f_{(\alpha)}(x)}{\partial x^i} dx^i \wedge dx^{\alpha_1} \wedge \ldots \wedge dx^{\alpha_k},$$

$$\partial^2\omega \equiv \partial(\partial\omega) = \sum_{(\alpha)} \sum_i \sum_j \frac{\partial^2 f_{(\alpha)}(x)}{\partial x^i\, \partial x^j} dx^j \wedge dx^i \wedge dx^{\alpha_1} \wedge \ldots \wedge dx^{\alpha_k}.$$

Les coefficients de la forme obtenue sont symétriques par rapport aux indices i et j; par conséquent, en vertu de *7.16d*, la forme $\partial^2\omega$ est nulle, ce qu'il fallait.

Le théorème de Poincaré que nous démontrerons dans *c* représente une inversion (locale) de ce lemme: il se trouve que si la forme ω est telle que $\partial\omega \equiv 0$, alors, au voisinage de n'importe quel point $x \in M_n$, il existe une forme θ telle que $\partial\theta = \omega$. Envisageons préalablement encore un lemme.

b. Lemme. *Si une k-forme différentielle antisymétrique* $\omega = \sum_{(\alpha)} \widetilde{a}_{(\alpha)}\, dx^{\alpha_1} \wedge \ldots \wedge dx^{\alpha_k}$ *sur une variété* M_n *est telle que* $\partial\omega = 0$, *et si* ω *ne contient pas* dx^n, *alors ses coefficients* $\widetilde{a}_{(\alpha)}$ *ne contiennent pas* x^n.

Démonstration. On a

$$\partial\omega = \sum_{(\alpha)} \sum_{j} \frac{\partial \widetilde{a}_{(\alpha)}(x)}{\partial x^j}\, dx^j \wedge dx^{\alpha_1} \wedge \ldots \wedge dx^{\alpha_k} = \sum_{(\alpha)} \frac{\partial \widetilde{a}_{(\alpha)}(x)}{\partial x^n}\, dx^n \wedge dx^{\alpha_1} \wedge \ldots \wedge dx^{\alpha_k} + \omega_1,$$

où ω_1 désigne une forme ne contenant pas dx^n. Comme $\partial\omega = 0$ par hypothèse, chaque coefficient de la forme obtenue en réduisant les termes semblables est nul. Les termes explicités dans le second membre n'ont pas de semblables; donc, $\frac{\partial \widetilde{a}_{(\alpha)}(x)}{\partial x^n} \equiv 0$, c.q.f.d.

c. Théorème (de Poincaré). *Si une k-forme antisymétrique* ω $(k \geqslant 1)$ *satisfait, dans un domaine d'une variété* M_n, *à l'équation* $\partial\omega = 0$, *alors, au voisinage de chaque point de ce domaine, il existe une* $(k-1)$*-forme* θ *telle que* $\partial\theta = \omega$.

Démonstration. Pour $n = 1$ et $k > 1$, le théorème est trivial; et pour $k = 1$, il se réduit à l'affirmation: toute 1-forme $f(x)\, dx$ est la différentielle d'une 0-forme $\theta(x)$. Cette affirmation est évidente: la 0-forme cherchée peut être donnée par l'expression $\theta(x) = \int_a^x f(\xi)\, d\xi$. Supposons maintenant que le théorème soit juste pour toutes les k-formes dans n'importe quelle variété $(n-1)$-dimensionnelle M_{n-1}, et démontrons-le pour une k-forme $\omega = \sum_{(\alpha)} f_{(\alpha)}(x)\, dx^{\alpha_1} \wedge \ldots \wedge dx^{\alpha_k}$ dans la variété M_n.

Explicitons dans la forme ω les termes contenant dx^n:

$$\omega = dx^n \wedge \omega_1 + \omega_2; \tag{1}$$

ici ω_1 est une $(k-1)$-forme et ω_2 une k-forme, les deux ne contenant pas dx^n. Le reste de la construction sera réalisé au voisinage d'un point donné $x_0 = (x_0^1, \ldots, x_0^n)$. Sans restreindre la généralité, on peut poser $x_0^1 = \ldots = x_0^n = 0$. Désignons par ∂_0 la différentiation par rapport aux variables $x^1, \ldots, x^{n-1}$. La restriction de la forme ω sur la surface $x^n = 0$ représente la k-forme ω_2 où l'on a posé $x^n = 0$; désignons cette nouvelle forme par ω_0. Montrons que $\partial_0(\omega_0) = 0$; en effet,

$$0 = \partial\omega = \partial\omega|_{x^n=0} = dx^n \wedge \partial\omega_1|_{x^n=0} + \partial_0\omega_2|_{x^n=0} + \frac{\partial\omega_2}{\partial x^n} \wedge dx^n|_{x^n=0} = \partial_0\omega_0. \tag{2}$$

Notons que, pour toute forme γ, l'opérateur de la différentiation ∂_0 commute avec la substitution $x^n = 0$, de sorte que $(\partial_0\gamma)_0 = \partial_0(\gamma_0)$. La surface $x^n = 0$ représente elle-même une variété $(n-1)$-dimensionnelle différentiable. Par l'hypothèse de récurrence, on trouve que, dans un voisinage du point 0, sur la surface $x^n = 0$, il existe une $(k-2)$-forme θ_0 (des variables $x^1, \ldots, x^{n-1}$) pour laquelle

$$\partial_0\theta_0 = \omega_0. \tag{3}$$

Cherchons la forme θ au voisinage du point 0 parmi les $(k-1)$-formes qui ne dépendent pas de dx^n. Pour chaque forme, on peut écrire

$$\partial\theta = \partial x^n \wedge \frac{\partial\theta}{\partial x^n} + \partial_0\theta, \tag{4}$$

où $\partial_0\theta$ représente une forme qui ne contient pas non plus dx^n, et le symbole $\frac{\partial\theta}{\partial x^n}$ signifie qu'on a dérivé chaque coefficient de la forme θ par rapport à x^n. Comme il s'agit de résoudre l'équation $\partial\theta = \omega$, la comparaison de (1) et (4) montre qu'il serait bon de résoudre préalablement l'équation

$$\frac{\partial\theta}{\partial x^n} = \omega_1. \tag{5}$$

L'équation (5) peut être traitée comme système d'équations différentielles ordinaires, pour lesquelles les quantités $x^1, \ldots, x^{n-1}$ sont des paramètres, et le nombre de fonctions inconnues est égal à celui des coefficients dans la forme θ (et dans ω_1). Pour obtenir une solution unique, il faut compléter (5) par une condition initiale. En tant que telle choisissons

$$\theta \mid_{x^n=0} = \theta_0, \tag{6}$$

où θ_0 est une solution de l'équation (3).

D'après le théorème fondamental d'existence de la solution pour une équation avec des paramètres (*1.65*), l'unique solution du système (5) avec la condition (6) existe dans un voisinage du point 0 et est dérivable par rapport aux paramètres $x^1, \ldots, x^{n-1}$.

Montrons que la forme θ qu'on obtient ainsi satisfait déjà à l'égalité $\partial\theta = \omega$. En effet, considérons la forme $\varphi = \partial\theta - \omega$. Son coefficient devant dx^n vaut $\frac{\partial\theta}{\partial x^n} - \omega_1 \equiv 0$. Avec la condition $\partial\omega = 0$ trouvons que $\partial\varphi = \partial^2\theta - \partial\omega = 0$. Par conséquent, en vertu du lemme *b*, les coefficients de la forme φ ne contiennent pas x^n. Posons $x^n = 0$, alors, en vertu de la condition initiale (6), on a :

$$\varphi = (\partial\theta - \omega)\mid_{x^n=0} = dx^n \wedge \frac{\partial\theta}{\partial x^n}\Big|_{x^n=0} + \\ + \partial_0\theta\mid_{x^n=0} - \omega\mid_{x^n=0} = \partial_0\theta_0 - \omega_0 = 0.$$

Donc, $\varphi \equiv 0$ au voisinage du point 0, et la forme que nous avons construite satisfait à l'équation $\partial\theta = \omega$, ce qu'il fallait.

Nous avons trouvé la forme θ dans un système donné de coordonnées. On peut la définir dans tout autre système de coordonnées d'après la règle décrite dans *7.21c*. Il faut prouver que l'égalité $\partial\theta = \omega$ a lieu dans chaque système de coordonnées. Or, pour une forme différentielle invariante qu'est devenue la forme θ, sa différentielle a elle aussi un sens invariant d'après *7.22b*. Par conséquent, l'égalité $\partial\theta = \omega$ obtenue dans le système de coordonnées initial, a lieu dans tout autre système de coordonnées, ce qui achève la démonstration.

7.25. Application à l'analyse vectorielle.

a. Rappelons-nous les opérations différentielles principales de l'analyse vectorielle classique dans l'espace tridimensionnel R_3. Dans les coordonnées orthonormales x^1, x^2, x^3, à un champ scalaire $\varphi(x)$ on fait correspondre le champ vectoriel du gradient

$$\operatorname{grad}\varphi(x) = \left(\frac{\partial\varphi}{\partial x^1}, \frac{\partial\varphi}{\partial x^2}, \frac{\partial\varphi}{\partial x^3}\right).$$

A un champ vectoriel $W = (W_1, W_2, W_3)$ on fait correspondre le champ vectoriel du rotationnel (*4.24d*)

$$\operatorname{rot} W = \left(\frac{\partial W_2}{\partial x^3} - \frac{\partial W_3}{\partial x^2}, \frac{\partial W_3}{\partial x^1} - \frac{\partial W_1}{\partial x^3}, \frac{\partial W_1}{\partial x^2} - \frac{\partial W_2}{\partial x^1}\right)$$

et le champ scalaire de la divergence (*4.15*)

$$\operatorname{div} W = \frac{\partial W_1}{\partial x^1} + \frac{\partial W_2}{\partial x^2} + \frac{\partial W_3}{\partial x^3}.$$

Toutes ces définitions peuvent être considérées comme cas spéciaux de la définition de la différentielle d'une forme différentielle, et cela même dans l'espace n-dimensionnel. Ainsi, si l'on interprète le champ scalaire $\varphi(x)$ comme 0-forme $\varphi(x)$, alors sa différentielle $\partial\varphi(x)$ est la 1-forme $\sum\limits_i \frac{\partial\varphi}{\partial x^i}\, dx^i$ (*7.23a*), qu'on peut faire correspondre au champ vectoriel $\operatorname{grad}\varphi(x) = \left(\frac{\partial\varphi}{\partial x^1}, \frac{\partial\varphi}{\partial x^2}, \frac{\partial\varphi}{\partial x^3}\right)$. Si l'on fait correspondre le champ vectoriel $W = (W_1, \ldots, W_n)$ à la 1-forme $\omega = \sum\limits_i W_i(x)\, dx^i$, alors, d'après *7.23b*, on a

$$\partial\omega = \sum_{\alpha<\beta} \left(\frac{\partial W_\beta(x)}{\partial x^\alpha} - \frac{\partial W_\alpha(x)}{\partial x^\beta}\right) dx^\alpha \wedge dx^\beta;$$

cette forme, pour $n = 3$, a trois composantes essentielles qui coïncident avec celles du vecteur rot W. Et si l'on fait correspondre le

même champ vectoriel $W = (W_1, \ldots, W_n)$ à la $(n-1)$-forme $\zeta = \sum_j W_j(x)(-1)^{j-1} dx^1 \wedge \ldots \wedge \widehat{dx^j} \wedge \ldots \wedge dx^n$ (où le signe $\widehat{\ }$ montre que le facteur correspondant fait défaut), on aura d'après *7.23c*:

$$\partial\zeta = \sum_j \frac{\partial W_j(x)}{\partial x^j} dx^1 \wedge \ldots \wedge dx^n,$$

ce qui correspond au champ scalaire div W.

L'égalité $\partial^2\varphi = 0$ pour le champ scalaire φ peut maintenant être notée comme rot grad $\varphi = 0$. L'égalité $\partial^2\omega = 0$, pour le champ vectoriel $W(x)$ dans l'espace R_3 peut être notée comme div rot $W(x) = 0$.

b. Un champ vectoriel $f = \{f_1(x), \ldots, f_n(x)\}$ qui est le gradient d'un champ scalaire $\theta(x)$ s'appelle *champ potentiel* (*4.21b*), et la fonction $\theta(x)$ *potentiel du champ* f. Le théorème de Poincaré fournit:

Conséquence. *Si, dans un domaine G de l'espace R_n, une 1-forme $f = \sum_i f_i(x)\, dx^i$ satisfait à la condition*

$$\partial f = \sum_i \sum_j \frac{\partial f_i}{\partial x^j} dx^j \wedge dx^i = 0$$

ou, ce qui est le même (*7.23b*),

$$\frac{\partial f_i}{\partial x^j} - \frac{\partial f_j}{\partial x^i} = 0 \quad (i, j = 1, \ldots, n),$$

alors chaque point $x_0 \in G$ admet un voisinage dans lequel le champ $f(x)$ est potentiel, autrement dit, dans lequel il existe un champ scalaire $\theta(x)$ tel que $f = \operatorname{grad}\theta$, ou, ce qui revient au même,

$$f_i(x) = \frac{\partial\theta(x)}{\partial x^i}.$$

Nous avons déjà vu ce résultat dans *4.21d* où il était conséquence du théorème de Frobenius.

c. Envisageons une $(n-1)$-forme

$$\omega = \sum_j (-1)^{j-1} f_j(x)\, dx^1 \wedge \ldots \wedge \widehat{dx^j} \wedge \ldots \wedge dx^n$$

ayant la divergence

$$\operatorname{div}\omega \equiv \partial\omega = \sum_{j=1}^{n} \frac{\partial f_j}{\partial x^j} dx^1 \wedge \ldots \wedge dx^n.$$

S'il existe une $(n-2)$-forme θ telle que $\partial\theta = \omega$, alors div $\omega = 0$, la réciproque étant elle aussi valable (au moins localement) d'après le théorème de Poincaré. Pour $n = 3$, la $(n-2)$-forme θ représente une 1-forme $\sum_{j=1}^{3} \theta_i(x)\, dx^i$ ou bien un champ vectoriel $(\theta_1, \theta_2, \theta_3)$, $\partial\theta$ étant associé au champ vectoriel rot θ. Un champ vectoriel $W = (W_1, W_2, W_3)$ qui est le rotationnel d'un champ vectoriel $\theta = (\theta_1, \theta_2, \theta_3)$ s'appelle *vectoriellement potentiel*, et le champ θ *potentiel vectoriel du champ* W (*4.43b*). Nous tirons du théorème de Poincaré :

Conséquence. *Si, dans un domaine* G *de l'espace* R_3, *un champ vectoriel* $W = (W_1, W_2, W_3)$ *satisfait à la condition*

$$\operatorname{div} W = \frac{\partial W_1}{\partial x^1} + \frac{\partial W_2}{\partial x^2} + \frac{\partial W_3}{\partial x^3} = 0,$$

alors chaque point $M \in G$ *admet un voisinage dans lequel le champ* W *est vectoriellement potentiel, autrement dit, dans lequel il existe un champ vectoriel* θ *tel que* $W = \operatorname{rot}\theta$, *ou bien, ce qui revient au même,*

$$W_1 = \frac{\partial\theta_3}{\partial x^2} - \frac{\partial\theta_2}{\partial x^3}, \quad W_2 = \frac{\partial\theta_1}{\partial x^3} - \frac{\partial\theta_3}{\partial x^1}, \quad W_3 = \frac{\partial\theta_2}{\partial x^1} - \frac{\partial\theta^1}{\partial x_2}.$$

Ce résultat nous est également familier (*4.66*). Proprement dit, c'est l'idée de la construction de *4.66*, dûment généralisée, qui amène à la démonstration du théorème de Poincaré.

§ 7. 3. Théorèmes intégraux

7.31. Intégration des formes différentielles.

a. Désignons par I^k le cube unité dans l'espace euclidien k-dimensionnel R_k qui est déterminé en coordonnées orthonormales $u^1, \ldots, u^k$, par les inégalités

$$0 \leqslant u_1 \leqslant 1, \ \ldots, \ 0 \leqslant u_k \leqslant 1.$$

Pour une fonction $f(u)$ continue dans le cube I^k, il existe l'intégrale k-uple de Riemann (*3.15b*)

$$\int_{I^k} f(u)\, du = \lim_{d(\Pi)\to 0} \sum_{j=1}^{N} f(\xi_j)\, |\Delta^j u|$$

où, comme d'habitude, $\Pi = \{\Delta^j u\}$ est une partition du cube I^k en pavés $\Delta^i u$, $j = 1, \ldots, N$; $|\Delta^j u|$ le volume du pavé $\Delta^j u$; $d(\Pi)$ le plus grand des diamètres des pavés $\Delta^j u$ dans la partition Π ; ξ_j un point du pavé $\Delta^j u$. On peut toujours admettre que le pavé $\Delta^j u$ soit déterminé par les vecteurs $h_1, \ldots, h_k$ ayant ξ_j pour origine commune. La quantité $\sum f(\xi_j)\, |\Delta^j u|$ s'avère donc égale à la valeur

de la forme linéaire antisymétrique $\omega(u; du^1, \ldots, du^k) = f(u)\, du^1 \wedge \ldots \wedge du^k$ pour $u = \xi_j$, sur les vecteurs $du^1 = h_1, \ldots, du^k = h_k$ (*7.13c*). Nous sommes amenés à une définition naturelle de l'intégrale de la forme k-linéaire antisymétrique ω :

$$\int_{I^k} \omega \equiv \int_{I^k} f(u)\, du^1 \wedge \ldots \wedge du^k = \int_{I^k} f(u)\, du. \tag{1}$$

Il est évident que l'intégrale obtenue de la forme ω possède les propriétés linéaires usuelles de l'intégrale d'une fonction ordinaire : l'intégrale de la somme de deux formes est égale à la somme des intégrales des termes, et un facteur numérique peut être mis en dehors du signe somme.

b. Dans la suite (*e*) nous définirons l'intégrale d'une forme k-linéaire antisymétrique

$$\omega = \sum_{(i)} f_{(i)}(x)\, dx^{i_1} \wedge \ldots \wedge dx^{i_k} \tag{2}$$

sur certaines « surfaces admissibles » dans une variété différentiable élémentaire n-dimensionnelle M_n ; nous allons définir ces surfaces.

Un ensemble $S \subset M_n$ sera appelé *k-surface admissible* s'il peut être représenté par les équations de la forme

$$x^1 = \varphi^1(u^1, \ldots, u^k), \ldots, x^n = \varphi^n(u^1, \ldots, u^k), \tag{3}$$

où les fonctions $\varphi^1, \ldots, \varphi^n$ sont définies, continues et possèdent les dérivées premières continues dans le cube I^k. On ne suppose pas que l'application (3) soit biunivoque ou qu'elle soit non dégénérée (i.e. que la matrice jacobienne $\frac{\partial(\varphi^1, \ldots, \varphi^n)}{\partial(u^1, \ldots, u^k)}$ ait partout le rang k). En particulier, la k-surface admissible peut représenter un point ou une surface sans angles ; par exemple, les fonctions

$$x^1 = u^1 \cos 2\pi u^2,\ x^2 = u^1 \sin 2\pi u^2,\ x^3 = \ldots = x^n = 0$$

appliquent le carré I^2 sur un cercle dans l'espace R_n, de sorte que le cercle soit une *2*-surface admissible. Remarquons encore que la définition d'une surface admissible dans M_n a, évidemment, un sens absolu, qui ne dépend pas du choix spécial des coordonnées dans M_n.

c. Soulignons qu'une surface admissible est non seulement un être géométrique dans la variété M_n, mais aussi le système des équations (3) qui la déterminent. Pour le moment, nous sommes forcés de considérer comme différentes une surface S représentée par les équations du type (3) et la même surface (au sens géométrique) S' représentée par d'autres équations du type (3). Ainsi, l'intervalle $x^1 = u$, $x^2 = u$, $0 \leqslant u \leqslant 1$, sur le plan R_2 n'est pas le même que l'intervalle sur ce même plan représenté pas les équations $x^1 = v^2 + v - 1$, $x^2 = v^2 + v - 1$, $0 \leqslant v \leqslant 1$. D'autre part, il ne semble pas raisonnable de conserver de telles distinctions

sans besoin majeur : c'est, tout de **même**, les propriétés géométriques des surfaces que nous étudions. Il faut définir dans quel cas nous avons affaire à « une même surface ». La première raison de considérer comme « une même surface » deux surfaces représentées par les équations du type (3) et coïncidant en tant que lieux géométriques de points dans M_n, ne convient pas, puisque le passage d'un système d'équations (3) à l'autre à l'aide d'une transformation assez bonne des coordonnées dans le cube I^k n'est pas encore garanti par la coïncidence des lieux géométriques. Nous adoptons donc une autre définition de « coïncidence », ou, ce qui est mieux, d'« équivalence » de deux surfaces admissibles :

Une surface S représentée par les équations (3), ou, en abrégé, $S = \{x \in M_n : x = \varphi(u),\ u \in I^k\}$ est dite *équivalente* à une surface $\widetilde{S} = \{x \in M_n : x = \psi(v),\ v \in I^k\}$ (on note $S \sim \widetilde{S}$) s'il existe un difféomorphisme (i.e. une application biunivoque, bicontinue et bidérivable) du cube I^k sur lui-même, $v = v(u)$, tel que $\varphi(u) = \psi(v(u))$ et que le jacobien de l'application $v(u)$ est positif.

Avec cette définition, la relation d'équivalence devient *réflexive* (S est équivalente à elle-même), *symétrique* (si S est équivalente à $\widetilde{S}$, alors $\widetilde{S}$ est équivalente à S) et *transitive* (si S est équivalente à $\widetilde{S}$, et $\widetilde{S}$ est équivalente à $\widetilde{\widetilde{S}}$, alors S est équivalente à $\widetilde{\widetilde{S}}$), et toutes les k-surfaces admissibles peuvent être partagées en les classes deux à deux disjointes de k-surfaces équivalentes. Les intervalles considérés plus haut sur le plan R_2 : $x^1 = u$, $x^2 = u$, $0 \leqslant u \leqslant 1$ et $x^1 = v^2 + v - 1$, $x^2 = v^2 + v - 1$, $0 \leqslant v \leqslant 1$ sont équivalents ; et, par exemple, les intervalles $x^1 = u$, $x^2 = u$, $0 \leqslant u \leqslant 1$ et $x^1 = v^3$, $x^2 = v^3$, $0 \leqslant v \leqslant 1$ ne le sont pas, car, dans le dernier cas, nous avons $u = v^3$, ce qui n'est pas un difféomorphisme de l'intervalle I^1 (l'application inverse n'est pas partout dérivable).

Nous avons encore à montrer que la définition d'équivalence a un sens absolu sur la variété M_n, i.e. qu'elle ne dépend pas du système de coordonnées admissible dans M_n.

Soient $S = \{x \in M_n : x = \varphi(u),\ u \in I^k\}$ et $\widetilde{S} = \{x \in M_n : x = \psi(v),\ v \in I^k\}$ deux surfaces équivalentes dans les coordonnées $x = (x^1, \ldots, x^n)$, et soit $x^{i'} = x^{i'}(x^1, \ldots, x^n)$ un autre système de coordonnées admissible. On peut mettre les équations $x^i = \varphi^i(u)$ sous la forme $x^{i'} = x^{i'}[\varphi^1(u), \ldots, \varphi^n(u)]$, et les équations $x^i = \psi^i(v)$ sous la forme $x^{i'} = x^{i'}[\psi^1(v), \ldots, \psi^n(v)]$. Les systèmes de fonctions $\varphi^{i'}(u)$ et $\psi^{i'}(v)$ possèdent eux aussi les dérivées premières continues. De plus, pour $v = v(u)$ on a $\psi^{i'}(v) = x^{i'}[\psi^1(v), \ldots, \psi^n(v)] = x^{i'}[\varphi^1(u), \ldots, \varphi^n(u)] = \varphi^{i'}(u)$, de sorte que les surfaces S et $\widetilde{S}$ restent équivalentes dans les nouvelles coordonnées.

d. A côté de la définition d'équivalence de deux k-surfaces $S = \{x \in M_n : x = \varphi(u), u \in I^k\}$ et $\widetilde{S} = \{x \in M_n : x = \psi(v), v \in I^k\}$,

on emploie la définition de leur *antiéquivalence* qui diffère de la précédente par ce que le difféomorphisme qui y figure $v = v(u)$ a le jacobien négatif. La relation d'antiéquivalence est symétrique mais non transitive, puisque l'antiéquivalence de S et $\widetilde{S}$ et de $\widetilde{S}$ et $\widetilde{\widetilde{S}}$ implique l'équivalence de $\widetilde{S}$ et $\widetilde{\widetilde{S}}$. L'antiéquivalence devient équivalence après un difféomorphisme supplémentaire, avec le jacobien négatif, du cube I^k, par exemple, le difféomorphisme qui transforme u^1 en $1 - u^1$ sans changer les autres coordonnées.

e. Passons à la définition de l'intégrale d'une forme k-linéaire antisymétrique ω (2) sur une k-surface admissible $S \subset M_n$ (3).

Nous posons

$$\int_S \omega \equiv \int_S \sum_{(i)} f_{(i)}(x)\, dx^{i_1} \wedge \ldots \wedge dx^{i_k} \equiv$$

$$\equiv \int_{I^k} \sum_{(i)} f_{(i)}(\varphi(u)) \sum_{j_1=1}^{k} \frac{\partial x^{i_1}}{\partial u^{j_1}}\, du^{j_1} \wedge \ldots \wedge \sum_{j_k=1}^{k} \frac{\partial x^{i_k}}{\partial u^{j_k}}\, du^{j_k}. \qquad (4)$$

L'invariance tensorielle de l'expression sous le signe somme implique l'indépendance de l'intégrale du système de coordonnées dans M_n.

En transformant le second membre de (4) on obtient:

$$\int_S \omega = \int_{I^k} \sum_{(i)} f_{(i)}(\varphi(u)) \sum_{(j)} \varepsilon^{1 \ldots k}_{j_1 \ldots j_k} \frac{\partial x^{i_1}}{\partial u^{j_1}} \ldots \frac{\partial x^{i_k}}{\partial u^{j_k}}\, du^1 \wedge \ldots \wedge du^k =$$

$$= \int_{I^k} \sum_{(i)} f_{(i)}(\varphi(u)) \det \left\| \frac{\partial x^{i_r}}{\partial u^s} \right\| du^1 \wedge \ldots \wedge du^k =$$

$$= \int_{I_k} \sum_{(i)} f_{(i)}(\varphi(u)) \det \left\| \frac{\partial x^{i_r}}{\partial u^s} \right\| du^1 \ldots du^k \qquad (5)$$

$$(r = 1, \ldots, k;\ s = 1, \ldots, k).$$

f. Montrons que la valeur de l'intégrale obtenue ne change pas en passant d'une surface S à une surface équivalente $\widetilde{S}$. Soit $u = u(v)$ un difféomorphisme du cube I^k qui transforme la représentation $x = \varphi(u)$ de la surface S en la représentation $x = \psi(v)$ de la surface $\widetilde{S}$. Avec la règle de dérivation d'une fonction composée (*1.33b*),

$$\det \left\| \frac{\partial x^{i_r}}{\partial u^s} \right\| = \det \left\| \frac{\partial x^{i_r}}{\partial v^q} \right\| \det \left\| \frac{\partial v^q}{\partial u^s} \right\| \qquad (r, s, q = 1, \ldots, k)$$

et celle de changement des variables dans une intégrale multiple (*3.58a*),

$$\int_{I^k} F(v)\, dv = \int_{I^k} F(v(u)) \det \left\| \frac{\partial v^q}{\partial u^s} \right\| du,$$

passons dans l'intégrale (5) des variables u aux variables v; nous aurons:

$$\int_S \omega = \int_{I^k} \sum_{(i)} f_{(i)}(\varphi(u)) \det \left\| \frac{\partial x^{i_r}}{\partial u^s} \right\| du =$$

$$= \int_{I^k} \sum_{(i)} f_{(i)}(\varphi(u)) \det \left\| \frac{\partial x^{i_r}}{\partial v^q} \right\| \det \left\| \frac{\partial v^q}{\partial u^s} \right\| du =$$

$$= \int_{I^k} \sum_{(i)} f_{(i)}(\psi(v)) \det \left\| \frac{\partial x^{i_r}}{\partial v^q} \right\| dv,$$

ce qu'il fallait.

g. Il est évident que l'intégrale d'une k-forme ω sur une k-surface admissible S possède les propriétés linéaires ordinaires d'une intégrale.

h. Les intégrales d'une k-forme donnée ω sur une k-surface S et sur une surface antiéquivalente ne se diffèrent que par le signe.

7.32. Exemples.

a. Le point $x = \psi(0)$, qui est l'image du point 0 représentant le cube I^0, est, évidemment, une 0-surface admissible dans l'espace M_n. L'intégrale d'une 0-forme $f(x)$ est, par définition, le nombre $f(\varphi(0))$.

b. Une ligne correspondant à un système d'équations

$$x^1 = \varphi^1(u), \ldots, x^n = \varphi^n(u), \quad 0 \leqslant u \leqslant 1 \tag{1}$$

représente, en général, une 1-surface admissible dans l'espace M_n. L'intégrale d'une 1-forme $\omega = \sum_i f_i(x)\, dx^i$ sur la 1-surface (1) est l'intégrale curviligne ordinaire (U9.91).

$$\int_L \omega = \int_L \sum_i f_i(x)\, dx^i = \int_0^1 \sum_{i=1}^n f_i(x(u)) \frac{d\varphi^i(u)}{du} du. \tag{2}$$

Si l'espace M_n est muni d'une métrique riemannienne $g_{ij}(x)$, et si la courbe (1) est non dégénérée, i.e. $\sum_{i=1}^n \left(\frac{d\varphi^i(u)}{du}\right)^2 > 0$, alors on peut y considérer la longueur d'arc s, $0 \leqslant s \leqslant S$ comme paramètre naturel; alors l'intégrale (2) prend la forme

$$\int_L \omega = \int_0^S \sum_{i=1}^n f_i(x(s)) \frac{dx^i}{ds} ds = \int_0^S (f, \tau)\, ds, \tag{3}$$

où f désigne le vecteur $(f_1(x), \ldots, f_n(x))$; et $\tau = \left(\frac{dx^1}{ds}, \ldots, \frac{dx^n}{ds}\right)$ le vecteur unitaire tangent à la courbe L. Dans l'analyse vectorielle (dans l'espace R_n), l'expression (3) s'appelle intégrale du champ vectoriel le long de la courbe L (4.21).

c. Considérons encore l'intégrale d'une $(n-1)$-forme

$$\omega = \sum_{i=1}^{n} (-1)^{i-1} f_i(x)\, dx^1 \wedge \ldots \wedge \widehat{dx^i} \wedge \ldots \wedge dx^n \qquad (4)$$

sur une $(n-1)$-surface admissible S décrite par les équations de la forme

$$\left.\begin{aligned} x^1 &= \varphi^1(u_1, \ldots, u^{n-1}) \\ &\cdots\cdots\cdots\cdots \\ x^n &= \varphi^n(u^1, \ldots, u^{n-1}), \end{aligned}\right\} \quad u \in I^{n-1}.$$

Cette intégrale est

$$\int_S \omega = \int_{I^{n-1}} \sum_{i=1}^{n} (-1)^{i-1} f_i(x(u)) \begin{vmatrix} \frac{\partial x^1}{\partial u^1} & \cdots & \frac{\partial x^1}{\partial u^{n-1}} \\ \cdots & \cdots & \cdots \\ \widehat{\frac{\partial x^i}{\partial u^1}} & \cdots & \widehat{\frac{\partial x^i}{\partial u^{n-1}}} \\ \cdots & \cdots & \cdots \\ \frac{\partial x^n}{\partial u^1} & \cdots & \frac{\partial x^n}{\partial u^{n-1}} \end{vmatrix} du^1 \ldots du^{n-1}.$$

La notion correspondante de l'analyse vectorielle est le flux du champ vectoriel f à travers la surface S (*4.12a*). Pour nous en convaincre, raisonnons comme suit. Supposons qu'il y ait, dans l'espace euclidien R_n, $n-1$ vecteurs ayant, dans une base orthonormale $e_1, \ldots, e_n$, les coordonnées

$$(\underset{1}{\xi^1}, \ldots, \underset{1}{\xi^n}) = \underset{1}{x}, \ldots, (\underset{n-1}{\xi^1}, \ldots, \underset{n-1}{\xi^n}) = \underset{n-1}{x}.$$

Rappelons-nous (*3.62c*) que l'expression

$$[\underset{1}{x}, \ldots, \underset{n-1}{x}] = \begin{vmatrix} e_1 & \underset{1}{\xi^1} & \cdots & \underset{n-1}{\xi^1} \\ \cdots & \cdots & \cdots & \cdots \\ e_n & \underset{1}{\xi^n} & \cdots & \underset{n-1}{\xi^n} \end{vmatrix} = \sum_{i=1}^{n} (-1)^{i-1} e_i \begin{vmatrix} \underset{1}{\xi^1} & \cdots & \underset{n-1}{\xi^1} \\ \cdots & \cdots & \cdots \\ \widehat{\underset{1}{\xi^i}} & \cdots & \widehat{\underset{n-1}{\xi^i}} \\ \cdots & \cdots & \cdots \\ \underset{1}{\xi^n} & \cdots & \underset{n-1}{\xi^n} \end{vmatrix}$$

s'appelle produit vectoriel des vecteurs $\underset{1}{x}, \ldots, \underset{n}{x}$. Dans l'espace euclidien, avec le produit scalaire

$$(x, y) = \left(\sum_{i=1}^{n} \xi^i e_i, \sum_{j=1}^{n} \eta^j e_j\right) = \sum_{i=1}^{n} (\xi^i, \eta^i),$$

le vecteur $[\underset{1}{x}, \ldots, \underset{n-1}{x}]$ est orthogonal à chacun des vecteurs x, et sa longueur représente le volume $(n-1)$-dimensionnel du parallélépipède construit sur ces vecteurs.

Supposons que la $(n-1)$-surface admissible $S \in M_n$ ne soit pas dégénérée, i.e. que le rang de la matrice de Jacobi $\frac{\partial (x^1, \ldots, x^n)}{\partial (u^1, \ldots, u^{n-1})}$ soit partout égal à $n-1$. Alors les vecteurs

$$\underset{1}{dx} = \left(\frac{\partial x^1}{\partial u^1}, \ldots, \frac{\partial x^n}{\partial u^1}\right) du^1, \ \ldots, \ \underset{n-1}{dx} = \left(\frac{\partial x^1}{\partial u^{n-1}}, \ldots, \frac{\partial x^n}{\partial u^{n-1}}\right) du^{n-1}$$

sont linéairement indépendants. Ils se trouvent dans le plan tangent à la surface S (à son point donné), et le vecteur normal N à ce même point peut être déterminé comme leur produit vectoriel:

$$N = [\underset{1}{dx}, \ldots, \underset{n-1}{dx}] = \begin{vmatrix} e_1 & \frac{\partial x^1}{\partial u^1} & \cdots & \frac{\partial x^1}{\partial u^{n-1}} \\ \cdots & \cdots & \cdots & \cdots \\ e_n & \frac{\partial x^n}{\partial u^1} & \cdots & \frac{\partial x^n}{\partial u^{n-1}} \end{vmatrix} du^1 \ldots du^{n-1} =$$

$$= \sum_{i=1}^{n} (-1)^{i-1} e_i \begin{vmatrix} \frac{\partial x^1}{\partial u^1} & \cdots & \frac{\partial x^1}{\partial u^{n-1}} \\ \cdots & \cdots & \cdots \\ \frac{\widehat{\partial x^i}}{\partial u^1} & \cdots & \frac{\widehat{\partial x^i}}{\partial u^{n-1}} \\ \cdots & \cdots & \cdots \\ \frac{\partial x^n}{\partial u^1} & \cdots & \frac{\partial x^n}{\partial u^{n-1}} \end{vmatrix} du^1 \ldots du^{n-1}.$$

Conformément à ceci, le vecteur normal unitaire, au même point, a la forme

$$m = \frac{N}{|N|}.$$

On appelle élément euclidien de la surface S le volume dS du parallélépipède $(n-1)$-dimensionnel construit sur les vecteurs $\underset{1}{dx}, \ldots, \underset{n-1}{dx}$ (cf. *3.62c*), i.e. la longueur du vecteur N. Le flux d'un champ vectoriel $f(x) = (f_1(x), \ldots, f_n(x))$ à travers la surface S dans l'espace riemannien M_n est, par définition, l'intégrale

$$\int_S (f, m)\, dS = \int_{I^{n-1}} \left(f, \frac{N}{|N|}\right) dS =$$

$$= \int_{I^{n-1}} \sum_{i=1}^{n} f_i(x(u)) (-1)^{i-1} \begin{vmatrix} \frac{\partial x^1}{\partial u^1} & \cdots & \frac{\partial x^1}{\partial u^{n-1}} \\ \frac{\widehat{\partial x^i}}{\partial u^1} & \cdots & \frac{\widehat{\partial x^i}}{\partial u^{n-1}} \\ \frac{\partial x^n}{\partial u^1} & \cdots & \frac{\partial x^n}{\partial u^{n-1}} \end{vmatrix} du^1 \ldots du^{n-1}. \quad (5)$$

Nous voyons que l'intégrale d'une $(n-1)$-forme ω (4) sur une $(n-1)$-surface S dans un espace riemannien M_n coïncide avec le flux du champ $f(x)$ à travers cette surface, ce qu'on affirmait.

Nous avons raisonné dans la situationoù l'application $x = x(u)$ n'est pas dégénérée. Or, le premier membre de l'égalité (5) a un sens indépendamment de l'application $x = x(u)$, si la surface S est lisse par morceaux, i.e. peut être décomposée en un nombre fini de parties, dans chacune desquelles elle possède une paramétrisation $x = x(u)$ sans dégénérescence. Le second membre de (5) a son sens indépendant, même pour une application $x = x(u)$ dégénérée. Montrons que, pour une surface S lisse par morceaux, représentée par une application $x = x(u)$ $(I^{n-1} \to S)$ qui ne peut être dégénérée que sur les faces du cube I^{n-1}, l'égalité (5) reste valable. En effet, dans les suppositions faites elle est valable pour le cube $I^{n-1}_{(\varepsilon)}$ concentrique avec le cube I^{n-1} et ayant les arêtes de longueur $1-\varepsilon$, et pour son image $x(I^{n-1}_{(\varepsilon)}) = S_{(\varepsilon)}$ (les intégrales correspondantes se déterminent sans difficulté); en passant à la limite pour $\varepsilon \to 0$, nous arrivons au résultat cherché.

7.33. Chaînes et frontières.

a. La formule de Stokes-Poincaré que nous établirons dans *7.38* liera l'intégrale d'une $(k+1)$-forme $\partial\omega$ sur une $(k+1)$-surface admissible dans M_n à l'intégrale de la forme ω elle-même sur la frontière de cette surface, de même que la formule classique de Newton-Leibniz

$$\int_a^b df(x) = f(b) - f(a) \qquad (1)$$

relie l'intégrale de la différentielle d'une fonction $f(x)$ sur un intervalle avec les valeurs de cette fonction sur la frontière de l'intervalle.

Nous avons à définir ce qu'est la frontière d'une $(k+1)$-surface S. Premièrement, cette frontière représentera un ensemble de k-surfaces dont les paramétrisations se déduiront d'après certaines règles de la paramétrisation de la surface S. Cela déterminera, d'après *7.31e*, l'intégrale de la forme ω sur chacune de ces k-surfaces. Deuxièmement, pour composer l'intégrale de la forme ω sur toute la frontière, certaines des intégrales mentionnées seront munies du signe $+$, d'autres du signe $-$ (comme dans la formule (1)). La définition rigoureuse de la frontière consiste à formuler une règle de détermination de ses parties k-dimensionnelles et une règle de signes.

En rapport à ce qu'on a dit, donnons la définition d'une chaîne. On appelle *chaîne*, plus exactement *k-chaîne*, une somme formelle de la forme $c = \sum_{i=1}^{m} \varepsilon_i S_i$ où $S_1, \ldots, S_m$ sont des k-surfaces admissibles dans M_n, et $\varepsilon_1, \ldots, \varepsilon_m$ des nombres égaux à 0, $+1$ ou -1.

En fait, une chaîne c détermine le schéma d'intégration de n'importe quelle k-forme ω :

$$\int_c \omega = \sum_{i=1}^{m} \varepsilon_i \int_{S_i} \omega. \tag{2}$$

L'expression (2) s'appelle *intégrale de la forme* ω *sur la chaîne* c.

b. Deux k-chaînes $c = \sum_{i=1}^{m} \varepsilon_i S_i$ et $c' = \sum_{j=1}^{p} \varepsilon_j S'_j$ sont dites *équivalentes* si pour toute k-forme ω continue sur M_n a lieu l'égalité

$$\int_c \omega = \int_{c'} \omega. \tag{3}$$

Il est évident que deux chaînes qui ne se diffèrent que par l'ordre de termes sont équivalentes. Ensuite, si une chaîne c' se déduit d'une chaîne c en remplaçant une k-surface S_i par une k-surface $\widetilde{S}_i$ équivalente (*7.31c*), alors la chaîne c' est équivalente à la chaîne c (*7.31f*) ; il en sera de même si la chaîne c' se déduit de c en remplaçant une k-surface S_i par une k-surface antiéquivalente $\widetilde{S}_i$, en même temps, le coefficient correspondant ε_i par $-\varepsilon_i$ (*7.31h*).

c. La définition de la frontière sera donnée non seulement pour les k-surfaces admissibles, mais aussi pour les k-chaînes. La frontière d'une chaîne c est désignée par ∂c.

Supposons d'abord que c soit un cube $I^k \subset R_n$. Le cube I^k est une k-surface admissible avec la paramétrisation

$$x^1 = u^1, \ \ldots, \ x^k = u^k, \quad x^{k+1} = \ldots = x^n = 0,$$
$$0 \leqslant u^i \leqslant 1, \quad i = 1, \ \ldots, \ k.$$

Posons, pour $i = 1, \ \ldots, \ k$:

$$I^k_{i,0} = \{x \in I^k : \ x^i = 0, \ 0 \leqslant x^j \leqslant 1, \ j = 1, \ \ldots, \ \hat{i}, \ \ldots, \ k\},$$
$$I^k_{i,1} = \{x \in I^k : \ x^i = 1, \ 0 \leqslant x^j \leqslant 1, \ j = 1, \ \ldots, \ \hat{i}, \ \ldots, \ k\}.$$

Les ensembles $I^k_{i,0}$ et $I^k_{i,1}$ sont les faces $(k-1)$-dimensionnelles du cube I^k.

Ils sont des $(k-1)$-surfaces admissibles si on les paramétrise comme suit

$$I^k_{i,0} = \{x \in R_n : x^1 = u^1, \ \ldots, \ x^{i-1} = u^{i-1}, \ x^i = 0, \ x^{i+1} = u^i, \ \ldots$$
$$\ldots, \ x^k = u^{k-1}, \ x^{k+1} = \ldots = x^n = 0\},$$
$$I^k_{i,1} = \{x \in R_n : x^1 = u^1, \ \ldots, \ x^{i-1} = u^{i-1}, \ x^i = 1, \ x^{i+1} = u^i, \ \ldots$$
$$\ldots, \ x^k = u^{k-1}, \ x^{k+1} = \ldots = x^n = 0\}.$$

Posons, par définition,

$$\partial I^k = \sum_{i=1}^{k} [(-1)^i I^k_{i,\,0} + (-1)^{i+1} I^k_{i,\,1}] \equiv \sum_{i=1}^{k} \sum_{\alpha=0}^{1} (-1)^{i+\alpha} I^k_{i,\,\alpha}.$$

Soit $x = x(u)$ une représentation paramétrique de n'importe quelle k-surface admissible S. Ici nous posons

$$\partial S = \partial (x(I^k)) = \sum_{i=1}^{k} \sum_{\alpha=0}^{1} (-1)^{i+\alpha} x(I^k_{i,\,\alpha}).$$

Enfin, pour toute k-chaîne $c = \sum_{j=1}^{m} \alpha_j S_j$, nous posons

$$\partial c = \sum_{j=1}^{m} a_j \partial S_j.$$

7.34. Lemme. *Si* $c \sim \widetilde{c}$, alors $\partial c \sim \partial \widetilde{c}$.

Démonstration. Soient, d'abord, c une k-surface S avec la paramétrisation

$$x^j = \varphi^j(u^1, \ldots, u^k), \quad j = 1, \ldots, n, \quad u \in I^k, \tag{1}$$

et $\widetilde{c} = \widetilde{S}$ la même k-surface avec la paramétrisation

$$x^j = \psi^j(v^1, \ldots, v^k), \quad j = 1, \ldots, n, \quad v \in I^k \tag{2}$$

où $\psi(v) = \varphi(u(v))$, et l'application $u = u(v)$ ou bien

$$\left.\begin{aligned} u^1 &= u^1(v^1, \ldots, v^k), \\ u^k &= u^k(v^1, \ldots, v^k) \end{aligned}\right\}$$

est un difféomorphisme du cube I^k. Comme l'application $u = u(v)$ est linéaire en sa partie principale et non dégénérée, elle transforme chaque point intérieur du cube I^k en un point intérieur, donc chaque point sur la frontière en un point sur la frontière. En appliquant le même raisonnement aux faces $(k-1)$-dimensionnelles, nous voyons que l'application $u = u(v)$ transforme chaque point intérieur d'une face en un point intérieur de la face correspondante, et chaque face en une face. L'application $u(v)$ étant biunivoque, elle transforme des faces disjointes en des faces disjointes, autrement dit, des faces opposées en des faces opposées. Supposons, par exemple, qu'une face $I^k_{i,\alpha}$ $(i = 1, \ldots, k;\ \alpha = 0, 1)$ soit transformée par l'application $u(v)$ en une face $I^k_{j,\beta}$ $(j = 1, \ldots, k;\ \beta = 0, 1)$:

$$\left.\begin{aligned} u^1 &= u^1(v^1, \ldots, v^{i-1}, \alpha, v^{i+1}, \ldots, v^k), \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ \beta &= u^j(v^1, \ldots, v^{i-1}, \alpha, v^{i+1}, \ldots, v^k), \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ u^k &= u^k(v^1, \ldots, v^{i-1}, \alpha, v^{i+1}, \ldots, v^k). \end{aligned}\right\} \tag{3}$$

Montrons que *cette application* $u\,(I^k_{i,\alpha}) = I^k_{j,\beta}$ *est un difféomorphisme* $\left(\text{i.e. que } \det \frac{\partial\,(u^1, \ldots, \beta, \ldots, u^k)}{\partial\,(v^1, \ldots, \alpha, \ldots, v^k)} > 0\right)$ *ou bien un antidifféomorphisme* $\left(\text{i.e } \det \frac{\partial\,(u^1, \ldots, \beta, \ldots u^k)}{\partial\,(v^1, \ldots, \alpha, \ldots, v^k)} < 0\right)$, *selon la parité ou imparité respectivement du nombre* $i + j + \alpha + \beta$.

En effet, il résulte de (2) que

$$\frac{\partial u^j\,(v^1, \ldots, v^{i-1}, \alpha, v^{i+1}, \ldots, v^k)}{\partial v^\nu} = 0 \quad \text{pour } \nu \neq i.$$

Etant donnée l'inégalité

$$0 < \begin{vmatrix} \frac{\partial u^1(v^1,\ldots,\alpha,\ldots,v^k)}{\partial v^1} & \cdots & \frac{\partial u^1(v^1,\ldots,\alpha,\ldots,v^k)}{\partial v^i} & \cdots & \frac{\partial u^1(v^1,\ldots,\alpha,\ldots,v^k)}{\partial v^k} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{\partial u^j(v^1,\ldots,\alpha,\ldots,v^k)}{\partial v^1} & \cdots & \frac{\partial u^j(v^1,\ldots,\alpha,\ldots,v^k)}{\partial v^i} & \cdots & \frac{\partial u^j(v^1,\ldots,\alpha,\ldots,v^k)}{\partial v^k} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{\partial u^k(v^1,\ldots,\alpha,\ldots,v^k)}{\partial v^1} & \cdots & \frac{\partial u^k(v^1,\ldots,\alpha,\ldots,v^k)}{\partial v^i} & \cdots & \frac{\partial u^k(v^1,\ldots,\alpha,\ldots,v^k)}{\partial v^k} \end{vmatrix} =$$

$$= \begin{vmatrix} \frac{\partial u^1(v^1,\ldots,\alpha,\ldots,v^k)}{\partial v^1} & \cdots & \frac{\partial u^1(v^1,\ldots,\alpha,\ldots,v^k)}{\partial v^i} & \cdots & \frac{\partial u^1(v^1,\ldots,\alpha,\ldots,v^k)}{\partial v^k} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \frac{\partial u^j(v^1,\ldots,\alpha,\ldots,v^k)}{\partial v^i} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{\partial u^k(v^1,\ldots,\alpha,\ldots,v^k)}{\partial v^1} & \cdots & \frac{\partial u^k(v^1,\ldots,\alpha,\ldots,v^k)}{\partial v^i} & \cdots & \frac{\partial u^k(v_1,\ldots,\alpha,\ldots,v^k)}{\partial v^k} \end{vmatrix},$$

nous développons le dernier déterminant suivant les éléments de la j-ème ligne et voyons que la quantité $\frac{\partial u^j\,(v^1, \ldots, \alpha, \ldots, v^k)}{\partial v^i}$ et le mineur

$$\begin{vmatrix} \frac{\partial u^1}{\partial v^1} & \cdots & \frac{\partial u^1}{\partial v^{i-1}} & \frac{\partial u^1}{\partial v^{i+1}} & \cdots & \frac{\partial u^1}{\partial v^k} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{\partial u^{j-1}}{\partial v^1} & \cdots & \frac{\partial u^{j-1}}{\partial v^{i-1}} & \frac{\partial u^{j-1}}{\partial v^{i+1}} & \cdots & \frac{\partial u^{j-1}}{\partial v^k} \\ \frac{\partial u^{j+1}}{\partial v^1} & \cdots & \frac{\partial u^{j+1}}{\partial v^{i-1}} & \frac{\partial u^{j+1}}{\partial v^{i+1}} & \cdots & \frac{\partial u^{j+1}}{\partial v^k} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{\partial u^k}{\partial v^1} & \cdots & \frac{\partial u^k}{\partial v^{i-1}} & \frac{\partial u^k}{\partial v^{i+1}} & \cdots & \frac{\partial u^k}{\partial v^k} \end{vmatrix} \tag{4}$$

sont non nuls, le mineur (4) ayant pour signe

$$(-1)^{i+j} \operatorname{sign} \frac{\partial u^j(v^1, \ldots, \alpha, \ldots, v^k)}{\partial v^i},$$

quelles que soient les valeurs $v^1, \ldots, v^{i-1}, v^{i+1}, \ldots, v^k$.

Or, si l'on envisage les possibilités présentées fig. 7.3-1, on voit que la quantité $\frac{\partial u^j(v^1, \ldots, \alpha, \ldots, v^k)}{\partial v^i}$ est positive si $\alpha = \beta = 0$ ou $\alpha = \beta = 1$ et négative si $\alpha = 0$, $\beta = 1$ ou $\alpha = 1$, $\beta = 0$. La fonction $\frac{\partial u^j(v^1 \ldots, \alpha, \ldots, v^k)}{\partial v^i}$ a donc le signe $(-1)^{\alpha+\beta}$.

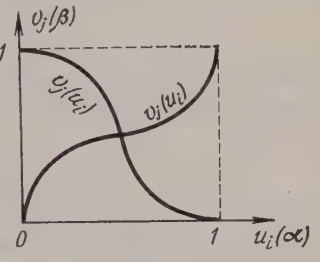


Fig. 7.3-1

Il en résulte que l'application de la face $I^k_{i,\alpha}$ dans la face $I^k_{j,\beta}$ d'après les formules (1) est *un difféomorphisme pour* $\alpha + \beta + i + j$ *pair, et un antidifféomorphisme pour* $\alpha + \beta + i + j$ *impair.* Par conséquent, *les k-surfaces* $\psi(I^k_{j,\beta}) = \psi(u(I^k_{i,\alpha}))$ *et* $(-1)^{\alpha+\beta+i+j}\varphi(I^k_{i,\alpha})$ *sont équivalentes.*

Définitivement, on trouve

$$\partial S = \sum_{i=1}^{k} \sum_{\alpha=0}^{1} (-1)^{i+\alpha} \varphi(I^k_{i,\alpha}),$$

$$\partial \widetilde{S} = \sum_{j=1}^{k} \sum_{\beta=0}^{1} (-1)^{j+\beta} \psi(I^k_{j,\beta}) \sim \sum_{j=1}^{k} \sum_{\beta=0}^{1} (-1)^{j+\beta} (-1)^{i+\alpha+j+\beta} \varphi(I^k_{i,\alpha}) =$$

$$= \sum_{j=1}^{k} \sum_{\beta=0}^{1} (-1)^{i+\alpha} \varphi(I^k_{i,\alpha}) = \sum_{i=1}^{k} \sum_{\alpha=0}^{1} (-1)^{i+\alpha} \varphi(I^k_{i,\alpha}) = \partial S,$$

donc $\partial \widetilde{S} \sim \partial S$.

Il en découle, selon la définition d'équivalence des chaînes c et $\widetilde{c}$, l'équivalence de ∂c et $\partial \widetilde{c}$, ce qui démontre le lemme.

Pour une 0-chaîne c, qui représente une somme algébrique finie de points isolés, à coefficients entiers, nous posons, par définition, $\partial c = 0$.

7.35. Le lemme suivant ressemble au lemme *7.24a.*

L e m m e. *Pour n'importe quelle k-chaîne c, on a* $\partial^2 c \equiv \partial(\partial c) \equiv 0$.

Il suffit d'examiner les cas $k \geqslant 2$ et $c = S = I^k$. Alors on a

$$\partial I^k = \sum_{i=1}^{k} \sum_{\alpha=0}^{1} (-1)^{i+\alpha} I^k_{i,\alpha},$$

$$\partial^2 I^k \equiv \partial(\partial I^k) = \sum_{i=1}^{k} \sum_{\alpha=0}^{1} (-1)^{i+\alpha} \partial I^k_{i,\alpha}.$$

Le symbole $I^k_{i,\alpha}$ désigne la $(k-1)$-surface avec la paramétrisation

$$I^k_{i,\alpha} = \{x \in R_n : x^1 = u^1, \ldots, x^{i-1} = u^{i-1}, x^i = \alpha, x^{i+1} = u^i, \ldots, x^k = u^{k-1}, x^{k+1} = \ldots = x^n = 0\}.$$

D'après la règle générale, on a

$$\partial I^k_{i,\alpha} = \sum_{j=1}^{k-1} \sum_{\beta=0}^{k} (-1)^{j+\beta} (I^k_{i,\alpha})_{j,\beta}$$

où $(I^k_{i,\alpha})_{j,\beta}$ est la $(k-2)$-surface avec la paramétrisation

$$(I^k_{i,\alpha})_{j,\beta} = \{x \in R_n : x^1 = u^1, \ldots, x^j = \beta, x^{j+1} = u^j, \ldots, x^{i-1} = u^{i-2}, x^i = \alpha, \ldots, x^k = u^{k-2}, \ldots, x^{k+1} = \ldots = x^n = 0\}, \quad (1)$$

si $j < i$, et avec la paramétrisation

$$(I^k_{i,\alpha})_{j,\beta} = \{x \in R_n : x^1 = u^1, \ldots, x^i = \alpha, x^{i+1} = u^i, \ldots, x^{j+1} = \beta, x^{j+2} = u^j, \ldots, x^k = u^{k-2}, x^{k+1} = \ldots = x^n = 0\}, \quad (2)$$

si $j \geqslant i$. Ainsi

$$\partial^2 I^k = \sum_{i=1}^{k} \sum_{\alpha=0}^{1} \sum_{j=1}^{k-1} \sum_{\beta=0}^{1} (-1)^{i+\alpha+j+\beta} (I^k_{i,\alpha})_{j,\beta}. \quad (3)$$

Fixons dans cette somme le terme $(I^k_{i,\alpha})_{j,\beta}$ avec α, β, i, j donnés et avec $j < i$. Il a la forme (1). Examinons ensuite le terme $(I^k_{j,\beta})_{i-1,\alpha}$. Comme on y a $i - 1 \geqslant j$, son expression est donnée par la formule (2):

$$(I^k_{j+1,\alpha})_{i,\beta} = \{x \in R_n : x^1 = u^1, \ldots, x^j = \beta, \ldots x^i = \alpha, \ldots, x^k = u^{k-2}, x^{k+1} = \ldots = x^n = 0\}.$$

Nous voyons que $(I^k_{j,\beta})_{i,\alpha} = (I^k_{i,\alpha})_{j,\beta}$. Ainsi, chaque terme de la somme (3) dont le deuxième indice latin est inférieur au premier s'annule en somme avec l'autre terme dont le deuxième indice latin est supérieur ou égal au premier. Définitivement, toute la somme (3) est nulle, ce qu'il fallait.

Une chaîne c dont la frontière est nulle s'appelle *cycle*.

Le lemme démontré affirme que *toute frontière est un cycle*.

7.36. Théorème. *Soit* ω *une* $(n-1)$*-forme différentiable dans le cube* $I^n \subset R^n$. *Alors*

$$\int_{I^n} \partial\omega = \int_{\partial I^n} \omega. \quad (1)$$

Démonstration. Sans restreindre la généralité, il suffit de considérer la forme:

$$\omega = f(x)\, dx^1 \wedge \ldots \wedge \widehat{dx^i} \wedge \ldots \wedge dx^n. \quad (2)$$

On a

$$d\omega = (-1)^{i-1} \frac{\partial f(x)}{\partial x^i} dx^1 \wedge \ldots \wedge dx^i \wedge \ldots \wedge dx^n.$$

En passant à une intégrale itérée, on trouve

$$\int_{I^n} d\omega = (-1)^{i-1} \int_{I^n} \frac{\partial f(x)}{\partial x^i} dx^1 \ldots dx^n =$$

$$= (-1)^{i-1} \int_{I^{n-1}} \left\{ \int_0^1 \frac{\partial f(x)}{\partial x^i} dx^i \right\} dx^1 \ldots dx^{i-1} dx^{i+1} \ldots dx^n =$$

$$= (-1)^{i-1} \int_{I^{n-1}} [f(x^1, \ldots, x^{i-1}, 1, x^{i+1}, \ldots, x^n) -$$

$$- f(x^1, \ldots, x^{i-1}, 0, x^{i+1}, \ldots, x^n)] dx^1 \ldots dx^{i-1} dx^{i+1} \ldots dx^n. \quad (3)$$

Or, nous savons que

$$\partial I^n = \sum_{j=1}^{n} [(-1)^j I^n_{j,0} + (-1)^{j+1} I^n_{j,1}],$$

où $I^n_{j,\alpha}$ est la $(n-1)$-surface avec la paramétrisation

$$x^1 = u^1, \ldots, x^j = \alpha, \ldots, x^n = u^{n-1} \quad (\alpha = 0, 1). \quad (4)$$

Par conséquent,

$$\int_{\partial I^n} \omega = \sum_{j=1}^{n} (-1)^j \int_{I^n_{j,0}} \omega + \sum_{j=1}^{n} (-1)^{j+1} \int_{I^n_{j,1}} \omega.$$

La réduction de ces intégrales à des intégrales $(n-1)$-uples ordinaires s'effectue à l'aide de la représentation (4). Puisque, dans le j-e terme, nous avons $dx^j = 0$, il ne reste de toute la somme que le terme avec $j = i$, car seulement dans ce cas l'expression sous le signe somme peut ne pas être nulle. Ainsi,

$$\int_{\partial I^n} \omega = (-1)^j \int_{I^n_{i,0}} f(x^1, \ldots, 0, \ldots, x^n)\, dx +$$

$$+ (-1)^{i+1} \int_{I^n_{i,1}} f(x^1, \ldots, 1, \ldots, x^n)\, dx = \int_{I^n} \partial\omega \quad (5)$$

et le théorème est démontré.

En fait, la formule (5) n'est rien d'autre que la formule d'Ostrogradski *4.13* (5) formulée en termes des formes différentielles pour le cas où le domaine d'intégration est un cube n-dimensionnel.

7.37. Lemme. *Pour une k-forme différentielle $\omega(x)$, l'opération ∂ et le passage aux nouvelles coordonnées $x = x(u)$ $(u = (u^1, \ldots, u^m))$ commutent: on peut d'abord effectuer le changement des variables, puis réaliser l'opération ∂ (par rapport aux nouvelles variables).*

Démonstration. Supposons d'abord que la forme ω soit une fonction (0-forme) $f(x)$. Nous avons alors

$$\partial f(x) = \sum_{i=1}^{n} \frac{\partial f(x)}{\partial x^i} dx^i,$$

$$\partial f(x(u)) = \sum_{i=1}^{n} \frac{\partial f(x(u))}{\partial x^i} \sum_{j=1}^{m} \frac{\partial x^i}{\partial u^j} du^j.$$

Or, en effectuant la différentiation après le changement des variables, nous obtenons

$$\partial f(x(u)) = \sum_{j=1}^{m} \frac{\partial f(x(u))}{\partial u^j} du^j = \sum_{j=1}^{m} \left\{ \sum_{i=1}^{n} \frac{\partial f(x(u))}{\partial x^i} \frac{\partial x^i(u)}{\partial u^j} \right\} du^j,$$

ce qui coïncide avec l'expression précédente. Ainsi, pour les 0-formes, la proposition est valable. Procédons ensuite par récurrence: supposons que le lemme soit valable pour toute k-forme et démontrons-le pour toute $(k+1)$-forme. Il suffit de considérer une $(k+1)$-forme $\omega \wedge dx^s$ où ω est une k-forme. Alors, selon *7.22g*, on a

$$\partial(\omega \wedge dx^s) = d\omega \wedge dx^s,$$

$$\partial(\omega \wedge dx^s)[u] = \partial\omega[u] \wedge \sum_{j-1}^{m} \frac{\partial x^s}{\partial u^j} du^j. \tag{1}$$

D'autre part,

$$(\omega \wedge dx^s)[u] = \omega[u] \wedge dx^s[u],$$

et, comme $\partial^2 = 0$, en différentiant par rapport aux variables u et compte tenu de l'hypothèse de récurrence, on trouve

$$\partial[(\omega \wedge dx^s)[u]] = \partial(\omega[u]) \wedge dx^s[u] =$$

$$(\partial\omega(x))[u] \wedge \sum_{j=1}^{m} \frac{\partial x^s}{\partial u^j} du^j. \tag{2}$$

En comparant (1) et (2) nous voyons que le lemme est juste.

7.38. Théorème (Stokes-Poincaré). *Soient ω une $(k-1)$-forme différentiable dans un domaine $G \subset M_n$ et c n'importe quelle k-chaîne dans G. Alors on a:*

$$\int_c \partial\omega = \int_{\partial c} \omega.$$

Il suffit de réaliser la démonstration pour le cas où la chaîne c est une k-surface admissible S. Soit $x = x(u)$ une représentation paramétrique de S, les paramètres u parcourant, comme d'ordinaire, tout le cube I^k. Conformément à la définition *7.31d* de l'intégrale d'une forme et d'après le lemme *7.37*, on a

$$\int_S \partial\omega(x) = \int_{I^k} (\partial\omega(x))[u] = \int_{I^k} \partial[\omega(x(u))].$$

Or,

$$\int_{\partial S} \omega(x) = \int_{\partial I^k} \omega(x(u)),$$

et le théorème découle de l'égalité

$$\int_{I^k} \partial[\omega(x(u))] = \int_{\partial I^k} \omega(x(u))$$

qui résulte du théorème *7.36* (pour $n = k$).

7.39. Cas particuliers du théorème de Stokes-Poincaré.

a. $k = 1$. Dans ce cas la chaîne c est une 1-chaîne. Supposons, pour simplifier, que c représente une ligne L. i.e. l'image $x = \varphi(u)$ de l'intervalle $I^1 = \{u: 0 \leqslant u \leqslant 1\}$:

$$L = \{x \in R_n: x^1 = \varphi^1(u), \ldots, x^n = \varphi^n(u)\},$$

où $\varphi^j(u)$, comme d'habitude, sont des fonctions continûment dérivables. La frontière de l'intervalle I^1 est la chaîne formée de deux points 0 et 1, le premier avec le signe —, le deuxième avec +. Par conséquent, la frontière de la ligne est la chaîne de deux points $\varphi(1) - \varphi(0)$. La forme ω est une 0-forme $\omega = f(x)$, et $\partial\omega = \sum_{i=1}^{n} \frac{\partial f(x)}{\partial x^i} dx^i$. Nous arrivons à la formule

$$\int_L \sum_{i=1}^{n} \frac{\partial f(x)}{\partial x^i} dx^i = f[\varphi(1)] - f[\varphi(0)]$$

qui représente la généralisation bien connue de la formule de Newton-Leibniz aux intégrales curvilignes.

b. $k = 2$. Ici la chaîne c est une 2-chaîne. Soit, pour simplifier, $c = S$ l'image $x = \varphi(u)$ du carré $I^2 = \{(u^1, u^2) \in R^2: 0 \leqslant u^1 \leqslant 1, 0 \leqslant u^2 \leqslant 1\}$:

$$S = \{x \in R_n: x^1 = \varphi^1(u^1, u^2), \ldots, x^n = \varphi^n(u^1, u^2)\}.$$

La frontière du carré I^2 est la chaîne $I^2_{2,0} + I^1_{1,1} - I^2_{2,1} - I^2_{1,0}$. Intégrer sur cette chaîne consiste à intégrer sur u^1 pour $u^2 = 0$ avec le signe $+$, sur u^2 pour $u^1 = 1$, avec le signe $+$, sur u^1 pour $u^2 = 1$, avec le signe $-$, sur u^2 pour $u^1 = 0$, avec le signe $-$; le résultat représente l'intégrale sur le contour fermé du carré I^2 dans

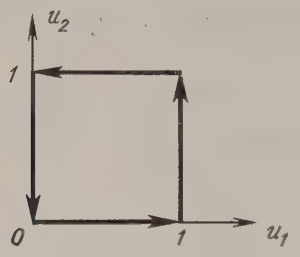


Fig. 7.3-2

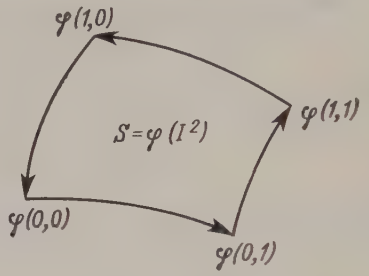


Fig. 7.3-3

le sens positif (fig. 7.3-2). L'intégration sur la chaîne ∂S (fig. 7.3-3) a un sens correspondant. La forme ω est ici une 1-forme $\sum_{i=1}^{n} f_i(x)\, dx^i$ pour laquelle

$$\partial\omega = \sum_{\alpha<\beta} \left(\frac{\partial f_\beta}{\partial x^\alpha} - \frac{\partial f_\alpha}{\partial x^\beta} \right) dx^\alpha \wedge dx^\beta.$$

Pour $n = 3$, on peut faire correspondre à la forme $\partial\omega$ le champ vectoriel

$$\operatorname{rot} f = \left\{ \frac{\partial f_2}{\partial x^3} - \frac{\partial f_3}{\partial x^2}, \frac{\partial f_3}{\partial x^1} - \frac{\partial f_1}{\partial x^3}, \frac{\partial f_1}{\partial x^2} - \frac{\partial f_2}{\partial x^1} \right\}.$$

Conformément à *7.32c*, l'intégrale de la forme $\partial\omega$ sur la surface S est le flux du champ vectoriel rot f à travers cette surface. Le sens de la normale à la surface S est déterminé par le produit vectoriel des vecteurs $\frac{\partial x}{\partial u^1}$ et $\frac{\partial x}{\partial u^2}$ tangents aux lignes des familles $u^2 = C$ et $u^1 = C$ respectivement et dirigés dans les sens de croissance des paramètres respectifs. Pour cette raison, dans un système de coordonnées à droite, la concordance de la direction de la normale et de celle de parcours de la frontière s'effectue d'après la règle: vu l'extrémité du vecteur m, le parcours de la frontière se fait dans le sens inverse des aiguilles d'une montre (fig. 7.3-4).

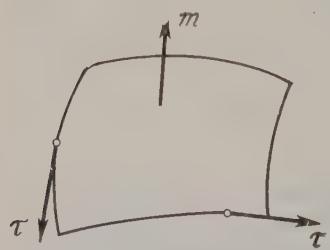


Fig. 7.3-4

Nous retrouvons la formule classique de Stokes *4.26* (5):

$$\int_S (\operatorname{rot} f, m)\, dS = \int_\Gamma (f, \tau)\, d\Gamma.$$

Dans la formule classique de Stokes il s'agit d'une surface lisse par morceaux S, avec une frontière lisse par morceaux Γ. La formule, que nous avons obtenue est écrite pour une surface lisse dont la frontière contient quatre points anguleux au plus; pour passer au cas général, il suffit de remarquer qu'une surface lisse par morceaux avec une frontière lisse par morceaux peut toujours être coupée en un nombre fini des parties qui représentent des surfaces lisses avec quatre points anguleux sur la frontière; en écrivant la formule de Stokes obtenue ici pour chacune de ces parties et en les additionnant nous arrivons à la formule de Stokes classique.

c. $k = n$. Dans ce cas, la chaîne c est une n-chaîne; supposons, pour simplifier, que $c = G$ soit l'image $x = \varphi(u)$ du cube $I^n = \{u \in R_n: 0 \leqslant u^1 \leqslant 1, \ldots, 0 \leqslant u^n \leqslant 1\}$:

$$G = \{x \in R_n: x^1 = \varphi^1(u^1, \ldots, u^n), \ldots, x^n = \varphi^n(u^1, \ldots, u^n)\}.$$

Supposons encore que le jacobien de l'application $x = \varphi(u)$ soit non négatif et ne s'annule que sur un ensemble Z de dimension $\leqslant n - 2$. Alors la normale extérieure à la face $I^n_{i,\alpha}$ du cube I^n se transformera partout, sauf sur l'ensemble $Z \cap I^n_{i,\alpha}$, en la normale extérieure au morceau correspondant $S_{i,\alpha}$ de la frontière de G. Comme la normale extérieure à la face $I^n_{i,\alpha}$ est colinéaire au vecteur de base e_i, à savoir elle représente le vecteur

$$(-1)^{\alpha-1} e_i = (-1)^{\alpha-1} [e_1, \ldots, e_{i-1}, e_{i+1}, \ldots, e_n] \cdot (-1)^{i-1} = \\ = (-1)^{\alpha+i} [e_1, \ldots, e_{i-1}, e_{i+1}, \ldots, e_n],$$

alors le vecteur correspondant de la normale extérieure au morceau $S_{i,\alpha}$ peut être donné par l'expression

$$\nu = \frac{N}{|N|}, \quad N = (-1)^{\alpha+i} \left[\frac{\partial x}{\partial u^1}, \ldots, \frac{\partial x}{\partial u^{i-1}}, \frac{\partial x}{\partial u^{i+1}}, \ldots, \frac{\partial x}{\partial u^n}\right].$$

La forme ω est ici une $(n-1)$-forme

$$\omega = \sum_{j=1}^{n} (-1)^{j-1} f_j(x)\, dx^1 \wedge \ldots \wedge dx^{j-1} \wedge dx^{j+1} \wedge \ldots \wedge dx^n$$

à laquelle nous faisons correspondre le champ vectoriel $f = \{f_1, \ldots, f_n\}$.

Conformément à *7.32c*, l'intégrale $\int_{\varphi(I^n_{i,\alpha})} \omega$ est égale au flux du champ f à travers la surface $S_{i,\alpha}$ avec la normale unitaire $m_{i,\alpha} = N_{i,\alpha}/|N_{i,\alpha}|$ où

$$N_{i,\alpha} = \left[\frac{\partial x}{\partial u^1}, \ldots, \frac{\partial x}{\partial u^{i-1}}, \frac{\partial x}{\partial u^{i+1}}, \ldots, \frac{\partial x}{\partial u^n}\right] = (-1)^{i+\alpha} N;$$

autrement dit, $\int\limits_{S_{i,\alpha}} (f, m_{i,\alpha})\, dS = \int\limits_{\varphi(I^n_{i,\alpha})} \omega$. Donc,

$$\int\limits_{\partial G} \omega = \sum_{i=1}^{n} \sum_{\alpha=0}^{1} (-1)^{i+\alpha} \int\limits_{\varphi(I^n_{i,\alpha})} \omega =$$

$$= \sum_{i=1}^{n} \sum_{\alpha=0}^{1} \int\limits_{S_{i,\alpha}} (f, (-1)^{i+\alpha} m_{i,\alpha})\, dS = \sum_{i=1}^{n} \sum_{\alpha=0}^{1} \int\limits_{S_{i,\alpha}} (f, m)\, dS =$$

$$= \int\limits_{S} (f, m)\, dS.$$

D'autre part, nous avons

$$\int\limits_{G} \partial\omega = \int\limits_{G} \sum_{j=1}^{n} \frac{\partial f_j(x)}{\partial x^j}\, dx^1 \wedge \ldots \wedge dx^n.$$

Ainsi, la formule de Stokes se réduit dans ce cas à la forme

$$\int\limits_{S} (f, m)\, dS = \int\limits_{G} \sum_{j=1}^{n} \frac{\partial f(x)}{\partial x^j}\, dG \qquad (1)$$

et représente la formule d'Ostrogradski *4.13* (5) :

$$\int\limits_{S} (f, \nu)\, dS = \int\limits_{G} \operatorname{div} f(x)\, dG.$$

Ici le domaine G est supposé possédant une frontière lisse par morceaux S. Dans la formule (1), la frontière a une forme un peu particulière (l'image de la frontière du cube ne peut pas être très « anguleux », de sorte que, par exemple, elle ne peut être un polyèdre à un grand nombre de faces). On peut donc passer dans (*b*) des domaines figurant dans (1) à n'importe quel domaine G avec la frontière lisse par morceaux en coupant G en un nombre de parties admettant la formule de Stokes, en écrivant (1) pour chaque telle partie et en additionnant les résultats ; nous omettons les détails.

§ 7. 4. Codifférentiation

7.41. Forme adjointe.

a. Soit l'espace R_n muni d'un produit scalaire avec un tenseur métrique g_{ij}. Alors, parmi toutes les bases, on dégage naturellement la classe des bases $\{e_\alpha\}$ orthonormales pour lesquelles $(e_\alpha, e_\beta) = \delta_{\alpha\beta}$.

Une forme antisymétrique $A(\underset{1}{x}, \ldots, \underset{k}{x})$ peut être écrite d'après *7.16* (2) :

$$A(\underset{1}{x}, \ldots, \underset{k}{x}) = \sum_{(\alpha)}{}^{\sim} a(\alpha)\, \xi^{\alpha_1} \wedge \ldots \wedge \xi^{\alpha_k},$$

où la sommation se fait sur les multinuméros strictement ordonnés $(\alpha) = (\alpha_1, \ldots, \alpha_k)$.

Définissons la forme adjointe $*A$ de $n - k$ variables en posant, dans une base orthonormale,

$$*A(\underset{1}{x}, \ldots, \underset{n-k}{x}) = \sum_{(\gamma)} \tilde{b}(\gamma)\, \xi^{\gamma_1} \wedge \ldots \wedge \xi^{\gamma_k}_{n-k} \qquad (1)$$

où $(\gamma) = (\gamma_1, \ldots, \gamma_{n-k})$ est le multinuméro complémentaire à (α) (*7.11c*) qui est aussi strictement ordonné, et

$$\tilde{b}_{(\gamma)} = (-1)^{\sum \alpha_i} \tilde{a}_{(\alpha)}. \qquad (2)$$

b. On peut montrer directement que la définition (1)-(2) ne dépend pas de la base orthonormale (exercice 16). Nous ferons autrement : une autre définition de la forme adjointe sera formulée pour toutes les bases à la fois ; son invariance découlera de sa notation tensorielle ; dans le cas de bases orthonormales elle se réduira à la définition (1)-(2). Utilisons la notation *7.16* (3) de la forme A :

$$A(\underset{1}{x}, \ldots, \underset{k}{x}) = \sum_{(i)} a_{(i)}\, \xi^{i_1} \wedge \ldots \wedge \xi^{i_k},$$

avec la sommation sur tous les $(k\text{-}n)$-numéros stricts (i), les coefficients $a_{(i)}$ étant des fonctions antisymétriques du $(k\text{-}n)$-numéro (i). Posons

$$*A(\underset{1}{x}, \ldots, \underset{n-k}{x}) = \sum_{(j)} b_{(j)}\, \xi^{j_1} \wedge \ldots \wedge \xi^{j_{n-k}}, \qquad (3)$$

où 107

$$b_{(j)} = C_{kn} \sum_{(i)} \sum_{(s)} a_{(s)} g^{i_1 s_1} \ldots g^{i_k s_k} \varepsilon^{1 \ldots k \ldots n}_{i_1 \ldots i_k\, j_1 \ldots j_{n-k}} \sqrt{G}. \qquad (4)$$

Dans cette formule, $G = \det \| g_{i,j} \|$ et C_{kn} est une constante dont la valeur sera précisée plus bas. Montrons que la définition (3)-(4) donne un tenseur $n-k$ fois covariant. Nous avons

$$\det \| g_{i'j'} \| = \det \| g_{ij} p^i_{i'} p^j_{j'} \| = \det \| g_{ij} \| \det{}^2 \| p^i_{i'} \| ;$$

puis

$$\varepsilon^{1' \ldots k' \ldots n'}_{i'_1 \ldots i'_k j'_1 \ldots j'_{n-k}} p^{i'_1}_{i_1} \ldots p^{i'_k}_{i_k} p^{j'_1}_{j_1} \ldots p^{j'_{n-k}}_{j_{n-k}} =$$

$$= \begin{vmatrix} p^{1'}_{i_1} \ldots p^{1'}_{i_k} \ldots p^{1'}_{j_{n-k}} \\ \ldots\ldots\ldots\ldots \\ p^{n'}_{i_1} \ldots p^{n'}_{i_k} \ldots p^{n'}_{j_{n-k}} \end{vmatrix} = \varepsilon^{1 \ldots k \ldots n}_{i_1 \ldots i_k \ldots j_{n-k}} \begin{vmatrix} p^{1'}_1 \ldots p^{1'}_n \\ \ldots\ldots \\ p^{n'}_1 \ldots p^{n'}_n \end{vmatrix},$$

d'où

$$\varepsilon^{1' \ldots k' \ldots n'}_{i'_1 \ldots i'_k \ldots j'_{n-k}} = p^{i_1}_{i'_1} \ldots p^{i_k}_{i'_k} \ldots p^{j_{n-k}}_{j'_{n-k}} \varepsilon^{1 \ldots k \ldots n}_{i_1 \ldots i_k \ldots j_{n-k}} \det \| p^{i'}_i \|.$$

Par conséquent,

$$\varepsilon^{1'\ldots n'}_{i'_1\ldots j'_{n-k}}\sqrt{\det\|g_{i'j'}\|}=$$

$$=p^{i_1}_{i'_1}\ldots p^{j_{n-k}}_{j'_{n-k}}\varepsilon^{1\ldots n}_{i_1\ldots j_{n-k}}\det\|p^{i}_{i'}\|\det\|p^{i'}_{i}\|\sqrt{\det\|g_{ij}\|}=$$

$$=p^{i_1}_{i'_1}\ldots p^{j_{n-k}}_{j'_{n-k}}\varepsilon^{1\ldots n}_{i_1\ldots j_{n-k}}\sqrt{\det\|g_{ij}\|},$$

de sorte que la quantité $\varepsilon^{1\ldots n}_{i_1\ldots j_{n-k}}\sqrt{\det\|g_{ij}\|}$ est un tenseur n fois covariant. Or, $a_{(i)}$ représente un tenseur k fois covariant (*7.16c*) et g^{ij} un tenseur deux fois contravariant ; la contraction (4) est donc un tenseur $n-k$ fois covariant, ce qu'il fallait montrer.

Dans une base orthogonale, nous avons $g^{ij}=g_{ij}=\delta_{ij}$ et la formule (4) devient

$$b_{(j)}=C_{kn}\sum_{(i)}a_{(i)}\varepsilon^{1\ldots k\ldots n}_{i_1\ldots i_k j_1\ldots j_{n-k}}.$$

Pour (j) donné, il n'est pas dépourvu de sens de considérer les multinuméros (i) complémentaires à (j). Si $O(i)=(\alpha)$, $O(j)=(\gamma)$, alors $a_{(i)}=\frac{1}{k!}\varepsilon^{\alpha_1\ldots\alpha_k}_{i_1\ldots i_k}\widetilde{a}_{(\alpha)}$ (*7.16* (5)), et, comme il n'y a qu'un seul multinuméro strictement ordonné (α) complémentaire à (γ):

$$b_{(\gamma)}=\frac{C_{kn}}{k!}\widetilde{a}_{(\alpha)}\sum_{(i)}\varepsilon^{\alpha_1\ldots\alpha_k}_{i_1\ldots,i_k}\varepsilon^{1\ldots k\ldots n}_{i_1\ldots i_k\gamma_1\ldots\gamma_{n-k}}.$$

L'expression sous le signe de la sommation intérieure ne dépend plus du multinuméro (i), car, d'après *7.11* (2), on a

$$\varepsilon^{\alpha_1\ldots\alpha_k}_{i_1\ldots i_k}\varepsilon^{1\ldots k\ldots n}_{i_1\ldots i_k\gamma_1\ldots\gamma_{n-k}}=\varepsilon^{\alpha_1\ldots\alpha_k}_{\alpha_1\ldots\alpha_k}\varepsilon^{1\ldots k\ldots n}_{\alpha_1\ldots\alpha_k\gamma_1\ldots\gamma_{n-k}}=$$

$$=\varepsilon^{1\ldots k\ldots n}_{\alpha_1\ldots\alpha_k\gamma_1\ldots\gamma_{n-k}}=(-1)^{\sum\alpha_i+\frac{k(k+1)}{2}}.$$

Ainsi

$$b_{(\gamma)}=\frac{C_{kn}}{k!}\widetilde{a}_{(\alpha)}k!\,(-1)^{\sum\alpha_i+\frac{k(k+1)}{2}}=C_{kn}\widetilde{a}_{(\alpha)}(-1)^{\sum\alpha_i+\frac{k(k+1)}{2}},$$

et, selon *7.16* (10),

$$\widetilde{b}_{(\gamma)}=C_{kn}(n-k)!\,b_{(\gamma)}=C_{kn}(n-k)!\,\widetilde{a}_{(\alpha)}(-1)^{\sum\alpha_i+\frac{k(k+1)}{2}}.$$

En posant maintenant $C_{kn}=(-1)^{\frac{k(k+1)}{2}}\frac{1}{(n-k)!}$ on trouve

$$\widetilde{b}_{(\gamma)}=(-1)^{\sum\alpha_i}\widetilde{a}_{(\alpha)},$$

ce qui coïncide avec la formule (2).

c. Exemples. Une 0-forme C (constante) a pour son adjointe la n-forme

$$C\xi^1 \wedge \ldots \wedge \xi^n.$$

L'adjointe d'une 1-forme $\sum_{i=1}^{n} a_i \xi^i$ est la $(n-1)$-forme qui, dans une base orthonormale, s'écrit:

$$\sum (-1)^i a_i \xi^1 \wedge \ldots \wedge \xi^{i-1} \wedge \xi^{i+1} \wedge \ldots \wedge \xi^n.$$

d. Dans le cas général, on a les égalités

$$*(aA + bB) = a(*A) + (*B) \tag{5}$$

(A et B sont des k-formes antisymétriques, a et b des nombres) et

$$*(*A) = (-1)^{\frac{n(n+1)}{2}} A. \tag{6}$$

L'égalité (5) est immédiate. L'égalité (6) résulte de

$$(-1)^{\sum \alpha_i + \sum \gamma_i} = (-1)^{1+\ldots+n} = (-1)^{\frac{n(n+1)}{2}}.$$

7.42. Forme adjointe dans un espace riemannien.

a. A présent, pour une k-forme différentielle donnée

$$\omega = \sum_{(i)} a_{(i)}(x)\, dx^{i_1} \wedge \ldots \wedge dx^{i_k} \equiv \sum_{(\alpha)} \widetilde{a}_{(\alpha)}(x)\, dx^{\alpha_1} \wedge \ldots \wedge dx^{\alpha_k} \tag{1}$$

dans un espace riemannien élémentaire M_n, nous définissons la $(n-k)$-forme adjointe $*\omega$ par l'égalité

$$*\omega = \sum_{(j)} b_{(j)}(x)\, dx^{j_1} \wedge \ldots \wedge dx^{j_{n-k}} \equiv$$

$$\equiv \sum_{(\gamma)} \widetilde{b}_{(\gamma)}(\widetilde{x})\, dx^{\gamma_1} \wedge \ldots \wedge dx^{\gamma_{n-k}}, \tag{2}$$

où, conformément à *7.41* (4),

$$b_{(j)}(x) = (-1)^{\frac{k(k+1)}{2}} \frac{1}{(n-k)!} \sum_{(i)} \sum_{(s)} a_{(s)}(x)\, g^{i_1 s_1} \ldots g^{i_k s_k} \times$$

$$\times \varepsilon^{1 \ldots k \ldots n}_{i_1 \ldots i_k j_1 \ldots j_{n-k}} \sqrt{\det \| g_{ij}(x) \|}. \tag{3}$$

Si, en un point donné, le système de coordonnées $\{x^j\}$ est orthonormal, c'est-à-dire que $g_{ij} = g^{ij} = \delta_{ij}$, alors, comme dans *7.41b*, on peut écrire une formule plus simple:

$$\widetilde{b}_{(\gamma)}(x) = (-1)^{\sum \alpha_i} \widetilde{a}_{(\alpha)}(x), \tag{4}$$

où (γ) est le multinuméro strictement ordonné complémentaire à (α).

b. Conformément à *7.41d*, ont lieu les relations:

$$*(a_1\omega_1 + a_2\omega_2) = a_1(*\omega_1) + a_2(*\omega_2) \tag{5}$$

pour n'importe quelles k-formes antisymétriques ω_1 et ω_2 et n'importe quelles constantes a_1 et a_2,

$$*(*\omega) = (-1)^{\frac{n(n+1)}{2}}\omega \tag{6}$$

pour n'importe quelle k-forme antisymétrique ω.

7.43. a. Définissons l'opération de la *différentiation adjointe* ou bien de la *codifférentiation* d'une forme antisymétrique sur un espace riemannien. Cette opération est définie par la relation

$$\delta\omega = *\,\partial *\,\omega. \tag{1}$$

b. La codifférentielle d'une k-forme est une $n - (1 + (n - k)) = (k - 1)$-forme. En particulier, la codifférentielle d'une 0-forme est 0. La codifférentielle d'une 1-forme $\omega = \sum f_j(x)\,dx^j$ est une 0-forme; calculons-la en coordonnées orthonormales:

$$*\omega = \sum_{j=1}^{n} (-1)^j f_j(x)\,dx^1 \wedge \ldots \wedge dx^{j-1} \wedge dx^{j+1} \wedge \ldots \wedge dx^n,$$

$$\partial(*\omega) = \sum_{j=1}^{n} (-1)^j \frac{\partial f_j(x)}{\partial x^j}\,dx^j \wedge dx^1 \wedge \ldots \wedge dx^{j-1} \wedge dx^{j+1} \wedge \ldots$$

$$\ldots \wedge dx^n = -\sum_{j=1}^{n} \frac{\partial f_j(x)}{\partial x^j}\,dx^1 \wedge \ldots \wedge dx^n,$$

$$\delta\omega = *d*\omega = (-1)^{\frac{n(n+1)}{2}} \sum_{j=1}^{n} \frac{\partial f_j(x)}{\partial x^j}.$$

c. La codifférentiation, comme les opérations d et $*$, est une opération linéaire. Ensuite, il résulte de *7.42* (6) et de la propriété associative que

$$\delta^2 = (*\partial *)(*\partial *) = (*\partial)(**)(\partial *) = (*\partial)(-1)^{\frac{n(n+1)}{2}}(\partial *) =$$

$$= (-1)^{\frac{n(n+1)}{2}}(*\partial)(\partial *) = (-1)^{\frac{n(n+1)}{2}} *(\partial\partial) * = 0.$$

7.44. Opérateur de Laplace.

a. Un rôle important est joué par l'opérateur $\partial\delta + \delta\partial$. Nous allons montrer que, *dans l'espace euclidien n-dimensionnel, cet opérateur, dans toute base orthonormée, à un facteur constant près, coïncide*

avec l'opérateur de Laplace appliqué à chaque coefficient de la forme:

$$(\partial\delta+\delta\partial)\sum_{(\alpha)} a_{(\alpha)}(x)\,dx^{\alpha_1}\wedge\ldots\wedge dx^{\alpha_k}=$$

$$=-(-1)^{\frac{n(n+1)}{2}}\sum_{(\alpha)}\sum_{j=1}^{n}\frac{\partial^2 a_{(\alpha)}(x)}{(\partial x^j)^2}\,dx^{\alpha_1}\wedge\ldots\wedge dx^{\alpha_k}. \qquad (1)$$

Evidemment, il suffit de considérer seules les formes monômes. Sans restreindre la généralité, une forme monôme ω peut être écrite comme suit:

$$\omega = a(x)\,dx^1\wedge\ldots\wedge dx^k. \qquad (2)$$

(On change, s'il le faut, la numération des coordonnées.)

Pour la forme (2) on a

$$*\omega = a(x)(-1)^{1+\ldots+k}\,dx^{k+1}\wedge\ldots\wedge dx^n,$$

$$\partial(*\omega)=(-1)^{1+\ldots+k}\sum_{j=1}^{k}\frac{\partial a(x)}{\partial x^j}\,dx^j\wedge dx^{k+1}\wedge\ldots\wedge dx^n,$$

$$*(\partial*\omega)=(-1)^{1+\ldots+k}\sum_{j=1}^{k}(-1)^{j+((k+1)+\ldots+n)}\frac{\partial a(x)}{\partial x^j}\times$$

$$\times\,dx^1\wedge\ldots\wedge\widehat{dx^j}\wedge\ldots\wedge dx^n,$$

où le signe $\widehat{dx^j}$ montre que le facteur dx^j est à omettre. Ensuite, $\partial\,\delta\omega=\partial*\partial*\omega=$

$$=(-1)^{\frac{n(n+1)}{2}}\sum_{j=1}^{k}(-1)^j\frac{\partial^2 a(x)}{(\partial x^j)^2}\,dx^1\wedge\ldots\wedge dx^j\wedge\ldots\wedge dx^k+$$

$$+(-1)^{\frac{n(n+1)}{2}}\sum_{j=1}^{k}(-1)^j\sum_{p=k+1}^{n}\frac{\partial^2 a(x)}{\partial x^j\,\partial x^p}\times$$

$$\times\,dx^1\wedge\ldots\wedge\widehat{dx^j}\wedge\ldots\wedge dx^k\wedge dx^p\,(-1)^{k-1}. \qquad (3)$$

D'une manière analogue,

$$\partial\omega=\sum_{p=k+1}^{n}\frac{\partial a(x)}{\partial x^p}(-1)^k\,dx^1\wedge\ldots\wedge dx^k\wedge dx^p,$$

$$*\partial\omega=\sum_{p=k+1}^{n}(-1)^k(-1)^{1+\ldots+k+p}\frac{\partial a(x)}{\partial x^p}\,dx^{k+1}\wedge\ldots$$

$$\ldots\wedge\widehat{dx^p}\wedge\ldots\wedge dx^n,$$

$$\partial * \partial \omega = (-1)^{1+\ldots+k} \sum_{p=k+1}^{n} (-1)^p \frac{\partial^2 a(x)}{(\partial x^p)^2} dx^{k+1} \wedge \ldots$$

$$\ldots \wedge dx^p \wedge \ldots \wedge dx^n + (-1)^{k+(1+\ldots+k)} \sum_{p=k+1}^{n} (-1)^p \times$$

$$\times \sum_{j=1}^{k} \frac{\partial^2 a(x)}{\partial x^j \partial x^p} dx^j \wedge dx^{k+1} \wedge \ldots \wedge \widehat{dx^p} \ldots \wedge dx^n,$$

$$\delta \partial \omega = * \partial * \partial \omega =$$

$$= -(-1)^{1+\ldots+k} \sum_{p=k+1}^{n} \frac{\partial^2 a(x)}{(\partial x^p)^2} (-1)^{(k+1)+\ldots+n} dx^1 \wedge \ldots \wedge dx^k +$$

$$+ (-1)^{k+(1+\ldots+k)} \sum_{p=k+1}^{n} \sum_{j=1}^{n} \frac{\partial^2 a(x)}{\partial x^j \partial x^p} (-1)^{j+(k+1)+\ldots+n} \times$$

$$\times dx^1 \wedge \ldots \wedge \widehat{dx^j} \wedge \ldots \wedge dx^k =$$

$$= -(-1)^{\frac{n(n+1)}{2}} \sum_{p=k+1}^{n} \frac{\partial^2 a(x)}{(\partial x^p)^2} dx^1 \wedge \ldots \wedge dx^k +$$

$$+ (-1)^{k+\frac{n(n+1)}{2}} \sum_{p=k+1}^{n} \sum_{j=1}^{k} (-1)^j \frac{\partial^2 a(x)}{\partial x^j \partial x^p} dx^1 \wedge \ldots \wedge \widehat{dx^j} \wedge \ldots \wedge dx^k. \tag{4}$$

En additionnant (3) et (4) les sommes doubles s'annulent, et l'on a

$$(\partial\delta + \delta\partial)\omega = -(-1)^{\frac{n(n+1)}{2}} \Delta\omega, \quad \Delta = \sum_{j=1}^{n} \frac{\partial^2}{(\partial x^j)^2}, \tag{5}$$

ce qu'il fallait.

b. Dans le cas général (espace riemannien et n'importe quel système de coordonnées), l'opérateur $\partial\delta + \delta\partial$ est appelé *opérateur de Laplace-Beltrami*. L'expression de l'opérateur de Laplace-Beltrami se compose de deux termes; dans le premier, on applique à la forme ω l'opérateur $\nabla^i\nabla_i$ (*6.47b, c*), dans le deuxième figure le tenseur de courbure de l'espace riemannien (cf. G. de Rham, Variétés différentiables, Paris, Hermann, 1955).

7.45. Construction d'une forme ω d'après $\partial\omega$ et $\delta\omega$ donnés.

a. Soient μ une $(k-1)$-forme et λ une $(k+1)$-forme données sur une variété riemannienne M_n; on se demande s'il existe une k-forme ω telle que $\delta\omega = \mu$, $\partial\omega = \lambda$.

Si une telle forme ω existe, alors $\delta\mu = \delta^2\omega = 0$ et $\partial\lambda = \partial^2\omega = 0$, donc les relations $\partial\mu = 0$, $\delta\lambda = 0$ représentent des conditions *nécessaires* de l'existence de la forme cherchée.

Il s'avère que ces conditions sont aussi bien suffisantes, au moins dans la supposition que les formes μ et λ soient *à support compact*, c'est-à-dire que leurs coefficients s'annulent en dehors d'un ensemble compact dans l'espace M_n.

Nous donnons dans *d* une simple démonstration de ce fait * pour le cas où $M_n = R_n$ est l'espace n-dimensionnel vectoriel. (Dans le cas général, la démonstration est beaucoup plus compliquée; cf., par exemple G. de R h a m, Variétés différentiables, Paris, Hermann, 1955).

b. Considérons d'abord un cas particulier où la forme est identiquement nulle.

Choisissons dans l'espace R_n une base orthonormale et posons

$$\mu(X) = \sum_{(\alpha)} \mu(x)\, dx^{\alpha_1} \wedge \ldots \wedge dx^{\alpha_{k-1}},$$

$$\varphi(x) = \int_{R_n} \frac{\mu(y)\, dy}{|x-y|^{n-2}} \equiv \sum_{(\alpha)} \left[\int_{R_n} \frac{\mu_{(\alpha)}(y)\, dy}{|x-y|^{n-2}} \right] dx^{\alpha_1} \wedge \ldots \wedge dx^{\alpha_{k-1}}.$$

Calculons $\partial\varphi$ et $\delta\varphi$. D'après *3.77g*, nous avons

$$\partial\varphi = \sum_{(\alpha)} \sum_{i=1}^{n} \frac{\partial}{\partial x^i} \left[\int_{R_n} \frac{\mu_{(\alpha)}(y)\, dy}{|x-y|^{n-2}} \right] dx^i \wedge dx^{\alpha_1} \wedge \ldots \wedge dx^{\alpha_{k-1}} =$$

$$= (n-2) \sum_{(\alpha)} \sum_{i=1}^{n} \left[\int_{R_n} \frac{y_i - x_i}{|x-y|^n} \mu_{(\alpha)}(y)\, dy \right] dx^i \wedge dx^{\alpha_1} \wedge \ldots \wedge dx^{\alpha_{k-1}}.$$

Désignons par $(\beta) = (\beta_1, \ldots, \beta_{n-k+1})$ le multinuméro complémentaire à (α) et par $\mu^*_{(\beta)}(y)$ la quantité $(-1)^{\alpha_1 + \ldots + \alpha_k} \mu_{(\alpha)}(y)$. Alors on a

$$*\varphi = \sum_{(\beta)} \left[\int_{R_n} \frac{\mu^*_{(\beta)}(y)\, dy}{|x-y|^{n-2}} \right] dx^{\beta_1} \wedge \ldots \wedge dx^{\beta_{n-k-1}},$$

$$\partial(*\varphi) = \sum_{(\beta)} \sum_{i=1}^{n} \frac{\partial}{\partial x^i} \left[\int_{R_n} \frac{\mu^*_{(\beta)}(y)\, dy}{|x-y|^{n-2}} \right] dx^i \wedge dx^{\beta_1} \wedge \ldots \wedge dx^{\beta_{n-k+1}} =$$

$$= \sum_{(\beta)} \sum_{i=1}^{n} \left[-\int_{R_n} \mu^*_{(\beta)}(y) \frac{\partial}{\partial y_i} \frac{1}{|x-y|^{n-2}}\, dy \right] dx^i \wedge dx^{\beta_1} \wedge \ldots \wedge dx^{\beta_{n-k+1}} =$$

* Proposée par I. Dorfman.

$$= \sum_{(\beta)} \Big[- \int_{R_n} \sum_{i=1}^{n} \frac{\partial}{\partial y^i} \Big(\mu^*_{(\beta)}(y) \frac{1}{|x-y|^{n-2}} \Big) dx^i\, dy \Big] \wedge dx^{\beta_1} \wedge \ldots \wedge dx^{\beta_{n-k+1}} +$$

$$+ \sum_{(\beta)} \Big[\int_{R_n} \frac{1}{|x-y|^{n-2}} \sum_{i=1}^{n} \frac{\partial \mu^*_{(\beta)}(y)}{\partial y^i} dx^i\, dy \Big] \wedge dx^{\beta_1} \wedge \ldots \wedge dx^{\beta_{n-k+1}}. \qquad (1)$$

Par hypothèse, les fonctions $\mu^*_{(\beta)}(y)$ s'annulent en dehors d'une boule S. En désignant par $\nu = (\nu_1, \ldots, \nu_n)$ la normale unitaire extérieure à la frontière de la boule et appliquant le théorème d'Ostrogradski, on trouve:

$$\int_{R_n} \sum_{i=1}^{n} \frac{\partial}{\partial y^i} \Big(\mu^*_{(\beta)}(y) \frac{1}{|x-y|^{n-2}} \Big) dx^i\, dy =$$

$$= \int_{\partial S} \mu^*_{(\beta)}(y) \frac{1}{|x-y|^{n-2}} \Big[\sum_{i=1}^{n} \nu_i\, dx^i \Big] dy = 0.$$

Or,

$$\sum_{i=1}^{n} \frac{\partial \mu^*_{(\beta)}(y)}{\partial y^i} dx^i \wedge dx^{\beta_1} \wedge \ldots \wedge dx^{\beta_{n-k+1}} = \partial (*\mu(x))|_{x=y} =$$

$$= (-1)^{\frac{n(n+1)}{2}} (**)\, \partial (*\mu) = (-1)^{\frac{n(n+1)}{2}} * \delta \mu = 0.$$

Ainsi, les deux intégrales dans (1) sont nulles. Par conséquent,

$$\delta \varphi = * \partial * \varphi = 0.$$

Appliquons ensuite la formule $\delta\partial + \partial\delta = -(-1)^{\frac{n(n+1)}{2}} \Delta$ (*7.44* (5)). La quantité $\Delta\varphi$ dans le présent cas peut être trouvée de *4.42* (9):

$$\Delta \varphi = \sum_{(\alpha)} \Big[\sum \frac{\partial^2}{(\partial x^i)^2} \int_{R_n} \frac{\mu(y)\, dy}{|x-y|^{n-2}} \Big] dx^{\alpha_1} \wedge \ldots \wedge dx^{\alpha_{k-1}} =$$

$$= S_n (2-n) \sum_{(\alpha)} \mu_{(\alpha)}(x)\, dx^{\alpha_1} \wedge \ldots \wedge dx^{\alpha_{k-1}} = S_n (2-n)\, \mu$$

où S_n est l'aire de la sphère unité dans R_n. Donc, $\delta\partial\varphi = -(-1)^{\frac{n(n+1)}{2}} \Delta\varphi = (n-2) S_n (-1)^{\frac{n(n+1)}{2}} \mu$. En posant $\omega = \frac{(-1)^{\frac{n(n+1)}{2}}}{(n-2) S_n} \partial\varphi$ on trouve $\delta\omega = \mu$, $\partial\omega = 0$, ce qu'il fallait.

c. Envisageons un autre cas particulier: il y a une $(k+1)$-forme λ avec $\partial\lambda = 0$ et une $(k-1)$-forme $\mu \not\equiv 0$; on se propose de trouver

une k-forme ω telle que $\partial\omega = \lambda$, $\delta\omega = 0$. La forme $*\lambda$ est du degré $n - k - 1$; pour elle, nous avons

$$\delta(*\lambda) = *\partial**\lambda = (-1)^{\frac{n(n+1)}{2}} *\partial\lambda = 0.$$

D'après ce qu'on a démontré dans b, il existe une $(n - k)$-forme g pour laquelle

$$\delta g = *\lambda, \qquad \partial g = 0.$$

Montrons que la k-forme $\omega = *g$ possède les propriétés voulues. En effet :

$$\partial\omega = *\delta**g = (-1)^{\frac{n(n+1)}{2}} *\delta g = \lambda,$$

$$\delta\omega = *\partial**g = (-1)^{\frac{n(n+1)}{2}} *\partial g = 0,$$

ce qu'il fallait.

d. Dans le cas général, s'il y a une $(k - 1)$-forme μ à support compact, avec $\delta\mu = 0$, et une $(k+1)$-forme λ à support compact, avec $\delta\lambda = 0$, trouvons d'abord deux k-formes ω_1 et ω_2 pour lesquelles

$$\delta\omega_1 = \mu, \qquad \partial\omega_1 = 0,$$

$$\delta\omega_2 = 0, \qquad \partial\omega_2 = \lambda,$$

et additionnons-les. La forme obtenue

$$\omega = \omega_1 + \omega_2 \tag{2}$$

résout le problème posé car

$$\partial\omega = \partial\omega_2 = \lambda, \qquad \delta\omega = \delta\omega_1 = \mu.$$

e. Etudions le problème d'unicité de solution. Supposons que, pour deux formes données μ et λ, il existe deux k-formes ω et Ω telles que

$$\partial\omega = \partial\Omega = \lambda, \qquad \delta\omega = \delta\Omega = \mu.$$

Alors la forme $\theta = \Omega - \omega$ vérifie les relations

$$\partial\theta = 0, \qquad \delta\theta = 0. \tag{3}$$

Les formes θ possédant les propriétés (3) sont appelées formes *harmoniques*. Si une forme θ est harmonique, alors $\Delta\theta = c(\partial\delta + \delta\partial)\,\theta = 0$ et tous les coefficients d'une forme harmonique sont des fonctions harmoniques. (La réciproque est, en général, fausse : on ne peut choisir pour coefficients d'une forme harmonique n'importe quelles fonctions harmoniques, car les coefficients sont liés par les égalités découlant de (3).)

La solution (2) que nous avons trouvée a, évidemment, les coefficients tendant vers zéro lorsque $|x| \to \infty$. Si une autre solution Ω possède la même propriété, alors il en est de même de la forme harmonique $\theta = \Omega - \omega$. Or, toute fonction harmonique tendant vers 0 lorsque $|x| \to \infty$ est identiquement nulle (*4.54c*). Donc, *dans la classe de toutes les k-formes dont les coefficients tendent vers* 0 lorsque $|x| \to \infty$, *la solution trouvée est unique.*

f. Le résultat obtenu peut être mis sous une autre forme:

Théorème. *Toute k-forme* Ω, $\partial\Omega = \lambda$ *et* $\delta\Omega = \mu$ *étant à support compact, peut être représentée comme somme* $f + g$ *de deux k-formes* f *et* g, *de sorte que* $\partial f = 0$, $\delta g = 0$. *Cette représentation est unique dans la classe de toutes les k-formes dont les coefficients tendent vers* 0 *lorsque* $|x| \to \infty$.

Pour le démontrer, remarquons que, comme $\partial\lambda = 0$ et $\delta\mu = 0$, on peut construire une telle forme $\omega = \omega_1 + \omega_2$ que

$$\delta\omega_1 = \mu, \qquad \partial\omega_1 = 0,$$

$$\delta\omega_2 = 0, \qquad \partial\omega_2 = \lambda.$$

En vertu de l'unicité déjà démontrée, les formes ω et Ω coïncident, ce qui fournit la représentation cherchée $\Omega = \omega_1 + \omega_2$. Elle est unique dans la classe mentionnée d'après les mêmes considérations que plus haut.

g. Le problème que nous sommes en train de traiter représente une généralisation directe du problème de construction d'un champ vectoriel d'après ses rotationnel et divergence. En effet, dans l'espace tridimensionnel euclidien, les opérations classiques de l'analyse vectorielle grad φ, div f et rot f s'expriment, ce qu'il est aisé de prouver, par ∂, δ et $*$ (au signe près) comme suit:

$$\operatorname{grad}\varphi = \partial\varphi, \quad \operatorname{div} f = *\,\partial * f \equiv \delta f, \quad \operatorname{rot} f = *\,\partial f.$$

Soient maintenant $u(x)$ $(U \in R_3 \to R_1)$ et $R(x)$ $(U \in R_3 \to R_3)$ des fonctions (à support compact), $\operatorname{div} R(x) \equiv 0$. On se propose de construire un champ vectoriel f tel que $\operatorname{div} f = u$, $\operatorname{rot} f = R$. On peut traiter le champ scalaire donné u comme 0-forme, donc, comme pour toute 0-forme, on a $\delta u = 0$. Le champ vectoriel donné R peut être traité comme 1-forme avec $\delta R = 0$ ou bien, ce qui revient au même, avec $\partial(*R) = 0$. Le problème posé est alors équivalent à la recherche d'un champ vectoriel ou d'une 1-forme f pour laquelle $\delta f = \varphi$, $*(\partial f) = R$ ou $\partial f = *R$. Or, c'est juste le problème considéré dans *a*, sous une forme particulière, donc un problème résoluble (en vertu de *d* pour $u(x)$ et $R(x)$ à support compact; en vertu de *e*, sa solution est unique dans la classe de tous les champs vectoriels $f(x)$ qui tendent vers zéro lorsque $|x| \to \infty$).

EXERCICES

1. Les formes $E_{(\alpha)}(\underset{1}{x}, \ldots, \underset{k}{x}) = \sum_{O(i)=(\alpha)} \xi^{i_1}_{1}, \ldots \xi^{i_k}_{k}$ sont symétriques; elles s'appellent *k-formes symétriques élémentaires*. Montrer que toute k-forme symétrique peut être représentée d'une façon unique comme combinaison linéaire de formes $E_{(\alpha)}$.

2. Quelle est la dimension de l'espace des k-formes symétriques dans R_n?

3. Montrer qu'il est possible de représenter toute forme bilinéaire $A(\underset{1}{x}, \underset{2}{x})$ comme somme d'une forme bilinéaire symétrique et d'une forme bilinéaire antisymétrique.

4. Pour $k > 2$ montrer qu'il existe une k-forme qu'on peut représenter comme somme des k-formes symétrique et antisymétrique.

5. Posons, pour n'importe quelle k-forme $A(\underset{1}{x}, \ldots, \underset{k}{x})$,

$$\operatorname{Sym} A(\underset{1}{x}, \ldots, \underset{k}{x}) = \frac{1}{k!} \sum_{(i)} A(\underset{i_1}{x}, \ldots, \underset{i_k}{x}).$$

Montrer que $\operatorname{Sym} A(\underset{1}{x}, \ldots, \underset{k}{x})$ est une forme symétrique; si $A(\underset{1}{x}, \ldots, \underset{k}{x})$ est symétrique, alors $\operatorname{Sym} A = A$, et réciproquement.

6. Un $(k\text{-}n)$-numéro ordonné (α) peut toujours être mis sous la forme

$$(\alpha) = (\underbrace{1, \ldots, 1}_{p_1 \text{ fois}}, \underbrace{2 \ldots, 2}_{p_2 \text{ fois}}, \ldots, \underbrace{n, \ldots, n}_{p_n}),$$

où les nombres $p_1, \ldots, p_n$ sont compris dans les limites $0 \leqslant p_i \leqslant k$ et $\sum p_i = k$. L'ensemble $(p_1, \ldots, p_k)$ s'appelle caractéristique du multinuméro (α). Montrer que

$$\operatorname{Sym} \xi^{i_1}_{1} \ldots \xi^{i_k}_{k} = \operatorname{Sym} \xi^{\alpha_1}_{1} \ldots \xi^{\alpha_k}_{k} = \frac{p_1! \ldots p_n!}{k!} E_{(\alpha)}(\underset{1}{x}, \ldots, \underset{k}{x}),$$

où $(\alpha) = O(i)$, $(p_1, \ldots, p_n)$ étant la caractéristique du multinuméro (α), $E_{(\alpha)}$ une forme symétrique élémentaire (exercice 1).

7. On appelle produit tensoriel symétrique d'une k-forme A et d'une m-forme B la $(k + m)$-forme

$$A \vee B = \operatorname{Sym}(A \times B).$$

Montrer que l'opération $\vee$ est distributive avec l'addition et associative.

8. Montrer que

$$\xi^{i_1} \vee \ldots \vee \xi^{i_k} = \frac{p_1! \ldots p_k!}{k!} E_{(\alpha)}(\underset{1}{x}, \ldots, \underset{k}{x}) = \operatorname{Sym} \xi^{i_1}_{1} \ldots \xi^{i_k}_{k}.$$

9. Montrer que toute $(n - 1)$-forme antisymétrique dans R_n peut être représentée comme suit:

$$A(\underset{1}{x}, \ldots, \underset{n-1}{x}) = \begin{vmatrix} \xi^1_1 & \ldots & \xi^1_{n-1} & a^1 \\ \cdot & \cdot & \cdot & \cdot \\ \xi^n_1 & \ldots & \xi^n_{n-1} & a^n \end{vmatrix},$$

où $a^1, \ldots, a^n$ sont des nombres fixes (pour la base donnée).

10. Vérifier la formule

$$A \wedge B = (-1)^k B \wedge A,$$

où k est le degré de la forme A.

11. Soient $P = \| a_{ij} \|$ une $n \times n$-matrice inversible et P^{-1} sa matrice inverse. Choisissons dans la matrice P le mineur M occupant les lignes et les colonnes de numéros $1, \ldots, k$, et dans la matrice P^{-1} le mineur T occupant les lignes et les colonnes de numéros $k+1, \ldots, n$. Démontrer que $T = M/\det P$.

12. Montrer que les coefficients $\widetilde{a}_{(\alpha)}$ de la deuxième notation canonique d'une forme antisymétrique (*7.16a*) se transforment, en effectuant un changement de coordonnées $\xi^i = p_{i'}^i \xi^{i'}$, selon la formule

$$(-1)^{\sum \alpha'_i} \widetilde{a}_{(\alpha')} = \frac{(-1)^{\sum \alpha_i}}{\det P} \sum_{(\alpha)} \widetilde{a}_{(\alpha)} \det \| p_{\beta_i}^{\beta'_i} \|,$$

où (β_i) et (β'_i) sont les multinuméros strictement ordonnés, complémentaires aux multinuméros (α) et (α') respectivement.

En particulier, les coefficients b^i d'une $(n-1)$-forme

$$A(\underset{1}{x}, \ldots, \underset{n-1}{x}) = \sum_i (-1)^{i-1} b^i \xi^1 \wedge \ldots \wedge \hat{\xi}^i \wedge \ldots \wedge \xi^n$$

se transforment comme les tenseurs une fois contravariants

$$b^{i'} = \frac{1}{\det P} \sum_i b^i p_i^{i'}$$

avec le « poids » $1/\det P$.

13. Pour une forme différentielle quelconque

$$\omega = \sum_{(i)} f_{(i)}(x)\, dx^{i_1}_{1} \ldots dx^{i_k}_{k},$$

on peut définir trois différentielles :

$$D\omega = \sum_{(i)} \sum_j \frac{\partial f_{(i)}(x)}{\partial x^j} dx^j_{1}\, dx^{i_1}_{2} \ldots dx^{i_k}_{k+1},$$

$$d\omega = \operatorname{Sym} D\omega, \qquad \partial\omega = \operatorname{Alt} D\omega.$$

Montrer que, pour une forme ω antisymétrique, la différentielle ∂ coïncide avec la différentielle définie dans *7.22*.

14. Montrer que

$$d\left(\sum_{(i)} f_{(i)}(x)\, dx^{i_1} \vee \ldots \vee dx^{i_k} \right) = \sum_{(i)} \sum_j \frac{\partial f_{(i)}(x)}{\partial x^j} dx^j \vee dx^{i_1} \vee \ldots \vee dx^{i_k}.$$

15. Montrer que

$$d(\omega_1 \times \omega_2) = d\omega_1 \vee \omega_2 + \omega_1 \vee d\omega_2.$$

16. Soit $A = \sum_{(\alpha)} \widetilde{a}_{(\alpha)} \xi^{\alpha_1} \wedge \ldots \wedge \xi^{\alpha_k}$. Posons $\widetilde{b}_{(\gamma)} = (-1)^{\sum \alpha_i} \widetilde{a}_{(\alpha)}$

dans toute base orthonormale, (γ) étant le multinuméro strictement ordonné,

complémentaire à (α). Montrer par le calcul direct que la forme $\sum_{(\nu)} b_{(\nu)}\, \xi^{\nu_1} \wedge \ldots$ $\ldots \wedge \xi^{\nu_{n-k}}$ est invariante par rapport aux transformations orthogonales des coordonnées.

17. Pour chaque k-forme $\omega \neq 0$ sur une variété différentiable élémentaire M_n, indiquer une k-chaîne c telle que $\int_c \omega \neq 0$.

18. Montrer que, pour toute k-chaîne c avec $\partial c = 0$ et pour toute k-forme ω avec $\partial\omega = 0$ sur une variété M_n, a lieu l'égalité

$$\int_c \omega = 0.$$

Historique

Une généralisation multidimensionnelle du théorème de Stokes est obtenue, pour la première fois, par Poincaré dans le III-e volume de ses « Méthodes nouvelles de la mécanique céleste » (1889). Ensuite, un autre énoncé de ce théorème, qui est déjà bien proche des énoncés modernes, est donné par E. Cartan (1899), le créateur du calcul des formes antisymétriques différentielles; la partie algébrique de ce calcul remonte à « Die Ansdehrungslehre » de Grassmann (1844, 1861) où, entre autres, apparaît pour la première fois « l'espace n-dimensionnel ». La notion de codifférentielle d'une forme est introduite par Brauer pour l'espace euclidien (1906) et par Weitzenböck pour l'espace riemannien (1923). Le problème de construction d'une forme différentielle d'après sa différentielle et sa codifférentielle est posé et résolu (pour l'espace riemannien) par G. de Rham dans son livre « Variétés différentiables » (1955).

INDICATIONS ET RÉPONSES

Chapitre 1

1. *Indication.* Considérer la suite de points

$$\varphi_n = 1/n, \quad r_n = 1/n.$$

2. *Indication.* De toute suite $\{\varphi_n, \rho_n\} \to 0$ $(0 \leqslant \varphi_n \leqslant 2\pi)$, on peut extraire une sous-suite partielle $\{\varphi_{n_k}, \rho_{n_k}\}$ telle que les nombres φ_{n_k} ont une limite.

3. *Indication.* a) y est linéaire sur toute demi-droite; b) cette fonction linéaire ne doit pas changer de pente suivant la droite; c) si elle est dérivable, elle doit coïncider avec sa partie linéaire principale.

4. *Indication.* La matrice de multiplication par un nombre complexe $a + rb$ a la forme $\begin{Vmatrix} a & -b \\ b & a \end{Vmatrix}$, celle de l'opérateur $f'(z)$, pour $w = u + iv$, a la forme $\begin{Vmatrix} u_x & u_y \\ v_x & v_y \end{Vmatrix}$. Se rappeler les conditions de Cauchy-Riemann (*U10.16*).

5. *Réponse.*

$$\frac{dx_2}{dx_1} = \frac{\dfrac{\partial f}{\partial x_1}}{\dfrac{\partial f}{\partial x_2}}.$$

6. *Réponse.* L'équation

$$[(x - a)^2 + y^2]\,[(x + a)^2 + y^2] = b^4$$

représente deux courbes séparées pour $b < a$, une lemniscate pour $b = a$, une courbe à quatre points d'inflexion pour $a < b < a\sqrt{2}$, une courbe convexe pour $b \geqslant a\sqrt{2}$.

7. *Réponse.* Par exemple,

$$x_1 = \cos\varphi, \quad x_2 = \sin\varphi, \quad 0 \leqslant \varphi \leqslant 2\pi.$$

8. *Indication.* Appliquer la formule de Newton-Leibniz.

9. *Indication.* Utiliser trois faits suivants: a) le gradient de la fonction est un vecteur unitaire; b) les diagonales d'un losange sont perpendiculaires l'une à l'autre; c) le gradient est orthogonal à une surface de niveau.

10. *Indication.* Supposons que la fonction $F(t, \xi)$ $(R_2 \to R_1)$ ait une dérivée continue en ξ. Trouver des points stationnaires pour la fonctionnelle

$$f(x) = \int_a^b F(t,\ x(t))\,dt$$

définie dans l'espace $R(a, b)$ des fonctions réelles continues $x = x(t)$, $a \leqslant t \leqslant b$.

11. *Réponse.* $x_1(t) = -\frac{1}{3}(t+1)$ est un maximum local ; $x_2(t) = \frac{1}{3}(t+1)$ est un minimum local.

12. *Indication.* Minimum local pour $x_0(t) \equiv 1$. Afin de le démontrer, poser $x(t) = 1 + \varepsilon(t)$ et, compte tenu de l'hypothèse, montrer que

$$f(x) - f(x_0) = \frac{3}{2}\int_0^1 \varepsilon^2(t)\,dt + \int_0^1 \varepsilon^3(t)\,dt.$$

13. *Indication.* Dans les conditions imposées, l'opérateur

$$Ax = x - [f'(a)]^{-1} f(a)$$

est contractant ; cf. la démonstration du théorème *1.53*.

14. *Indication.* Considérer cette fonction sur une droite passant par l'origine.

15. *Indication.* Dans l'espace hilbertien on a $|x| = \sqrt{(x, x)}$.

16. *Indication.* Choisir le déplacement de la forme $(0, \ldots, 0, h, 0, \ldots)$, $|h| < \delta$. Utiliser la forme générale d'une fonctionnelle linéaire dans l_1.

17. *Indication.* Remplacer y par $A_0 x^{r_0}$ dans l'équation (1) et montrer que l'équation obtenue

$$\varphi(x, u) \equiv f(u, A_0 x^{r_0} u) = 0$$

est vérifiée par les valeurs $x = 0$, $u = 1$, et que

$$\frac{\partial \varphi(0, 1)}{\partial u} \neq 0 ;$$

puis appliquer le théorème sur la fonction implicite.

18. *Indication.* Remplacer y par $A_0 x^{r_0}(1 + B_0 x^{s_0} u)$ dans l'équation (1); puis agir comme dans l'exercice 17.

19. *Indication.* Utiliser *1.34d* et *1.45*.

20. *Indication.* Toutes les fonctions $f(x)$ avec une même valeur $f'(a)$ appartiennent à une même classe.

21. *Indication.* Si une fonction $f(x) \in \mathfrak{A}$ est nulle dans un voisinage V du point x, et qu'une fonction $h(x) \in \mathfrak{A}$ est nulle en dehors de V, de façon que $h(a) = 1$, on aboutit au résultat cherché en appliquant la fonctionnelle D à $f(x) \cdot h(x)$. Montrer que toute fonction $f(x) \in J(a)$ est la limite sur la norme de $\mathfrak{A}$ de fonctions s'annulant au voisinage du point a.

22. *Indication.* D'après l'exercice 20, on a

$$Df = (f'(a), y) = \frac{\partial f(a)}{\partial y}.$$

23. *Réponse.* Par exemple, dans un espace hilbertien H avec une base orthonormale $e_1, \ldots, e_n, \ldots$, soit $S_n = \{x \in H : |x - \frac{2e_n}{n}| \leqslant \frac{1}{n}\}$, et considérons une fonction $\theta_n(t) : R_1 \to R_1$ dérivable, comprise entre 0 et 1, valant 0 pour $|t - \frac{2}{n}| \geqslant \frac{1}{n}$ et 1 pour $t = \frac{2}{n}$. Pour $x \in S_n$, l'opérateur linéaire $f(x)$ est donné par la formule $f(x)\,h = \frac{1}{n} e_n (h, e_n)\,\theta_n(|x - \frac{e_n}{n}|)$; pour $x \in \bigcup_{n=1}^{\infty} S_n$, on pose $f(x)\,h \equiv 0$.

24. *Réponse.* Par exemple, sur toute demi-droite d'argument φ, on peut définir la fonction $f(x, y)$ par un graphique du type montré fig. 1. Alors dans le plan $\{x, y\}$ la fonction $f(x, y)$ a, en l'origine des coordonnées, la dérivée non nulle suivant la courbe dérivable qui est formée par les points de maximum des graphiques mentionnés.



Fig. 1

25. *Indication.* Utiliser l idée de construction de l'exercice 24; définir la fonction cherchée sur la demi-droite correspondant au vecteur de base e_n par un graphique du type montré fig. 1.

26. *Indication.* Utiliser *1.34c*.

27. *Réponse.* Puisque le rang de la matrice de Jacobi des fonctions données n'est constant en aucun voisinage du point (0, 0), elles ne sont pas dépendantes ni indépendantes.

28. *Indication.* Vérifier directement que si $z = \varphi(y)$ est une solution de l'équation (1), alors $w = z - \varphi(y)$ est une solution de l'équation (2). Inversement, si $w(y, z)$ est une solution de l'équation (2), alors pour $w_z \neq 0$ (cas général), l'égalité $w(y, z) = 0$ s'écrit comme $z = \varphi(y)$. Ecrire la dérivée et multiplier par Φ.

29. *Indication.* Soit $d_n = \inf_{m \neq n} \rho(x_n, x_m)$; prouver que $d_n > 0$ et poser $\rho_n = d_n/2$.

30. *Indication.* Dans la boule $S(x_n, \rho_n)$ poser $f(x) = \varphi_n(\rho(x, x_n))$ où $\varphi_n(\rho)$ est continue, nulle pour $\rho \geqslant p_n$ et égale à n pour $\rho = 0$.

31. *Indication.* La même idée que dans la solution de l'exercice 30.

Chapitre 2

1. *Indication.* Dans l'espace correspondant de dimension paire, le produit des racines caractéristiques de l'opérateur est négatif, donc il existe des racines positives ainsi que négatives; or, ce sont les coefficients canoniques de la forme.

2. *Réponse.*

$$d^2x^{-1} = x^{-1}hx^{-1}kx^{-1} + x^{-1}kx^{-1}hx^{-1}.$$

3. *Réponse.*

$$\varphi''(y)\,pq = -[f'(x)]^{-1} \cdot f''(x) \cdot [f'(x)]^{-1} \cdot p \cdot [f'(x)]^{-1}\,q,\ p,\ q \in Y.$$

4. *Indication.* La première partie résulte du théorème sur l'extrémum lié. Pour construire un exemple, considérer la fonction $y = x_2$ $(R_2 \to R_1)$ sous les conditions $x_2 = Cx_1^2$ avec C différents.

5. *Indication.* Appliquer le critère *2.63a*.

6. *Indication.* Les fonctions $a_1(x)$ et $a_2(x)$ ne sont pas continues.

7. *Réponse.*

$$\left(\frac{\partial \Phi}{\partial x_1} + \frac{\partial \Phi}{\partial y}\Phi\right) h_1 k_1 = \left(\frac{\partial \Phi}{\partial x_1} + \frac{\partial \Phi}{\partial y}\Phi\right) k_1 h_1$$

$$(h_1,\ k_1 \in X_1).$$

8. *Indication.* Dériver la relation $(y'h,\ y'k) = \lambda^2(h,\ k)$ par rapport à n'importe quel vecteur l, puis effectuer le changement cyclique des variables h, k, l. Le vecteur l étant arbitraire, obtenir la relation $y''hk = 0$.

9. *Indication.* Commencer par l'inversion; utiliser l'exercice 26 du chapitre 1.

10. *Indication.* Dériver la relation $(y'h,\ y'k) = 0$ par rapport à l, puis faire le changement cyclique de h, k, l.

11. *Indication.* Utiliser le fait que le vecteur l de l'exercice 10 peut être un vecteur arbitraire orthogonal à h et k.

Réponse.

$$\mu = \frac{\lambda'(x)\,k}{\lambda(x)}, \qquad \nu = \frac{\lambda'(x)\,h}{\lambda(x)}.$$

12. *Indication.* Dériver l'égalité de l'exercice 11 par rapport à l arbitraire en tenant compte des valeurs trouvées de μ et ν. Se convaincre que la fonction vectorielle $\rho''lk \cdot y'h$ est symétrique par rapport à l et h. Utiliser l'indépendance de $y'h$ et $y'l$.

13. *Indication.* Prouver que $\rho''hk = \sigma(x)\,(h,\ k)$, puis dériver par rapport à l arbitraire et utiliser la symétrie de la dérivée troisième.

14. *Indication.* Intégrer le résultat de l'exercice 13.

15. *Indication.* Appliquer le résultat de l'exercice 14 à la fonction $y(x)$ et à son inverse.

16. *Indication.* Pour $\alpha \neq 0$, $\beta \neq 0$, la fonction

$$|y(x)| = \int_0^{|x|} \frac{dt}{\alpha t^2 + \beta}$$

serait une fonction transcendente, l'exercice 15 montre qu'elle est une fonction algébrique.

17. *Réponse.* Le théorème *2.61a* fournit une condition nécessaire de compatibilité d'un système, quelles que soient les conditions initiales (au voisinage du point donné). Le système en question n'a pas de solution avec la condition initiale $z(0,\ 0) = \varepsilon > 0$.

18. *Indication.* Appliquer le théorème *2.61a*.

Chapitre 3

1. *Réponse.* $I = \dfrac{1}{2^n n!} \displaystyle\int_0^x (x^2 - u^2)^n f(u)\,du.$

2. *Réponse.* $I = \dfrac{\pi^{n/2}}{\sqrt{\det \|A\|}}.$

3. *Réponse.* $I = S_{n-1}(1) \displaystyle\int_0^\infty r^{n-1} f(r)\,dr.$

4. *Réponse.* $I = S_{n-2}(1) \displaystyle\int_{-1}^{1} f(h)\sqrt{1-h^2}\,dh.$

5. *Indication.* $|\Pi_{h\alpha}|/2\pi = 1 + C\,\dfrac{\alpha^5}{h^2} + \ldots$

6. *Indication.* Il est possible de choisir une telle suite exhaustive de domaines que les valeurs positives de la fonction $f(x)$ soient dominantes, on peut faire de même pour les valeurs négatives.

Sur $[0,\ \infty)$, on a une autre définition de la convergence de l'intégrale (le choix de domaines admissibles est moins riche).

7. *Réponse.* Par exemple, l'ensemble P est une cellule non jordanienne.

8. *Indication.* Le segment limite n'est pas homogène.

9. *Indication.* Il suffit d'envisager le cas $k = 1$; appliquer le principe de Cavalieri et la relation (*U12.47b*)

$$\lim_{n\to\infty} \frac{\int_{\varepsilon}^{\pi/2} \cos^{2n} t\, dt}{\int_{0}^{\pi/2} \cos^{2n} t\, dt} = 0.$$

10. *Indication.* Idem.

11. *Indication.* Utiliser le principe de Cavalieri et la récurrence sur n.

12. *Indication.* Approcher les fonctions $f_1(x), \ldots, f_n(x)$ par des fonctions constantes par morceaux. Appliquer *3.55c*.

13. *Indication.* Intégrer d'abord sur la sphère de rayon r, puis par rapport à r de 0 à ∞. Pour la fonction $f(x) \equiv g(r)$ symétrique sphérique, obtenir la formule

$$\varphi(\sigma) = \int_0^{\infty} \left\{ \int_{S_r} e^{-i|\sigma| r\cos\theta}\, d\omega \right\} g(r)\, r^{n-1}\, dr.$$

14. *Réponse.*

$$\varphi(\sigma) = \frac{4\pi}{\rho} \int_0^{\infty} rg(r) \sin \rho r\, dr, \qquad \rho = |\sigma|.$$

15. *Indication.* Appliquer la règle de dérivation d'une intégrale impropre par rapport au paramètre (*3.74f*).

16. *Indication.* Dans la nouvelle base, écrire la forme donnée comme somme de carrés. Utiliser la méthode *U15.32*.

Chapitre 4

1. *Réponse.*

$$\text{rot } P(A) = [\rho\varphi'(\rho) + 2\varphi(\rho)]\, g,$$

où $\varphi(\rho) = \dfrac{f(\rho)}{\rho}$, g est le vecteur unitaire colinéaire à la droite λ. L'égalité rot $P(A) = 0$ a lieu pour $f(\rho) = C/\rho$.

2. *Indication.* Pour la courbe $y = \varphi(x)$, la courbure est donnée par la formule

$$k = \frac{\varphi''(x)}{[1 + \varphi'(x)]^{3/2}}.$$

Utiliser la formule *4.15* (2).

3. *Indication.* Si l'équation de la famille de surfaces a la forme $\varphi(M) = C$, l'hypothèse peut être écrite comme suit :

$$\text{grad } \varphi = \lambda R.$$

Pour éliminer λ et φ, appliquer l'opération rot.

4. *Indication.* Considérer sur la surface S le contour fermé composé de la ligne L, de deux courbes $\gamma(M)$ et d'un arc arbitraire et appliquer le théorème de Stokes.

5. *Indication.* Utiliser l'exercice 4.

6. *Indication.* Le domaine n'est pas simplement connexe; le potentiel local arctg (y/x) ne se prolonge pas dans tout le domaine en tant que fonction univoque.

7. *Indication.* Appliquer au domaine en question la formule de Stokes.

8. *Réponse.*

$$F(y)=\begin{cases} -\dfrac{e(0,\,y)}{|\,y\,|^2} & \text{pour } |\,y\,|\geqslant r, \\ -\dfrac{e(0,\,y)\,|\,y\,|}{r^2} & \text{pour } |\,y\,|\leqslant r. \end{cases}$$

9. *Indication.* Utiliser l'inégalité

$$(r-|\,y\,|)^n\leqslant|\,x-y\,|^n\leqslant(r+|\,y\,|)^n.$$

10. *Indication.* Passer, dans l'inégalité de l'exercice 9, à la limite pour $r\to\infty$.

11. *Indication.* Dans la formule de Poisson, passer au gradient et évaluer grad P $(x,\,y)$ pour $y=z$.

12. *Indication.* Appliquer le résultat de l'exercice 11 aux boules intérieures d'un rayon fixe, puis utiliser le théorème sur le recouvrement fini.

13. *Indication.* Vu le résultat de l'exercice 12, appliquer le théorème d'Arzelà.

14. *Indication.* Utiliser l'inégalité de l'exercice 9 et le résultat de l'exercice 11.

15. *Réponse.* Non, puisqu'il n'y a pas sur cette surface de vecteur normal continu.

Chapitre 5

1. *Réponse.*

$$K=-\frac{z_{xx}z_{yy}-z_{xy}^2}{(1+z_x^2+z_y^2)}.$$

2. *Réponse.*

$$K=-\frac{1}{(F_x^2+F_y^2+F_z^2)}\begin{vmatrix} F_{xx} & F_{xy} & F_{xz} & F_x \\ F_{yx} & F_{yy} & F_{yz} & F_y \\ F_{zx} & F_{zy} & F_{zz} & F_z \\ F_x & F_y & F_z & 0 \end{vmatrix}.$$

3. *Réponse.*

$$K=-\frac{1}{2\varphi(u,\,v)}\left(\frac{\partial^2\ln\varphi(u,\,v)}{\partial u^2}+\frac{\partial^2\ln\varphi(u,\,v)}{\partial v^2}\right).$$

4. *Indication.* Utiliser la formule *5.26* (5).

5. *Réponse.*

$$k_{\text{par}}=\frac{1}{\rho\sqrt{1+z_\rho^2}},\qquad k_{\text{mér}}=0.$$

6. *Réponse.* La géodésique s'enroule sur la caténoïde en s'approchant indéfiniment de sa ligne de striction (fig. 2).

7. *Réponse.* Pour une fonction dérivable $z(\rho)$: les parallèles sur lesquelles la fonction $z(\rho)$ atteint son maximum non nul, et seulement ces parallèles.

8. *Indication.* La coordonnée du centre est:

$$v = -\frac{(R_u, l_u)}{|l_u|^2}.$$

9. *Indication.* Dans le cas en question, l'équation de la ligne de striction a la forme:

$$|l_u|^2 (R_{uu}, l_u) + |l_u|^2 (R_u, l_{uu}) - 2 (R_u, l_u) (l_u, l_{uu}) = 0.$$

10. *Indication.* Montrer que, dans les hypothèses formulées, la ligne de striction coupe les génératrices orthogonalement. Choisir cette ligne pour directrice et montrer qu'elle est droite.

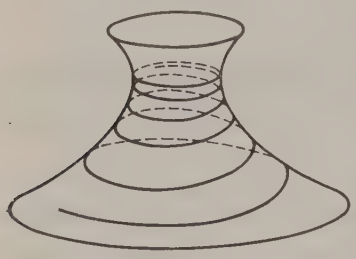


Fig. 2

11. *Indication.* Intégrer l'équation

$$dm = \lambda\, dr$$

le long de la ligne de courbure donnée.

12. *Indication.* A tout point B de l'espace tangent Π (A) faire correspondre un point C sur la variété d'après la règle: mener, à partir du point de tangence A, dans le sens du vecteur AB, une géodésique et définir sur elle un paramètre canonique, de sorte que la dérivée par rapport à lui au point A coïncide avec le vecteur AB; choisir le point correspondant à la valeur $r = 1$. (Cette application est appelée exponentielle.) Montrer que sa dérivée est non dégénérée.

13. *Indication.* Utiliser le résultat de l'exercice 12, choisir un voisinage sphérique assez petit du point A.

14. *Indication.* Composer la forme ds^2. Si r (u) et ρ (v) décrivent, par exemple, des circonférences, il n'y a pas de segments de droite sur la surface S.

15. *Indication.* Le lieu géométrique des points d'intersection de la surface donnée π_2 avec n'importe quel plan bidimensionnel normal est formé des points isolés.

16. *Indication.* Sans nuire à la généralité, on peut admettre que les vecteurs de base du plan tangent coïncident avec l premiers vecteurs de base de l'espace R_n. Alors chaque vecteur unitaire normal a la forme

$$m = (0, \ldots, 0, \omega_{l+1}, \ldots, \omega_n),$$

où

$$\omega_{l+1}^2 + \ldots + \omega_n^2 = 1.$$

En appliquant le théorème de la fonction implicite exprimer les coordonnées des points de la section normale complète comme fonctions d'un paramètre.

17. *Indication.* Appliquer la méthode de *5.31.*

18. *Indication.* Faire le calcul analogue à celui de *5.32.*

19. *Indication.* Idem.

20. *Indication.* Procéder par analogie avec la démonstration du théorème de Bonnet *5.34.*

21. *Indication.* Il y a un point singulier sur le cône.

Chapitre 6

1. *Indication.* Utiliser les coordonnées sphériques.

2. *Indication.* Utiliser la définition des dérivées covariantes.

3. *Indication.* Idem.

4. *Indication.* Pour cette définition de la connexion, a lieu le parallélisme absolu ; mais la construction *6.45* peut conduire à un écart essentiel des points C et E.

5. *Indication.* Utiliser l'équation différentielle de la translation.

6. *Indication.* L'application correspondante de l'espace tangent sur la variété transforme la dérivation ordinaire au point A en la dérivation covariante.

Chapitre 7

1. *Indication.* Appliquer l'égalité $\sum\limits_{(i)} = \sum\limits_{(\alpha)} \sum\limits_{O\,(i)=\alpha}$.

2. *Réponse.* C^k_{n+k-1}.

3. *Indication.* Considérer $A\underset{1}{(x,}\ \underset{2}{x)} + A\underset{2}{(x,}\ \underset{1}{x)}$ et $A\underset{1}{(x,}\ \underset{2}{x)} - A\underset{2}{(x,}\ \underset{1}{x)}$.

4. *Indication.* Comparer les dimensions de trois espaces correspondants.

5. *Indication.* La première affirmation est vérifiée par un calcul, la troisième résulte de la première.

6. *Indication.* Développer la forme $\mathrm{Sym}\ \underset{1}{\xi^{\alpha_1}} \ldots \underset{k}{\xi^{\alpha_k}}$ suivant les formes symétriques élémentaires. Comparer les coefficients.

7. *Indication.* Prouver pour les formes de base, par analogie avec *7.15*. Utiliser le résultat de l'exercice 5.

8. *Indication.* Itérer le résultat de l'exercice 5 écrit pour les formes du premier degré.

9. *Indication.* Utiliser *7.13b*.

10. *Indication.* Utiliser les formes $\xi^{i_1} \wedge \ldots \wedge \xi^{i_k}$.

11. *Indication.*

$$M = \begin{vmatrix} a_{11} & \ldots & a_{1k} \\ \cdot & \cdot & \cdot \\ a_{k1} & \ldots & a_{kk} \end{vmatrix} = \begin{vmatrix} a_{11} & \ldots & a_{1k} & a_{1,k+1} & \ldots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{k1} & \ldots & a_{kk} & a_{k,k+1} & \ldots & a_{kn} \\ 0 & \ldots & 0 & 1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \ldots & 0 & 0 & \ldots & 1 \end{vmatrix}, \qquad P^{-1} = \left\| \frac{A_{ij}}{\det P} \right\|.$$

12. *Indication.* Utiliser l'exercice 11.

13. *Indication.* Utiliser *7.16* (3).

14. *Indication.* Utiliser l'exercice 8.

15. *Indication.* Idem.

16. *Indication.* Utiliser les exercices 11 et 12 en considérant P comme matrice orthogonale.

17. *Indication.* Considérer d'abord une forme monôme.

18. *Indication.* Appliquer le théorème de Stokes-Poincaré.

Index

Additivité d'une mesure 199
Aire d'une sphère 262
  d'une surface 256
  d'un tore 262
Alternation 490
Application 12
  conforme 195
  contractante 54
  inverse 43
  jordanienne 227
  sphérique 382

Bivecteur 433
  unité 436
Branches d'une courbe 125

Caractéristique d'un multinuméro 537
Caténoïde 368
Cellule 199
Centre d'une droite 422
  de gravité 240
Chaîne 515
Chaînes équivalentes 516
Champ de Biot et Savart 339
  harmonique 314
  newtonien 334
  potentiel 314
  symétrique sphérique 332
  tensoriel 443
  vectoriel 301
Changement de variables dans une intégrale 250
Charge 199
  normale 213
Charges équivalentes 218
Circulation 316
Codifférentielle 530
Coefficient de distorsion 228
Coefficients de connexion 386
Composition de fonctions 16
Condition de Lipschitz 53
Connexion affine 452
  — symétrique 452
  riemannienne 453
Contraction d'un tenseur 430
Contravariance 428
Convergence absolue d'une intégrale impropre 278
Coordonnées sphériques 254, 298
Corps jordanien 213
Couche 269
Courbe lisse 46
Courbure d'une courbe sur une surface 369
  d'un espace riemannien dans une direction bidimensionnelle 475
  forcée 397
Courbure géodésique 397
  de Gauss 382
  — formelle 390
  moyenne 375
  normale 372
  totale 375
Courbures principales 375
Covariance 428
Cycle 520

Demi-anneau 199
Densité 225
Dérivation 29
  formelle 126
Dérivée 28
  contravariante 465
  covariante 465
  d'une fonction implicite 67
  — inverse 68
  d'ordre supérieur 129, 144
  partielle 26
  par rapport à un sous-espace 55
  suivant une direction 35
  — une ligne 46
Deuxième forme quadratique 369
Différentielle 26, 29
  absolue 449
  d'une forme 449
  géodésique 416
  d'ordre supérieur 131, 148
  — — partielle 45, 132
  , son invariance 41
Direction asymptotique 372
Directions principales 375
Distance d'un point à un ensemble 209
Divergence 303
Domaine admissible 275-277

Elément inverse 43
  — à droite 43
  — à gauche 43
Ensemble élémentaire 205
  jordanien 213
  négligeable 210
Equation différentielle 104
  de Poisson 338
Equations tensorielles 432
Equivalence affine 452
Espace chargé 199
  normalement chargé 213
  riemannien de courbure constante 479
  — élémentaire 445
Espaces riemanniens équivalents 445
Exemple de Schwartz 297
Extrémum 96
  lié 98

Flux d'un champ vectoriel 302
Fonction 11
additive 199
caractéristique 213
composée 16
continue 14
dérivable 25-27
— p fois 144
— par rapport à un sous-espace 55
fortement additive 255
harmonique 341
implicite 63
intégrable 200
inverse 62
linéaire 17
numérique 12
de n variables réelles 12
réelle 12
d'une variable réelle 12
vectorielle 12
Forme antisymétrique 489
— canonique 490
de degré p 147
différentielle 497
— adjointe 526
multilinéaire 146, 488
— antisymétrique 489
— symétrique 147
— — élémentaire 537
Formule d'Euler 375
de Gauss 388
— de dérivation 386
de Green 342, 343
de Meusnier 371
d'Ostrogradski 307-308
de Peterson-Codazzi 387
de Poisson 339, 346
de Stokes 320
de Taylor 133
Formules de dérivation 384
Frontière d'une chaîne 516
d'un ensemble 213

Géométrie intrinsèque 365
Gradient 26, 29
Graphique 12, 13

Hamiltonien 325
Hélicoïde 368
Homothétie 230

Identité de Poincaré 503
de Ricci 437
de Stokes-Poincaré 522
Image réciproque complète 83
Indicatrice de Dupin 376
Inégalité de Harnack 356
Intégrale 197, 199
d'un champ vectoriel 314
de Dirichlet 237
impropre avec une singularité variable 290
— de 1-ère espèce 275
— de 2-ème espèce 275
— de 3-ème espèce 276
itérée 219
n-uple 202
sur un ensemble jordanien 215
Intégrale de surface 256
Intégration des formes différentielles 508

Jacobien 31

Lemniscate 13
Ligne géodésique 396, 450, 457
de courbure 384
de niveau 13
de la plus rapide ascendance 36
de striction 422

Matrice de Jacobi 30
Maximum local 96, 138
— lié 98
Mesure 199
Méthode itérative 125
Minimum local 96, 138
— lié 98
Moment statique 240
Multinuméro 486
complémentaire 487
ordonné 486
strict 486
strictement ordonné 486
Multiplication de tenseurs 430

Notation canonique (première) d'une forme antisymétrique 495
— (deuxième) — — — 496
Noyau d'une application 83
Numéro 486

Opérateur hamiltonien 324
de Laplace 331, 530
— -Beltrami 532
Ordonnation 487
Orientabilité 304
Ovales de Cassini 13

Paraboloïde elliptique 12
hyperbolique 12
de révolution 12
Parallélépipède k-dimensionnel 241
Parallèles géodésiques 400
Parallélisme absolu 455
Paramètre canonique 458
Partition suivante 208
Pavé 202
Plan équilibrant 238
Pli 93
Point elliptique 373
hyperbolique 373
jordanien 224
méplat 374
parabolique 272
stationnaire 97
Potentiel 314
Première forme quadratique 363
Principe de Cavalieri 238
de localisation pour les intégrales impropres 279
Problème inverse de l'analyse vectorielle 352
Produit cartésien 15
d'espaces chargés 204
généralisé 23
tensoriel des formes 492
— — alterné 492
vectoriel 260
Projecteur 19
Propriétés absolues 441
Pseudo-sphère 410, 482

Rang d'un tenseur 428
Règle de Sylvester 136
Rigidité de surfaces multidimensionnelles 395

Rotation d'un champ vectoriel 313
Rotationnel 319
Ruban de Möbius 356

Schéma aux différences 187
Section normale 372
— complète 423
— élémentaire 423
Simplexe 244
Somme directe 19
Suite exhaustive 276
Suites en forme de delta 224
Surface admissible 509
fermée 304
de niveau 13
de révolution 368
Surfaces antiéquivalentes 511
équivalentes 510
Symboles de Christoffel 386
Symétrie de dérivées mixtes 153
de la dérivée seconde 152
de dérivées d'ordre supérieur 155
Systèmes de coordonnées admissibles 439
semi-géodésiques de coordonnées 403

Tangente 31
Tenseur 427
antisymétrique 429
de courbure 467
Tenseur métrique 435
— dérivé 436
symétrique 429
du type Ricci 437
Théorème de Bonnet 392
de Clairaut 407
sur la fonction implicite 62
de Frobenius 176
de Gauss (sur la courbure totale) 390
— (sur le triangle géodésique) 421
de Hilbert 410
de Janet-E. Cartan 396
de Levi-Civita 419
de Meusnier 372
de la moyenne 50
du rang 85
Torsion d'une connexion 453
Tractrice 409
Transformation de Fourier n-uple 299
Translation 416, 447, 452

Valeur moyenne d'une fonction 225
Variété différentiable élémentaire 439
Variétés équivalentes 440
Volume d'une boule 245
d'un ensemble jordanien 213
d'un parallélépipède 241
d'un simplexe 244
d'un tore 255

Table des matières

Avant-propos 5

Première partie

CALCUL DIFFÉRENTIEL ET INTÉGRAL

Chapitre 1. Dérivées du premier ordre 11
§ 1.1. Fonctions continues 11
§ 1.2. Fonctions dérivables 24
§ 1.3. Théorèmes généraux sur les fonctions dérivables 37
§ 1.4. Théorème de la moyenne 50
§ 1.5. Théorème sur la fonction implicite 62
§ 1.6. Structure locale d'une fonction dérivable 78
§ 1.7. Valeurs stationnaires de fonctions numériques 96
§ 1.8. Equations différentielles (théorèmes locaux) 103
§ 1.9. Equations différentielles (théorèmes non locaux) 117
Exercices 123
Historique 127

Chapitre 2. Dérivées d'ordre supérieur 129
§ 2.1. Dérivées d'ordre supérieur d'une fonction numérique de n variables 129
§ 2.2. Définition générale des dérivées d'ordre supérieur 144
§ 2.3. Propriétés des dérivées d'ordre supérieur 150
§ 2.4. Dérivées par rapport aux champs vectoriels 158
§ 2.5. Théorème de Frobenius 170
§ 2.6. Systèmes d'équations à dérivées partielles et applications géométriques 177
§ 2.7. Théorème de Taylor et son inversion 185
Exercices 194
Historique 196

Chapitre 3. Intégration dans les espaces à plusieurs dimensions 197
§ 3.1. Intégrale de Riemann sur un espace chargé 197
§ 3.2. Théorèmes d'existence 207
§ 3.3. Ensembles jordaniens 213
§ 3.4. Applications des espaces chargés 224
§ 3.5. Intégrale de Riemann dans un espace euclidien 228
§ 3.6. Intégrale de surface 256
§ 3.7. Intégrales impropres 275
Exercices 297
Historique 299

Chapitre 4. L'intégration et la dérivation 301
§ 4.1. Formule d'Ostrogradski 301
§ 4.2. Rotationnel du champ vectoriel 313
§ 4.3. Opérateur hamiltonien 324
§ 4.4. Certains types de champs vectoriels 332
§ 4.5. Champs et fonctions harmoniques 341
§ 4.6. Construction d'un champ vectoriel dans R_3 d'après ses rotationnel et divergence 352
Exercices 355
Historique 357

Deuxième partie

DES ESPACES VECTORIELS AUX VARIÉTÉS DIFFÉRENTIABLES

Chapitre 5. Géométrie différentielle classique 361
§ 5.1. Première forme quadratique 361
§ 5.2. Deuxième forme quadratique 369
§ 5.3. Relations entre la première et la deuxième formes quadratiques 384
§ 5.4. Lignes géodésiques et systèmes de coordonnées qui y sont liés 396
§ 5.5. Surfaces bidimensionnelles de courbure constante 408
§ 5.6. Translation des vecteurs et théorème de Levi-Civita . . . 416
Exercices 422
Historique 424

Chapitre 6. Géométrie riemannienne 426
§ 6.1. Théorie algébrique des tenseurs 426
§ 6.2. Variété différentiable élémentaire 439
§ 6.3. Espace riemannien élémentaire 445
§ 6.4. Espace à connexion affine 451
§ 6.5. Courbure 465
§ 6.6. Espaces riemanniens de courbure constante 478
Exercices 484
Historique 485

Chapitre 7. Différentiation et intégration sur les variétés 486
§ 7.1. Formes antisymétriques 486
§ 7.2. Formes différentielles 497
§ 7.3. Théorèmes intégraux 508
§ 7.4. Codifférentiation 526
Exercices 537
Historique 539
Indications et réponses 540
Index 548

Imprimé en Union Soviétique

DANS
LA MÊME
COLLECTION

Ouvrages parus :

COURS DE MATHÉMATIQUES SUPÉRIEURES
par V. SMIRNOV,
de l'Académie des Sciences de l'U.R.S.S.,
tome I, tome II,
tome III (en deux parties)

MÉTHODES DE LA THÉORIE DES FONCTIONS D'UNE VARIABLE COMPLEXE
par M. LAVRENTIEV,
de l'Académie des Sciences de l'U.R.S.S.,
et B. CHABAT,
docteur ès sciences physiques et mathématiques

ÉQUATIONS DIFFÉRENTIELLES ORDINAIRES
par V. ARNOLD,
membre correspondant de l'Académie des Sciences de l'U.R.S.S.

RECUEIL DE PROBLÈMES SUR LA THÉORIE DES FONCTIONS ANALYTIQUES
par M. EVGRAFOV (et coll.),
docteur ès sciences physiques et mathématiques

A paraître :

ÉLÉMENTS DE MATHÉMATIQUES APPLIQUÉES
par I. ZELDOVITCH,
de l'Académie des Sciences de l'U.R.S.S.,
et A. MYCHKIS,
docteur ès sciences physiques et mathématiques

ANALYSE MATHÉMATIQUE

FONCTIONS DE PLUSIEURS VARIABLES RÉELLES
1ère ET 2e PARTIES

De même que les volumes précédents, le présent livre est destiné à ceux qui possèdent déjà les connaissances de base dans le calcul différentiel et intégral et qui tiennent à les approfondir.

On considère, dans ce volume, la dérivation et l'intégration des fonctions du point d'un espace normé, notamment, des fonctions de plusieurs variables réelles, l'analyse vectorielle et la géométrie différentielle classiques, la géométrie riemannienne, certains problèmes de l'analyse sur les variétés différentiables, en particulier la théorie des formes différentielles antisymétriques avec ses théorèmes intégraux.