

Frédéric Pham

MAÎTRISE

Géométrie et calcul différentiel sur les variétés

Cours, études et exercices



2^e édition

DUNOD

Géométrie et calcul différentiel sur les variétés

Cours, études et exercices corrigés

Frédéric Pham

Professeur à l'université
de Nice-Sophia Antipolis

2^e édition

DUNOD

Du même auteur

Les différentielles, Masson, 1996

Découvrir l'algèbre linéaire, Diderot, 1996 (avec Hervé Dillinger)

Approche de la résurgence, Hermann, 1993 (avec B. Canderpergher et J.C. Nosmas)

Singularités des systèmes différentiels de Gauss-Manin, Birkhäuser, 1979

Conseiller scientifique

SINNOU DAVID

Ce pictogramme mérite une explication. Son objet est d'alerter le lecteur sur la menace que représente pour l'avenir de l'écrit, particulièrement dans le domaine de l'édition technique et universitaire, le développement massif du **photocopillage**.

Le Code de la propriété intellectuelle du 1^{er} juillet 1992 interdit en effet expressément la photocopie à usage collectif sans autorisation des ayants droit. Or, cette pratique s'est généralisée dans les établissements



d'enseignement supérieur, provoquant une baisse brutale des achats de livres et de revues, au point que la possibilité même pour les auteurs de créer des œuvres nouvelles et de les faire éditer correctement est aujourd'hui menacée.

Nous rappelons donc que toute reproduction, partielle ou totale, de la présente publication est interdite sans autorisation du Centre français d'exploitation du

droit de copie (CFC, 20 rue des Grands-Augustins, 75006 Paris).

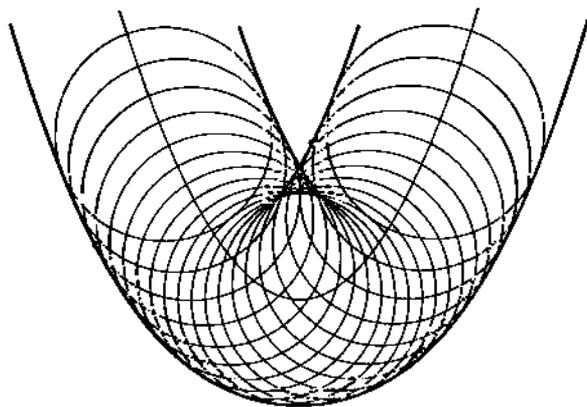
© InterÉditions, Paris, 1992

© Dunod, Paris, 1999

ISBN 2 10 004129 0

Toute représentation ou reproduction intégrale ou partielle faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite selon le Code de la propriété intellectuelle (Art L. 122-4) et constitue une contrefaçon réprimée par le Code pénal. • Seules sont autorisées (Art L. 122-5) les copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective, ainsi que les analyses et courtes citations justifiées par le caractère critique, pédagogique ou d'information de l'œuvre à laquelle elles sont incorporées, sous réserve, toutefois, du respect des dispositions des articles L. 122-10 à L. 122-12 du même Code, relative à la reproduction par reprographie.

À René Thom



« Bien que les pieds de l'homme n'occupent qu'un petit coin de la Terre, c'est par tout l'espace qu'il n'occupe pas que l'homme peut marcher sur la Terre immense. »

T'ehouang T'seu (traduction de Liou Kia-hway).

Préface à la deuxième édition

Relu aujourd'hui, le début de la première édition (chapitres 0, I, II) m'apparaît un peu comme un inventaire à la Prévert : quelques rappels (expédiés à la hâte) ; quelques bons exemples ; une profession de foi (que je ne renie pas) ; des bribes de formalisme ; des parenthèses polémiques (à l'usage des collègues)...

Tout cela est un peu chaotique, reflétant les tiraillements entre le physicien que j'ai commencé par être et le « bourbakiste¹ » que j'ai appris à devenir.

Six ans de réflexion (et d'expérience d'enseignement en amont de la maîtrise) m'ont conduit à présent à une vision plus claire et plus sereine des « contradictions » grâce auxquelles les mathématiques sont un organisme vivant. C'est pour exprimer cette vision apaisée, où les points de vue « ancien » et « moderne » se distinguent sans nécessairement s'opposer, que j'ai ré-écrit le début du livre ; le chapitre 1, presque entièrement nouveau, réalise méthodiquement et tranquillement le programme de « retour aux origines » que l'ancien chapitre I ne faisait que suggérer ; tout ce qui dans le début de la première édition se rattachait au point de vue « moderne » est maintenant exposé dans le chapitre 2, qui acquiert ainsi davantage de cohérence. La suite du livre est inchangée, à quelques menus détails près.

Juin 1998

¹ N'en déplaise aux collègues qui refusent de s'avouer bourbakistes, j'utilise ici cet adjectif comme synonyme de « mathématicien » !

Préface à la première édition

La théorie des variétés différentielles est un sujet que j'ai toujours trouvé particulièrement difficile à enseigner. Mais pourquoi l'enseigner, et à qui ?

Ce cours aura rempli son objectif si un étudiant moyen, au terme de son année de Maîtrise, en retire une bonne compréhension de la dualité entre le point de vue de la *géométrie*, qui s'intéresse aux « figures » et à leurs transformations, et celui du *calcul*, qui travaille avec des coordonnées et des équations.

Tel devrait être déjà l'objectif d'un cours de licence de calcul différentiel à n variables². Mais trop marqué qu'il est par l'idéologie ensembliste, notre enseignement du calcul différentiel dans \mathbb{R}^n ne sait pas faire la distinction entre un « point » et le n -uplet de ses coordonnées, entre un « vecteur » et le n -uplet de ses composantes. Émerveillés d'avoir appris à couler toutes les idées mathématiques dans le moule ensembliste, nous n'enseignons que la *forme* des idées, pensant que le *sens* suivra de lui-même. Parler du « vecteur vitesse » d'un « point mobile » est une atteinte à la pureté des mathématiques ! Utiliser la notation différentielle à la manière de nos pères, c'est-à-dire en considérant dx_1, \dots, dx_n comme des *variables*, est un « truc de physicien » !

Le malheur est que lorsqu'un étudiant formé à ce purisme arrive en Maîtrise, où toutes les notions évoquées plus haut prennent leur place naturelle dans le langage des variétés différentielles, on attend de lui non seulement une compréhension formelle de ce langage mais aussi une bonne intuition du savoir-faire que ce langage recouvre (savoir articuler judicieusement raisonnements géométriques et calculs en coordonnées). Faut-il s'étonner qu'il n'ait pas cette intuition alors qu'on n'a rien fait auparavant pour la lui faire acquérir ?

Avant de se poser pour le calcul différentiel, le problème se pose déjà pour l'enseignement de l'algèbre linéaire. C'est un problème difficile, auquel je ne prétends pas avoir de réponse miracle. Mais c'est le souci de ne pas l'é luder qui a inspiré la *première partie* de ce cours, dont j'aimerais pouvoir dire qu'il s'agit de *révisions et approfondissements des acquis essentiels du calcul différentiel de licence* (mais j'é mets là un vœu plus que je n'énonce une réalité !).

La *deuxième partie* présente la théorie intrinsèque des variétés (avec comme objectif essentiel la compréhension des notions de *fibré tangent* et *fibré normal*), et enchaîne

² et des cours d'algèbre linéaire en premier cycle !

sur les premiers rudiments de la topologie algébrique (homotopie et revêtements). Elle se termine par une ébauche de *théorie de l'intégration sur les variétés*, où l'on fait connaissance avec l'*homologie* et la *cohomologie*.

Parallèlement aux « chapitres » proprement dits, qui cherchent à présenter de façon cohérente les concepts qui forment l'ossature de la théorie, les « études » insérées entre les chapitres ont pour but de montrer ces concepts *en action dans un contexte*. Bien entendu, mes goûts personnels ont joué un rôle important dans le choix des contextes (où l'on pourra reconnaître l'influence de R. Thom, ainsi que de V. Arnold...). Mais parmi mes thèmes de prédilection, j'ai essayé de me limiter à ceux qui me paraissent mériter d'occuper une place centrale dans la culture d'un géomètre (au sens large de ce mot, allant de la géométrie algébrique à la mécanique), tout en restant assez « naïfs » pour qu'on puisse les raccrocher aux aspects plus élémentaires de la culture mathématique (calcul différentiel de premier cycle, géométrie au lycée,...). C'est dire que ces « études » sont pour moi bien davantage que de simples « compléments » ou « illustrations » des chapitres. C'est par leur interaction avec les exposés théoriques contenus dans les chapitres que j'ai essayé de montrer ce qu'est une démarche de mathématicien.

Peut-être les spécialistes trouveront-ils que certains thèmes d'étude mériteraient de plus amples développements, ou regretteront-ils l'absence dans les chapitres de tel ou tel résultat théorique important.

Inversement, les étudiants trouveront sans doute que ce cours est déjà bien gros, et que c'est beaucoup leur demander que d'attendre d'eux qu'ils maîtrisent autant de concepts abstraits, et sachent les appliquer dans autant de situations riches et complexes.

Aux uns comme aux autres, je répondrai qu'ils ont raison, mais que je n'ai voulu ni écrire un ouvrage de référence exhaustif sur le sujet, ni présenter aux étudiants une « somme de connaissances » à ingurgiter. J'aimerais que ce cours puisse être considéré simplement comme une porte ouverte sur un monde où les étudiants de niveaux variés pourront tous trouver à faire un bout de chemin³.

Remerciements

Bien que je sois le seul responsable de sa forme finale, ce livre doit beaucoup aux collègues avec qui j'ai eu l'occasion de partager des expériences d'enseignement, notamment François Rouvière, Georges Elencwajg, Hervé Dillinger, pour ne citer que ceux avec qui j'ai récemment collaboré sur ce sujet. Je remercie également Joël Briangon, dont le cours photocopié « Introduction aux Variétés différentiables » (Publication Pédagogique n° 4, Nice, 1988) m'a été très utile. Enfin je voudrais remercier mes étudiants, et tout particulièrement ceux de l'année 1991-1992, dont le travail opiniâtre et la soif de comprendre m'ont incité à remodeler maintes fois cet ouvrage et m'ont donné le courage de le mener à son terme.

³ Pour aider les étudiants à faire un premier tri dans l'ampleur du matériel qui leur est présenté, on a distingué sous le nom d'« exercices tests » des exercices qui doivent pouvoir être résolus *facilement* et *brèvement* si le concept auquel ils se rapportent a été assimilé.

Préface à la deuxième édition	VII
Préface à la première édition	IX
PREMIÈRE PARTIE : SOUS-VARIÉTÉS DE \mathbb{R}^N	1
Introduction	2
<i>Chapitre 1 – Calcul différentiel « à l'ancienne »</i>	3
0. Défense et illustration du calcul « à l'ancienne »	3
1. Différentielles	6
2. Interprétation géométrique	18
3. Théorème des fonctions implicites	27
Étude 1 : Points critiques et extrema locaux des fonctions	31
Étude 1^{bis} : Lemme de Morse	36
<i>Chapitre 2 – Applications différentiables et sous-variétés</i>	45
1. Applications différentiables entre ouverts d'espaces numériques	45
2. Point de vue géométrique sur les sous-variétés	51
3. Applications différentiables entre sous-variétés	55
4. Lc « foncteur tangent »	61
5. Transversalité	67
Exercices	69
Étude 2 : Contours apparents des surfaces	71
<i>Chapitre 3 – Champs de vecteurs</i>	75
1. Intégration d'un champ de vecteurs et dérivation le long d'un champ de vecteurs	75
2. Transformation des champs de vecteurs par difféomorphismes	79
Appendice : Compléments sur la notion de « flot »	82
Exercices	84
Étude 3 : Champs de gradients	85
Étude 3^{bis} : Mécanique de Newton et orbites planétaires	86
<i>Travaux pratiques : À propos de hérissons...</i>	92
<i>Intermède – Indices des champs de vecteurs et caractéristique d'Euler des surfaces</i>	93
1. Indice d'un champ de vecteurs le long d'un lacet du plan	93
2. Indice d'un point singulier isolé d'un champ de vecteurs du plan	96
3. Caractéristique d'Euler d'une surface compacte	100

DEUXIÈME PARTIE : POINT DE VUE INTRINSÈQUE SUR LES VARIÉTÉS	105
Introduction	107
Chapitre 4 – Variétés différentielles	109
0. Variétés topologiques	110
1. Structures différentielles sur une variété	112
2. Morphismes de variétés	115
Étude 4 : Un peu de géométrie plane	122
Étude 4^{bis} : La sphère de Riemann $S \approx \mathbb{C}P^1$	124
Appendice : Rappels sur les fonctions holomorphes	127
Étude 4^{ter} : Structures complexes sur le tore \mathbb{T}^2	127
Chapitre 5 – Tangence	130
0. Tangence de deux applications, jets d'applications	131
1. Foncteurs tangent et cotangent	133
2. Le théorème des fonctions implicites entre en scène	137
3. Fibré tangent et champs de vecteurs	138
4. Fibré cotangent et formes différentielles de degré 1	140
Étude 5 : Familles à un paramètre de droites du plan	143
Étude 5^{bis} : Champs de vecteurs constants sur les tores	154
Chapitre 6 – Espaces fibrés	156
0. De la notion de famille d'espaces à celle d'espace fibré	158
1. Fibrés vectoriels	163
2. Espaces fibrés, dans le contexte différentiel	166
Étude 6 : Enveloppes des familles de courbes	169
Appendice : Éléments de contact sur une variété et condition d'intégrabilité de Legendre	187
Chapitre 7 – Homotopie et revêtements	190
0. La catégorie d'homotopie	190
1. Groupe fondamental	194
2. Revêtements	200
3. Orientations	212
Étude 7 : Où une relecture de l'intermède nous ramène au tore \mathbb{T}^2	217
Chapitre 8 – Introduction au calcul intégral sur les variétés	220
0. L'algèbre des formes extérieures sur un espace vectoriel	221
1. L'algèbre graduée des formes différentielles sur une variété	223
2. Formes volumes	227
3. Intégration des formes différentielles	230
4. Introduction à l'homologie et à la cohomologie	234
Étude 8 : Mécanique hamiltonnienne et géométrie symplectique	241

PREMIÈRE PARTIE
SOUS-VARIÉTÉS DE \mathbb{R}^n

Introduction

Pourquoi \mathbb{R}^n ? Pour $n = 2$ ou 3 , c'est le modèle cartésien de l'espace euclidien usuel, habitacle des « figures » du géomètre (courbes, surfaces,...). Cette modélisation (du plan par \mathbb{R}^2 , de « l'espace » par \mathbb{R}^3) n'est pas canonique, le géomètre étant libre de son choix d'un « repère orthonormé ».

Mais il est d'autres contextes où la modélisation par \mathbb{R}^n est quasiment imposée par le problème : les sciences dites « exactes », au premier rang desquelles la physique, s'intéressent à des *grandeurs numériques* « variables » (par exemple en thermodynamique, la pression P , le volume V , la température T) liées par diverses *relations* (par exemple la relation $PV = RT$ en thermodynamique des gaz parfaits). Les variables du problème que l'on pourra désigner par x_1, \dots, x_n , définissent donc un *point* « variable » $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, mais l'existence de relations entre ces variables contraint ce point à rester sur un *sous-ensemble* $S \subset \mathbb{R}^n$ (de « dimension » d'autant plus petite que les relations sont nombreuses). Dans les bons cas ce sous-ensemble sera *lisse* (chap. 1), ce qui permettra de faire du *calcul différentiel* sur S (chap. 1 et 2), en considérant que les « accroissements infinitésimaux » (ou « différentielles ») dx_1, \dots, dx_n de ces variables sont liés par des relations *linéaires* (approximations linéaires des relations liant les x_1, \dots, x_n). Un problème d'*évolution* (des variables x_1, \dots, x_n en fonction du temps) sera gouverné par le *champ des vitesses* $v_i = dx_i/dt$ (chap. 3), champ de vecteurs *tangents* à la *variété* S des contraintes.

Calcul différentiel « à l'ancienne »

« Les différentielles du mathématicien
ne sont pas celles du physicien... »

(folklore universitaire, deuxième moitié du vingtième siècle)

0. Défense et illustration du calcul « à l'ancienne »

0.1. Petit exemple naïf : calcul de la tangente à une courbe plane

Il y a une dizaine d'années, j'ai raconté à des étudiants de première année d'université (filière « Sciences pour l'ingénieur ») le petit exemple que voici¹.

Soit à calculer la tangente au cercle d'équation :

$$x^2 + y^2 + 2x + y = 10 \quad (1)$$

au point de coordonnées (2, 1).

En différentiant l'équation (1), on trouve :

$$2x dx + 2y dy + 2 dx + dy = 0 \quad (2)$$

Pour $x = 2$, $y = 1$ cela donne $6 dx + 3 dy = 0$, c'est-à-dire $\frac{dy}{dx} = -2$. La tangente cherchée est donc donnée par la figure 1.

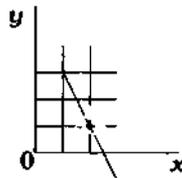


Figure 1.

« La méthode est drôlement simple » m'ont dit les étudiants. « Pourquoi est-ce qu'on ne nous l'a pas apprise en terminale ? »

¹ (publié dans P.C.P.É n°6, Nice déc. 1988)

« Cette méthode n'est pas du niveau du premier cycle universitaire » m'ont dit mes collègues, « car elle cache un théorème profond qui est le théorème des fonctions implicites ».

0.2. Discussion de l'exemple précédent

Pour n'importe quelle valeur de (x, y) solution de l'équation (1), la même méthode donne l'équation de la tangente à la courbe au point de coordonnées (x, y) :

$$(2x + 2) dx + (2y + 1) dy = 0 \quad (2')$$

Les valeurs de x et y étant fixées, (2') est une équation linéaire à deux inconnues dx, dy , inconnues que l'on peut interpréter comme des « accroissements de x et y le long de la tangente ». En tout point de la courbe où $2y + 1 \neq 0$ on peut résoudre l'équation (2') sous la forme $dy = p dx$, c'est-à-dire :

$$\frac{dy}{dx} = p \quad (3)$$

ce qui donne la pente p de la tangente à la courbe.

Aux points où $2y + 1 = 0$ l'équation (2') équivaut à $dx = 0$, ce qui veut dire que la tangente est *verticale* : dans notre exemple cela se produit aux deux points A_+, A_- de coordonnées respectives $(x_+, -\frac{1}{2})$ et $(x_-, -\frac{1}{2})$, où $x_{\pm} = -1 \pm \frac{\sqrt{45}}{2}$.

Début de justification

La formule (1) définit une **relation** entre les variables x et y . Si cette relation était **fonctionnelle en y** , c'est-à-dire si on pouvait l'écrire sous une forme « donnant y comme fonction de x » :

$$y = g(x)$$

la formule (3) définirait la *dérivée* $p = g'(x)$, qui est bien la pente de la tangente au graphe de g (au point considéré). Dans notre cas on peut effectivement considérer y comme fonction de x , à condition de restreindre convenablement les intervalles de variation de x et y : en l'occurrence, on pourra imposer à x et y de vérifier les inégalités :

$$x_- < x < x_+, \quad y > -\frac{1}{2}$$

ou si l'on préfère :

$$x_- < x < x_+, \quad y < -\frac{1}{2}$$

c'est-à-dire se restreindre à l'arc de cercle ouvert situé au-dessus (ou au-dessous) du diamètre A_+A_- (cf. fig.2).

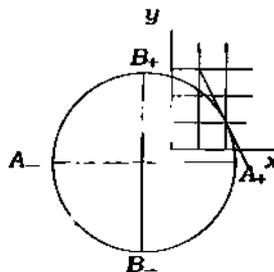


Figure 2.

Cette méthode n'est en défaut qu'aux points A_+ et A_- , où $y = -\frac{1}{2}$. Mais au voisinage de ces points on peut faire un raisonnement analogue en intervertissant les rôles de x et y , c'est-à-dire en cherchant à considérer x comme fonction de y (cela sera possible sur chacun des deux arcs de cercle ouverts délimités par le diamètre vertical B_+B_- de la fig. 2). En particulier aux points A_+ et A_- on trouve ainsi $\frac{dx}{dy} = 0$, ce qui veut bien dire que la tangente est parallèle à l'axe des y .

L'idée essentielle à retenir de cet exemple est que nous avons calculé des dérivées de fonctions *sans écrire d'expressions explicites de ces fonctions* : il nous a suffi pour cela de « différentier » la formule (1), dont on peut dire qu'elle définit « implicitement » ces fonctions.

Pour bien mesurer l'économie de calcul que cela représente, essayez de retrouver les résultats précédents en appliquant les règles de calcul des dérivées à une expression explicite comme :

$$y = \frac{-1 + \sqrt{1 - 4(x^2 + 2x + 10)}}{2}$$

(qui donne y comme fonction de x sur l'arc de cercle $A_+B_+A_-$) ! Et encore n'avons-nous considéré ici qu'un exemple où l'équation (1) est polynomiale du second degré, et peut donc sans difficulté être résolue par radicaux ! Dans des cas plus généraux (équations polynomiales de degré élevé) il ne sera même pas possible d'écrire de telles expressions explicites, et seule nous restera la méthode « implicite » précédente.

EXERCICE 0.1.— En quels points la courbe d'équation :

$$(x - y^2)^2 + y^2 = 1$$

a-t-elle une tangente parallèle à l'axe des x ? à l'axe des y ? À partir de ces informations, essayez de dessiner l'allure de la courbe.

EXERCICE 0.2.— Mêmes questions pour les courbes suivantes², données paramétriquement par $\begin{cases} x = \rho(\theta) \cos \theta \\ y = \rho(\theta) \sin \theta \end{cases}$, où :

- $\rho(\theta) = e^\theta$ (spirale logarithmique) ;
- $\rho(\theta) = 2 + \cos \theta$;
- $\rho(\theta) = \cos \theta$;
- $\rho(\theta) = 1 + \cos \theta$ (cardioïde).

EXERCICE 0.3.— En quels points la surface d'équation :

$$x^2 + (x + y)^2 + (x + y + z)^2 = 1$$

a-t-elle un plan tangent parallèle à l'un des plans de coordonnées ? En quels points son plan tangent est-il parallèle à l'axe des z ? Dessinez, en projection sur le plan des (x, y) , l'ensemble de ces points.

² On pourra aussi en déterminer les points où la tangente est parallèle à la première ou deuxième bissectrice.

Objectifs de ce chapitre

Le discours qui précède utilise un langage un peu désuet, qui fait jouer un rôle capital aux « variables » x, y, dx, dy . Traduire ce discours dans le langage des « maths modernes » (celui qui s'est imposé dans les années soixante) lui ferait inmanquablement perdre en simplicité³, et sans doute est-ce la raison pour laquelle cette méthode de calcul est rarement enseignée aujourd'hui en premier cycle universitaire.

Mais pourquoi ce langage si simple est-il tombé en désuétude ? D'aucuns vous diront que c'est parce qu'il manque de rigueur. Il est vrai que beaucoup d'auteurs du début du siècle s'autorisaient des « glissements » mal contrôlés entre raisonnements déductifs et interprétations intuitives. Mais je crois (et je vais essayer de montrer dans ce chapitre) qu'on peut *maîtriser* ces allers et retours entre *forme* et *sens*, grâce aux méthodes d'introspection que nous fournit la *logique mathématique*⁴ (notion d'**interprétation**, où les *variables* jouent un rôle crucial). Nous pourrions ainsi goûter en toute sécurité, et en toute liberté, les beautés du formalisme différentiel de Leibniz⁵.

Petit test linguistique

Vous aurez souvent l'occasion de lire dans ce chapitre des phrases comme :

1. « Substituons à la variable y son expression $g(x)$ comme fonction de x »
2. « La formule (1) définit une relation R entre les variables x et y , dont nous noterons S le graphe ».

Si le sens de telles tournures vous paraît clair, si rien n'y heurte l'idée que vous vous faites du « bon langage », continuez tranquillement à lire ce chapitre. Si quelque chose en elles vous dérange, lisez le *commentaire linguistique* en fin de chapitre.

1. Différentielles

Nous nous proposons ici de dégager les aspects « calculatoires » de la notion de différentielle. L'interprétation géométrique fera l'objet du paragraphe suivant.

³ Ou alors il faudrait introduire des concepts sophistiqués, comme ceux que nous développerons dans la suite de ce livre (chap. 5 § 4).

⁴ L'un des aspects paradoxaux des « maths modernes » est la façon dont elles ont conduit la plupart des mathématiciens à se désintéresser de la logique, réduite dans leur esprit au simple rôle de caution juridique garantissant le « bien-fondé » des usages en vigueur. À cette attitude il est réconfortant d'opposer celle du grand géomètre russe Yu. MANIN, qui dans l'introduction de son merveilleux livre « *A course in Mathematical Logic* » (traduction anglaise éditée par Springer-Verlag, 1977) en définit les objectifs en ces termes :

In this book mathematical logic is presented both as a part of mathematics and the result of its self-perception. (...)

Foundational problems are for the most part passed over in silence. Most likely, logic is capable of justifying mathematics to the same extent that biology is capable of justifying life.

⁵ Gottfried Wilhelm LEIBNIZ, philosophe et mathématicien allemand (1646-1716). Malgré les débats sans fin suscités par les idées de Leibniz sur les « infiniment petits », ses notations se sont rapidement installées au centre du paysage mathématique, et même les « maths modernes » n'ont pas réussi à les en chasser (malgré des commentaires du genre « Je n'ai jamais compris ce que c'est que le dx dans $\int f(x) dx$ », qu'on entendait souvent dans les milieux mathématiques à une certaine époque).

1.1. Différentielles de fonctions (comment les calculer)

Supposant connue la notion de *différentielle d'une fonction*⁶, nous allons ici focaliser notre attention sur les *propriétés formelles* qui permettent les calculs pratiques de différentielles.

Quand nous parlerons de « fonctions », il s'agira de fonctions de n variables réelles x_1, \dots, x_n , à valeurs réelles. En règle générale nous ne préciserons pas leurs domaines de définition, sauf lorsque ce sera indispensable. Rappelons que la propriété pour une fonction d'être différentiable est préservée par les opérations usuelles sur les fonctions : *sommes, produits*, dont il sera question tout de suite ; *substitutions* (composition des fonctions), dont il sera question un peu plus loin.

Propriétés algébriques

- o) La *différentielle d'une constante est nulle*.
- i) $d(f + g) = df + dg$.
- ii) $d(fg) = g df + f dg$.

Ces propriétés permettent de calculer les différentielles des fonctions polynomiales.

Exemple $d((x + y^2)^2 + y^2) = 2(x + y^2) d(x + y^2) + d(y^2) =$
 $2(x + y^2)(dx + 2y dy) + 2y dy = 2(x + y^2) dx + 2y(2(x + y^2) + 1) dy.$

De façon générale, si une fonction f de x_1, \dots, x_n est différentiable, sa différentielle df s'écrit d'une façon et d'une seule sous la forme :

$$p_1(x_1, \dots, x_n) dx_1 + \dots + p_n(x_1, \dots, x_n) dx_n \quad (1.1)$$

où les p_i sont des fonctions appelées **dérivées partielles de f** , et notées $p_i = \frac{\partial f}{\partial x_i}$, ou $p_i = \partial_{x_i} f$.

Dans le cas particulier $n = 1$ on a :

$$df = p(x) dx$$

où $p(x) = f'(x)$ est la fonction **dérivée** de f , encore notée $\frac{df}{dx}$.

Point de vue formel sur les différentielles Dans la pratique du calcul différentiel il est souvent commode de considérer une expression telle que (1.1) comme un objet formel, *sans chercher à donner un sens aux expressions* dx_1, \dots, dx_n : autrement dit on considère (1.1) comme une *forme linéaire à n indéterminées* dx_1, \dots, dx_n , à coefficients fonctions de x_1, \dots, x_n .

Un tel objet formel est ce qu'on appelle une **forme différentielle** (à n variables x_1, \dots, x_n). L'intérêt du point de vue formel est de permettre de distinguer la *pratique* des calculs de leur *interprétation*. On verra qu'un même calcul se prête à plusieurs *interprétations*, et c'est ce qui fait la grande force du formalisme de Leibniz, formalisme qui est sans doute l'une des inventions les plus géniales de l'histoire des mathématiques.

⁶ Cf. par exemple les chapitres 0 et 1 de mon livre *Les différentielles* (Masson 1996) ; nous reviendrons sur la définition de la différentielle au chap. 2 ci-après (§ 1.1).

Différentielles exactes Une forme différentielle (1.1) est dite **exacte** s'il existe une fonction f dont c'est la différentielle. Le problème de caractériser parmi toutes les formes différentielles celles qui sont exactes est pour $n > 1$ un problème délicat⁷.

Nous reviendrons sur ce problème au chap. 5 (§ 4.3). Mais nous aborderons dès le chap. 3 (étude 3 : « champs de gradients ») une variante géométrique du même problème, d'une très grande importance en mécanique et en physique.

Principe de substitution des différentielles Considérons une fonction f des variables x_1, \dots, x_n . En substituant aux variables x_i des fonctions de variables t_1, \dots, t_ℓ ⁸,

$$x_i := g_i(t_1, \dots, t_\ell), \quad i = 1, 2, \dots, n$$

on obtient une fonction \tilde{f} de t_1, \dots, t_ℓ .

Théorème Si f et les g_i sont différentiables, \tilde{f} l'est aussi, et sa différentielle $d\tilde{f}$ se déduit de df par les substitutions :

$$\begin{aligned} x_i &:= g_i(t_1, \dots, t_\ell) \\ dx_i &:= dg_i \end{aligned} \quad \square$$

Différentielle en un point Interprétons $x_1, \dots, x_n, dx_1, \dots, dx_n$ comme des *variables réelles*, et attribuons à (x_1, \dots, x_n) une valeur $a = (a_1, \dots, a_n)$. Pour toute valeur $u = (u_1, \dots, u_n)$ de (dx_1, \dots, dx_n) , la valeur correspondante de df est notée $df_a(u)$ et appelée **dérivée de f en a le long du vecteur u** ⁹:

$$df_a(u) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) u_i \quad (1.2)$$

(on en verra une interprétation géométrique au § 2.1).

Cette façon de voir les différentielles est le point de départ du point de vue moderne, où la **différentielle de f au point a** est la fonction linéaire :

$$\begin{aligned} df_a : \quad \mathbb{R}^n &\rightarrow \mathbb{R} \\ (u_1, \dots, u_n) &\mapsto df_a(u) \end{aligned}$$

⁷ Pour $n = 1$ toutes les formes différentielles « raisonnablement régulières » (par exemple toutes celles de la forme $\omega = p(x) dx$ avec p continue) sont exactes (car dire que ω est exacte revient à dire que la fonction p admet une primitive).

⁸ Rien n'interdit que certaines de ces variables t_k soient prises dans la liste x_1, \dots, x_n . C'est pour pouvoir le faire en toute liberté que j'utilise la notation « $x_i := \dots$ » (d'usage courant en informatique), que l'on doit comprendre ici comme une *décision d'affectation* (décision de remplacer x_i par l'expression écrite à droite du signe « := »).

⁹ On dit parfois « **dérivée directionnelle** de f etc. ». Mais il ne faudrait pas croire qu'elle ne dépend que de la *direction* du vecteur u : en fait il s'agit d'une fonction *linéaire* de u .

1.2. Différentier des relations

Relisons les calculs du § 0.1, en nous interrogeant sur la nature de la déduction logique qui permet de passer de la relation (1) à la relation (2). On pourrait être tenté de dire que (1) « implique » (2), ou (pour prendre un exemple plus simple) que la relation $x^2 + y^2 = 1$ « implique » la relation $2x dx + 2y dy = 0$. Mais ces relations ne portent pas sur les mêmes variables ! Bien sûr on pourrait dire que la première relation porte, comme la seconde, sur les variables x, y, dx, dy , car rien n'oblige une relation à faire intervenir effectivement toutes les variables qu'elle est censée relier (sinon on ne pourrait jamais parler de « fonction constante » !). Malheureusement l'implication $x^2 + y^2 = 1 \Rightarrow 2x dx + 2y dy = 0$ ainsi comprise est fautive¹⁰ !

Il y a là un « paradoxe » qu'il faut élucider, pour apprendre à raisonner juste quand on fait du calcul différentiel.

Dans ce paragraphe nous distinguerons soigneusement deux types de relations : les premières porteront sur les n variables x_1, \dots, x_n , et nous les noterons par des majuscules romaines telles que R ;

les secondes porteront sur les $2n$ variables $x_1, \dots, x_n, dx_1, \dots, dx_n$, et nous les noterons (au moins pour le moment) par des majuscules grecques telles que Σ ; nous les distinguerons des précédentes en les appelant **relations différentielles**.

Une **interprétation** φ des variables x_1, \dots, x_n (sous-entendu : « ...comme fonctions différentiables de variables t_1, \dots, t_ℓ ») est un dictionnaire qui à chacune des variables x_i fait correspondre une fonction différentiable $\varphi_i(t_1, \dots, t_\ell)$, et à chaque relation R entre les x_i fait correspondre la relation $R^\varphi(t_1, \dots, t_\ell)$ qui s'en déduit par les substitutions $x_i := \varphi_i(t_1, \dots, t_\ell)$ ($i = 1, \dots, n$). Si la relation R^φ ainsi obtenue est *identiquement vraie*, c'est-à-dire vérifiée pour toutes les valeurs de (t_1, \dots, t_ℓ) , nous dirons que l'**interprétation** φ **vérifie la relation** R .

Exemple Soit $R(x, y)$ la relation $x^2 + y^2 = 1$. Interprétons x et y comme les fonctions \cos et \sin (d'une variable réelle t). Cette interprétation vérifie la relation R .

Cas particulier : interprétations constantes Dans le cas particulier où chacune des fonctions φ_i est une constante α_i on retrouve la notion usuelle de « valeurs des variables vérifiant la relation R ».

Étant donné une interprétation φ des variables x_1, \dots, x_n , j'appellerai **interprétation dérivée** le dictionnaire qui étend le dictionnaire précédent de la façon que voici : à chacune des variables dx_i on fait correspondre la différentielle $d\varphi_i$ de la fonction φ_i ; à chaque relation différentielle Σ on fait correspondre la relation $\Sigma^\varphi(t_1, \dots, t_\ell, dt_1, \dots, dt_\ell)$ qui s'en déduit par les substitutions $x_i := \varphi_i(t_1, \dots, t_\ell)$, $dx_i := d\varphi_i$ ($i = 1, \dots, n$). Si la relation Σ^φ ainsi obtenue est *identiquement vraie*, c'est-à-dire vérifiée pour toutes les valeurs de $(t_1, \dots, t_\ell, dt_1, \dots, dt_\ell)$, nous dirons que l'**interprétation dérivée de φ vérifie la relation** Σ .

¹⁰ Par exemple pour $x = \frac{1}{\sqrt{2}}$, $y = \frac{1}{\sqrt{2}}$ la première équation est vérifiée pour toutes les valeurs de (dx, dy) , alors que la seconde ne l'est que si l'on donne à dx et dy des valeurs opposées.

DÉFINITION (implication différentielle) Soit $R(x_1, \dots, x_n)$ une relation, et soit $\Sigma(x_1, \dots, x_n, dx_1, \dots, dx_n)$ une relation différentielle.

Nous dirons que Σ dérive de R (ou que R **implique différentiellement** Σ), et nous écrirons :

$$R \stackrel{d}{\Rightarrow} \Sigma$$

si pour toute interprétation φ de x_1, \dots, x_n qui vérifie R l'interprétation dérivée de φ vérifie Σ .

Exemple $x^2 + y^2 = 1 \stackrel{d}{\Rightarrow} 2x dx + 2y dy = 0$. Plus généralement :

Lemme « d » Soit f une fonction différentiable des variables x_1, \dots, x_n . Alors :

$$f(x_1, \dots, x_n) = 0 \stackrel{d}{\Rightarrow} df = 0$$

Preuve

C'est une conséquence immédiate du principe de substitution (§ 1.1). □

Exercice de style Pour illustrer le langage précédent, voici une démonstration détaillée de la formule classique :

$$(\arcsin)'(x) = \frac{1}{\sqrt{1-x^2}}$$

Soit R la relation $x = \sin \theta$ entre les variables θ et x restreintes aux intervalles :

$$-\frac{\pi}{2} < \theta < +\frac{\pi}{2} \qquad -1 < x < 1$$

Évidemment (lemme « d »), $R \stackrel{d}{\Rightarrow} dx = \cos \theta d\theta$, donc :

$$R \stackrel{d}{\Rightarrow} \Sigma : \begin{cases} x = \sin \theta \\ dx = \cos \theta d\theta \end{cases} \iff \begin{cases} x = \sin \theta \\ dx = \sqrt{1-x^2} d\theta \end{cases}$$

(où pour établir l'équivalence on a utilisé le fait que $\cos \theta > 0$).

Considérons l'interprétation suivante de x et θ comme fonctions de x : ($x := x$, $\theta := \arcsin x$) ; par définition de arcsin cette interprétation vérifie R , donc son interprétation dérivée vérifie Σ , ce qui veut dire que l'on a :

$$dx = \sqrt{1-x^2} d(\arcsin x). \text{ donc } d(\arcsin x) = \frac{1}{\sqrt{1-x^2}} dx \quad \square$$

1.3. Différentielles d'ordre supérieur

1.3.1. Différentielles itérées d'une fonction

Une fonction f de n variables x_1, \dots, x_n est dite *deux fois différentiable* si elle est différentiable et si ses dérivées partielles $p_i(x) = \frac{\partial f}{\partial x_i}(x)$ sont différentiables. On peut alors différentier l'expression (1.1) de df . En convenant de traiter les dx_i de cette expression comme des constantes, on trouve ainsi une forme quadratique en dx_1, \dots, dx_n (polynôme homogène de degré deux), appelée *différentielle seconde* de f :

$$d^2f = \sum_{j=1}^n dp_j dx_j = \sum_{j=1}^n \sum_{i=1}^n p_{ij}(x_1, \dots, x_n) dx_i dx_j \quad (1.3)$$

Les coefficients p_{ij} sont les *dérivées partielles secondes* de f , notées :

$$p_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} f$$

Lemme de Schwarz

$$p_{ij} = p_{ji}$$

□

Autrement dit,

$$\frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} = \frac{\partial}{\partial x_j} \frac{\partial}{\partial x_i}$$

ce qui permet d'écrire symboliquement :

$$d^2f = \left(dx_1 \frac{\partial}{\partial x_1} + \dots + dx_n \frac{\partial}{\partial x_n} \right)^2 f$$

En itérant cette construction on obtient la notion de *fonction r fois différentiable*, et de *différentielle d'ordre r d'une fonction* : f est r fois différentiable si ses dérivées partielles d'ordre $r - 1$ sont différentiables ; sa *différentielle d'ordre r* est alors un polynôme homogène de degré r en dx_1, \dots, dx_n :

$$d^r f = \left(dx_1 \frac{\partial}{\partial x_1} + \dots + dx_n \frac{\partial}{\partial x_n} \right)^r f = \sum_{i_1, \dots, i_r} p_{i_1, \dots, i_r}(x_1, \dots, x_n) dx_{i_1} \dots dx_{i_r} \quad (1.4)$$

où les coefficients p_{i_1, \dots, i_r} sont les *dérivées partielles d'ordre r de f* , encore notées :

$$p_{i_1, \dots, i_r} = \frac{\partial^r}{\partial x_{i_1} \dots \partial x_{i_r}} f$$

Cas particulier $n = 1$: $d^2f = f''(x) dx^2$, etc. ... $d^r f = f^{(r)}(x) dx^r$, où $f^{(r)}$ est la dérivée r -ième de f , encore notée $f^{(r)}(x) = \frac{d^r f}{dx^r}(x)$.

Formule de Taylor-Young

Supposons que f soit r fois différentiable au point $a = (a_1, \dots, a_n)$. Alors :

$$f(a_1 + u_1, \dots, a_n + u_n) = f(a_1, \dots, a_n) + \frac{1}{1!} df_a(u_1, \dots, u_n) + \frac{1}{2!} d^2 f_a(u_1, \dots, u_n) + \dots + \frac{1}{r!} d^r f_a(u_1, \dots, u_n) \text{ mod. } o(\|u\|^r)^{11}$$

Pour $r = 1$ cette proposition n'est que l'expression de la définition de la différentielle df_a (cf. chap. 2, § 1.1).

Pour r quelconque le second membre de cette égalité est appelé **développement limité** à l'ordre r de la fonction f au point a .

Fonctions différentiables de classe C^r Une fonction f est dite de **classe C^0** si elle est continue. Elle est dite de **classe C^r** (r entier ≥ 1) si elle est r fois différentiable, et si ses dérivées partielles d'ordre r sont continues (dans le cas $r = 1$ on dit aussi que f est **continûment différentiable**). Elle est dite de classe C^∞ (ou **indéfiniment différentiable**) si elle est de classe C^r pour tout r . La série infinie $\sum_{r=0}^{\infty} \frac{1}{r!} d^r f_a$ est alors

appelée **série de Taylor de f en a** . La fonction f est dite **analytique**¹² si sa série de Taylor en tout point a converge vers f au voisinage de ce point.

Si l'on travaille sur \mathbb{C} au lieu de \mathbb{R} toutes les notions précédentes sont équivalentes, en ce sens que pour toute fonction de n variables complexes la \mathbb{C} -différentiabilité (de classe C^1 par exemple) équivaut à l'analyticité.

1.3.2. Implications différentielles itérées

Selon une opinion très répandue, les vertus « magiques » de la notation « d » de Leibniz seraient limitées au calcul différentiel d'ordre 1. Pourtant, si l'on prend bien soin (comme nous l'avons fait jusqu'à présent) de distinguer les *formules* de leurs *interprétations*, on peut en toute sécurité étendre aux ordres supérieurs la démarche du § 1.2 : on adjoint aux variables x_1, \dots, x_n des variables dx_1, \dots, dx_n , puis d^2x_1, \dots, d^2x_n , etc.; pour toute interprétation des x_i comme fonctions r fois différentiables de variables t_1, \dots, t_ℓ on a une notion naturelle « d'interprétation dérivée » des $dx_i, d^2x_i, \dots, d^r x_i$ comme différentielles d'ordre supérieur desdites fonctions ; la définition de « l'implication différentielle » $\stackrel{d}{\Rightarrow}$ se généralise de façon évidente.

Exemple $x^2 + y^2 = 1 \stackrel{d}{\Rightarrow} 2x dx + 2y dy = 0 \stackrel{d}{\Rightarrow} 2x d^2x + 2y d^2y + 2(dx)^2 + 2(dy)^2 = 0$.

EXERCICE 1.1.–

- Tester ces équations sur l'interprétation $x = \cos t, y = \sin t$.
- Retrouver sur ces équations l'expression de la dérivée seconde de la fonction $y(x) = \sqrt{1 - x^2}$.

Attention !!! Il faudra poser $d^2x = 0$; comprenez-vous pourquoi ?

¹¹ La notation $o(\alpha)$ (lire « petit o de α ») sert à qualifier toutes les fonctions φ qui tendent vers 0 plus vite que la fonction α , c'est-à-dire telles que φ/α tend vers 0 quand la variable tend vers la valeur qui nous intéresse (ici la variable est $u = (u_1, \dots, u_n)$, et la valeur qui nous intéresse est $u = (0, \dots, 0)$).

¹² On dit parfois « de classe C^ω », la lettre « omega » suggérant « l'au-delà de l'infini ».

Exercice de style : dérivée seconde d'une fonction réciproque

Soient x et y deux variables réelles liées par une relation à la fois fonctionnelle en y ($y = y(x)$) et en x ($x = x(y)$). Ces deux fonctions étant supposées deux fois différentiables, exprimer les dérivées première et seconde de l'une en fonction des dérivées première et seconde de l'autre.

SOLUTION $y = y(x) \stackrel{d}{\Rightarrow} dy = y'(x) dx \stackrel{d}{\Rightarrow} d^2y = y''(x)(dx)^2 + y'(x) d^2x$. Dans l'interprétation $x := x(y)$, $y := y$ on a $d^2y = 0$, donc $d^2x = -\frac{y''(x)}{y'(x)}(dx)^2 = -\frac{y''(x)}{y'(x)^3}(dy)^2$, de sorte que $x''(y) = -y''(x)/y'(x)^3$ (et bien sûr $x'(y) = \frac{1}{y'(x)}$).

EXERCICE 1.2.— Supposant les fonctions ci-dessus indéfiniment différentiables, pousser un cran plus loin la comparaison de leurs dérivées successives. Plus généralement, montrer que pour tout r la connaissance du développement limité à l'ordre r de $x(y)$ équivaut à celle du développement limité à l'ordre r de $y(x)$.

Application géométrique : tangence à l'ordre r de deux courbes planes Interprétant x, y comme des coordonnées cartésiennes dans le plan, on considère une courbe C passant par le point $a = (a_1, a_2)$, donnée localement au voisinage de ce point par une équation de la forme $y = g(x)$, où g est une fonction de classe C^∞ . Soit \tilde{C} une autre courbe passant aussi par a , donnée par une équation du même type $y = \tilde{g}(x)$.

Les courbes C et \tilde{C} sont dites *tangentes à l'ordre r en a* si les développements limités à l'ordre r en a_1 des deux fonctions g et \tilde{g} coïncident : pour $r = 1$ c'est la notion usuelle de tangence ; pour $r = 2$ on dit aussi que les courbes sont *osculatrices*. Il résulte de l'exercice précédent que la notion de tangence à l'ordre r de deux courbes n'est pas affectée par une inversion des rôles de x et y (en supposant C et \tilde{C} données par des équations $x = h(y)$ et $x = \tilde{h}(y)$).

EXERCICE 1.3.— (cercle osculateur)¹³ La courbe C étant donnée comme ci-dessus, montrez que par le point a passe un et un seul cercle osculateur à C en a , dont vous explicitez la position (centre ω et rayon r) en fonction des dérivées première et seconde de g en a_1 .

Indication En différentiant deux fois l'équation du cercle :

$$(x - \omega_1)^2 + (y - \omega_2)^2 = r^2 \quad (= (a_1 - \omega_1)^2 + (a_2 - \omega_2)^2)$$

on calculera le développement limité à l'ordre 2 en a_1 de la fonction $y(x)$.

EXEMPLES Calculez et dessinez le cercle osculateur

- à la parabole $y = x^2$, d'abord au point $(0, 0)$ puis au point $(1, 1)$;
- à la courbe de l'exercice 0.1, en divers points remarquables de cette courbe.

EXERCICE 1.4.— On considère quatre variables x, y, ρ, θ liées par la relation :

$$\begin{cases} x = \rho \cos \theta \\ y = \rho \sin \theta \end{cases}$$

(formule de passage des « coordonnées polaires » ρ, θ aux « coordonnées cartésiennes » x, y : cf. § 2.0 ci-après).

¹³ Ici on interprète x, y comme des coordonnées dans un repère orthonormé d'un plan euclidien.

- a) Dédurre de cette relation (au sens de l'implication différentielle $\stackrel{d}{\Rightarrow}$ une expression de dx, dy en fonction de $\rho, \theta, d\rho, d\theta$, puis une expression de d^2x, d^2y en fonction de $\rho, \theta, d\rho, d\theta, d^2\rho, d^2\theta$.
- b) Une courbe plane C est supposée donnée en coordonnées cartésiennes par une équation $y = y(x)$, et en coordonnées polaires par une équation $\rho = \rho(\theta)$.
Dédurre de a) une expression de la dérivée seconde $y''(x) = \frac{d^2y}{dx^2}$ en fonction des dérivées premières et secondes de la fonction $\rho(\theta)$.
- c) En utilisant le résultat de l'exercice 1.3, déduire de b) l'expression en coordonnées polaires du rayon du cercle osculateur à C .

EXEMPLES Calculez et dessinez le cercle osculateur à chacune des courbes de l'exercice 0.2, en quelques points remarquables de ces courbes.

1.3.3. Mise en garde

Généraliser aux ordres supérieurs le principe de substitution des différentielles ne doit pas se faire sans précautions. Soit par exemple à calculer la différentielle seconde $d^2\tilde{f}$ de la fonction \tilde{f} déduite de f par les substitutions $x_i := g_i(t_1, \dots, t_\ell)$ (les notations sont celles du § 1.1). On pourrait être tenté d'écrire tout simplement l'expression (1.3) de d^2f , et d'y faire les substitutions $x_i := g_i(t_1, \dots, t_\ell), dx_i := dg_i$. Cette façon de calculer est erronée, car elle revient à oublier le terme en d^2x_i dans l'expression formelle ci-dessous de d^2f :

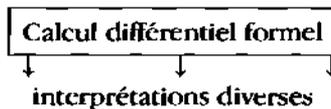
$$df = \sum_i p_i(x) dx_i \stackrel{d}{\Rightarrow} d^2f = \sum_i dp_i(x) dx_i + \sum_i p_i(x) d^2x_i \quad (1.5)$$

Dans l'interprétation $x_i := g_i(t_1, \dots, t_\ell)$ le terme d^2x_i n'est pas nul, mais vaut :

$$d^2x_i = \sum_{j,k} \frac{\partial^2 g_i}{\partial t_j \partial t_k} dt_j dt_k$$

La formule (1.3) était correcte dans le contexte du § 1.3.1, où la variable x_i était interprétée comme « fonction d'elle-même » (application identique $x_i \mapsto x_i$), de sorte que $d^2x_i = 0$.

On évitera ce type d'erreur en se fixant la règle de conduite suivante : dans tout calcul de différentielles où l'on va être amené à interpréter des variables de plusieurs façons, commencer par procéder formellement, en n'écrivant que des formules vraies pour toute interprétation des variables ; n'interpréter qu'à la fin.



1.4. Différentielle d'une relation lisse

Soit R une relation entre n variables x_1, \dots, x_n . Parmi toutes les relations différentielles Σ qui dérivent de R (cf. § 1.2) nous allons, sous certaines hypothèses dites de « lissité », en construire une « plus forte que les autres » que nous appellerons **différentielle de la relation** R et que nous noterons DR . Dire que cette relation est « plus forte que les autres » signifie ceci.

Propriété universelle de la différentielle

Pour toute relation différentielle Σ telle que $R \stackrel{d}{\Rightarrow} \Sigma$, on a $DR \Rightarrow \Sigma$.

Notons que parmi toutes les relations différentielles qui dérivent de R il ne peut en exister qu'une qui ait la propriété universelle (en effet deux telles relations s'impliquent mutuellement).

Définition (différentielle d'une relation fonctionnelle)

Considérons un partage de l'ensemble des variables en « variables de départ » x_{i_1}, \dots, x_{i_m} et « variables d'arrivée » x_{j_1}, \dots, x_{j_p} . La relation R est dite **fonctionnelle différentiable** en x_{j_1}, \dots, x_{j_p} , ou « de type $\text{Dif}(x_{j_1}, \dots, x_{j_p} | x_{i_1}, \dots, x_{i_m})^{14}$ », si elle est fonctionnelle en les variables x_{j_1}, \dots, x_{j_p} , donnant celles-ci comme fonctions différentiables des autres variables.

Sous cette hypothèse la différentielle DR de la relation R peut être définie par la formule :

$$\begin{cases} x_{j_1} = g_1(x_{i_1}, \dots, x_{i_m}) \\ \dots \\ x_{j_p} = g_p(x_{i_1}, \dots, x_{i_m}) \\ dx_{j_1} = dg_1 \\ \dots \\ dx_{j_p} = dg_p \end{cases} \quad (1.6)$$

où l'on a noté g_1, \dots, g_p les fonctions évoquées plus haut, et dg_1, \dots, dg_p leurs différentielles.

Le fait que la relation DR ainsi définie dérive de R résulte immédiatement du lemme « d ».

PROPOSITION (universalité de DR)

La relation DR définie par (1.6) a la propriété universelle.

Preuve

Considérons l'interprétation évidente de x_1, \dots, x_n comme fonctions des « variables de départ » x_{i_1}, \dots, x_{i_m} (les variables de départ s'interprétant comme elles-mêmes et celles d'arrivées x_{j_1}, \dots, x_{j_p} comme les fonctions g_1, \dots, g_p respectivement). Par construction cette interprétation φ vérifie R . Puisque $R \stackrel{d}{\Rightarrow} \Sigma$, l'interprétation dérivée de φ vérifie Σ . Mais cette interprétation dérivée consiste justement à substituer à x_{j_1}, \dots, x_{j_p} , $dx_{j_1}, \dots, dx_{j_p}$ leurs expressions données par les membres de droites de la formule (1.6); dire qu'elle vérifie Σ , c'est donc bien dire que $(1.6) \Rightarrow \Sigma$. \square

Remarque Bien que le partage des variables en « variables de départ » et « variables d'arrivée » nous ait été utile pour écrire la formule (1.6) qui donne DR , notez bien que DR est caractérisée par la propriété universelle, où ce partage n'intervient nullement. Deux partages différents, de type $\text{Dif}(\dots | \dots)$, donneront donc la même relation sous deux formes différentes.

¹⁴ Notation inventée pour les besoins de ce cours. Nous écrivons Dif_C au lieu de Dif si la différentiabilité est de classe C^r .

Un exemple important de cette situation est fourni par les « changements de variables différentiables ».

DÉFINITION On appelle **changement de variables différentiable** la donnée, notée $x_1, \dots, x_n \xrightarrow{R} y_1, \dots, y_n$, de deux listes de variables et d'une relation $R(x_1, \dots, x_n, y_1, \dots, y_n)$ entre ces variables, relation à la fois de type $\text{Dif}(y_1, \dots, y_n | x_1, \dots, x_n)$ et de type $\text{Dif}(x_1, \dots, x_n | y_1, \dots, y_n)$.

La formule (4) du § 0 fournit un exemple de changement de variables qui n'est pas différentiable (il est de type $\text{Dif}(x|y)$ mais pas $\text{Dif}(y|x)$, car la fonction $\sqrt{\quad}$ n'est pas différentiable en 0).

Exemple Soit R la relation $x = \sin \theta$ entre les variables θ et x . Elle est évidemment de type $\text{Dif}(x|\theta)$, de sorte que sa différentielle peut s'écrire :

$$DR : \begin{cases} x = \sin \theta \\ dx = \cos \theta d\theta \end{cases}$$

Mais si l'on restreint les domaines de variations des variables aux intervalles :

$$-\frac{\pi}{2} < \theta < +\frac{\pi}{2} \qquad -1 < x < 1$$

elle est aussi de type $\text{Dif}(\theta|x)$, de sorte que sa différentielle peut aussi s'écrire

$$DR : \begin{cases} \theta = \arcsin x \\ d\theta = (\arcsin)'(x) dx \end{cases}$$

Dire que ces deux formules sont équivalentes revient à dire que l'on a $(\arcsin)'(x) = \frac{1}{\sqrt{1-x^2}}$ (comparez à « l'exercice de style » du § 1.2).

DÉFINITION Soit $a = (a_1, \dots, a_n)$ un point du graphe de la relation R .

La relation R est dite **lisse au point** a s'il existe un partage comme ci-dessus tel que R soit de type $\text{Dif}(x_{j_1}, \dots, x_{j_p} | x_{i_1}, \dots, x_{i_m})$ **localement au voisinage de** a , c'est-à-dire après restriction des domaines de variation de chaque x_i à un intervalle ouvert convenable contenant a_i .

La relation R est dite **lisse** si elle est lisse en chacun des points de son graphe (on précisera « lisse de classe C^r » si la différentiabilité Dif est en fait de type Dif_{C^r}).

Différentielle d'une relation lisse Dans chacun des domaines (produits cartésiens d'intervalles) où R a une propriété $\text{Dif}(\dots | \dots)$, on a vu (formule (1.6)) comment construire la différentielle DR . Dans l'intersection de deux tels domaines, correspondant à deux partages différents en variables de départ et variables d'arrivée, les deux constructions donneront le même résultat (cf. remarque ci-avant). La différentielle se trouve ainsi définie partout.

Les deux propriétés suivantes de la relation DR se lisent immédiatement sur la formule (1.6).

Propriété 1 Cette relation n'est vérifiée que si (x_1, \dots, x_n) vérifie R .

Propriété 2 À (x_1, \dots, x_n) fixé, vérifiant R , elle est linéaire homogène en (dx_1, \dots, dx_n) .

Précisons cette deuxième propriété. Pour tout point $a = (a_1, \dots, a_n)$ du graphe de R , on notera DR_a , et on appellera **différentielle de** R en a la relation entre dx_1, \dots, dx_n déduite de DR en fixant (x_1, \dots, x_n) à la valeur (a_1, \dots, a_n) .

Alors le graphe de DR_a est un sous-espace vectoriel de \mathbb{R}^n , dont la dimension est localement constante quand a parcourt le graphe de R^{15} (elle est donc constante sur chaque composante connexe du graphe de R).

DÉFINITION Cette dimension, notée $\dim_a R$, est appelée **dimension de R en a** , ou simplement **dimension de R** si elle est indépendante de a (ce qui est le cas notamment si le graphe de R est connexe).

Connaissant la dimension de R , la proposition suivante donne une caractérisation commode de sa différentielle.

PROPOSITION (caractérisation de DR)

Soit Σ une relation différentielle dérivant de R et ayant la propriété 1, ainsi que la propriété 2 ainsi précisée :

Propriété 2^{bis} Pour tout point $a = (a_1, \dots, a_n)$ du graphe de R , l'ensemble des valeurs de (dx_1, \dots, dx_n) qui vérifient Σ_a^{16} est un sous-espace vectoriel de dimension $\dim_a R$.

Alors $\Sigma = DR$.

Preuve

Le seul cas où les graphes de DR_a et Σ_a ne sont pas vides est le cas où a appartient au graphe de R (propriété 1). Dans ce cas ce sont des espaces vectoriels de même dimension (propriété 2^{bis}). Comme le premier est inclus dans le second (propriété universelle : $DR_a \Rightarrow \Sigma_a$), ils sont égaux. \square

Exemple La relation $R : x^2 + y^2 = 1$ est lisse de dimension 1 (raisonner comme dans l'exemple du § 0.1, en considérant le cercle comme union de quatre arcs de demi-cercles ouverts). La relation différentielle $\Sigma : \begin{cases} x^2 + y^2 = 1 \\ 2x dx + 2y dy = 0 \end{cases}$ dérive de R , et a évidemment la propriété 1. Elle a aussi la propriété 2^{bis}, car si (a_1, a_2) est une valeur de (x, y) telle que $a_1^2 + a_2^2 = 1$, on a $(a_1, a_2) \neq (0, 0)$, de sorte que l'équation linéaire $2a_1 dx + 2a_2 dy = 0$ définit une droite vectorielle de \mathbb{R}^2 . Donc $\Sigma = DR$.

Introduction au théorème des fonctions implicites

Considérons, plus généralement, une relation R de la forme :

$$s(x_1, \dots, x_n) = 0$$

où s est une fonction différentiable, dont on notera :

$$ds = p_1(x_1, \dots, x_n) dx_1 + \dots + p_n(x_1, \dots, x_n) dx_n$$

la différentielle. La relation différentielle :

$$\Sigma : \begin{cases} s(x_1, \dots, x_n) = 0 \\ p_1(x_1, \dots, x_n) dx_1 + \dots + p_n(x_1, \dots, x_n) dx_n = 0 \end{cases}$$

¹⁵ En effet cette dimension n'est autre que le nombre m de « variables de départ » dans la définition de la lissité au voisinage de a .

¹⁶ On a noté Σ_a la relation déduite de Σ en fixant (x_1, \dots, x_n) à la valeur (a_1, \dots, a_n) .

dérive de R , et a évidemment la propriété 1.

Hyp. (H) Supposons que pour tout point $a = (a_1, \dots, a_n)$ du graphe de R l'un au moins des $p_i(a_1, \dots, a_n)$ soit $\neq 0$.

Alors l'ensemble des solutions (dx_1, \dots, dx_n) de l'équation :

$$p_1(a_1, \dots, a_n) dx_1 + \dots + p_n(a_1, \dots, a_n) dx_n = 0$$

est, pour tout a du graphe, un sous-espace vectoriel de dimension $n - 1$.

Si nous savions que R est lisse de codimension 1 nous pourrions donc en conclure que $\Sigma = DR$. Le § 3 va combler nos désirs, en montrant que sous les hypothèses ci-dessus (qu'il faudra un peu préciser) la relation R est effectivement lisse de codimension 1.

2. Interprétation géométrique

2.0. Interprétation de x_1, \dots, x_n comme « coordonnées »

Interprétation géométrique « naïve » Pour $n = 2$ ou 3 il est agréable d'interpréter les variables x_1, \dots, x_n comme les *coordonnées cartésiennes* de points du plan (cas $n = 2$) ou de l'espace à trois dimensions où nous vivons (cas $n = 3$).

On remarquera que parler des « coordonnées cartésiennes » d'un point n'a de sens qu'une fois choisi un *repère cartésien*, et que changer de repère se traduit en termes de coordonnées par un *changement de variables* qui en l'occurrence est non seulement différentiable mais *affine*, c'est-à-dire donné par des formules polynomiales de degré 1 :

$$\begin{cases} y_1 = a_{11}x_1 + \dots + a_{1n}x_n + b_1 \\ \dots \\ y_n = a_{n1}x_1 + \dots + a_{nn}x_n + b_n \end{cases}$$

Mais les coordonnées cartésiennes ne sont pas la seule façon de repérer la position des points : en géométrie euclidienne plane d'autres coordonnées très utilisées sont les *coordonnées polaires* ρ, θ , reliées aux coordonnées cartésiennes dans un repère orthonormé par la formule de changement de coordonnées rappelée dans l'exercice 1.4.

EXERCICE – En restreignant convenablement les domaines de variation de chacun des deux couples (ρ, θ) et (x, y) , s'assurer que la relation mentionnée ci-dessus est bien un changement de variables différentiable.

Interprétation canonique L'interprétation la plus usuelle des variables x_1, \dots, x_n consiste à considérer qu'elles sont censées désigner des nombres réels. Attribuer un n -uplet de valeurs à (x_1, \dots, x_n) , c'est donc choisir un « point » de \mathbb{R}^n . Dire « point » de \mathbb{R}^n (au lieu d'« élément » de \mathbb{R}^n) est une façon d'importer dans \mathbb{R}^n le langage géométrique suggéré par l'interprétation « naïve » précédente.

Voir \mathbb{R}^n comme un objet géométrique est tellement passé dans les mœurs que dire « faisons un dessin dans \mathbb{R}^2 » est devenu un tic de langage courant chez les mathématiciens¹⁷.

¹⁷ DEVINETTE Quelle est la différence entre un architecte, un tagueur, et un mathématicien ?

RÉPONSE L'architecte dessine sur sa table à dessin, le tagueur dessine sur un mur, le mathématicien dessine dans \mathbb{R}^2 !

Synthèse des deux interprétations précédentes Les règles du « calcul vectoriel » qu'on apprend au lycée, avec les règles du type « relation de Chasles » etc. qui régissent la correspondance $A, B \rightarrow \overrightarrow{AB}$ entre « points » et « vecteurs », se voient conférer un statut mathématique précis (en dimension n quelconque) par la notion d'espace affine à n dimensions¹⁸.

Dans un espace affine E à n dimensions on sait définir une notion de repère cartésien, d'où découle la notion de coordonnées cartésiennes d'un point $A \in E$ dans le repère considéré. Le choix d'un espace affine E et d'un tel repère fournit alors une interprétation de nos variables x_1, \dots, x_n comme les fonctions coordonnées dans le repère considéré ($x_i : E \rightarrow \mathbb{R}$ étant la fonction qui à chaque point $A \in E$ associe sa i -ème coordonnée $x_i(A)$ dans le repère considéré).

« L'interprétation canonique » évoquée plus haut peut être considérée comme un cas particulier de cette situation générale : on choisit $E = \mathbb{R}^n$, avec son « repère canonique » dans lequel la i -ème coordonnée du point $A = (a_1, \dots, a_n)$ n'est autre que a_i ! Dans cette interprétation, la i -ème fonction coordonnée $x_i : \mathbb{R}^n \rightarrow \mathbb{R}$ n'est autre que la i -ème projection canonique :

$$\begin{aligned} \text{pr}_i : \quad \mathbb{R}^n &\rightarrow \mathbb{R} \\ (a_1, \dots, a_n) &\mapsto a_i \end{aligned}$$

Dans le cas particulier $n = 1$ on arrive ainsi à l'interprétation un peu étrange de la variable x comme application identique de \mathbb{R} dans \mathbb{R} !

Interprétation géométrique des différentielles Ayant interprété les variables x_1, \dots, x_n comme coordonnées (cartésiennes) d'un point d'un espace affine à n dimensions, il est naturel d'interpréter le $2n$ -uplet de variables $(x_1, \dots, x_n, dx_1, \dots, dx_n)$ comme coordonnées d'un couple (A, \overrightarrow{dA}) , où :

$$\begin{aligned} A \in E &\text{ est le point de coordonnées } (x_1, \dots, x_n) \\ \overrightarrow{dA} \in \vec{E} &\text{ est le vecteur de coordonnées } (dx_1, \dots, dx_n) \end{aligned}$$

La donnée d'un tel couple (A, \overrightarrow{dA}) équivaut évidemment à la donnée du couple de points (A, A') tel que $A' = A + \overrightarrow{dA}$ (translaté du point A par la translation de vecteur \overrightarrow{dA}). On la représente graphiquement par une flèche d'origine A et d'extrémité $A' = A + \overrightarrow{dA}$. C'est ce que dans la terminologie « à l'ancienne » on appelle un **vecteur lié** (d'origine A)¹⁹.

2.1. Vecteur vitesse d'un point mobile

Supposons maintenant que le « point » x de coordonnées x_1, \dots, x_n soit « mobile », c'est-à-dire « fonction du temps ». Mathématiquement, cela revient à se donner une interprétation des variables x_1, \dots, x_n comme fonctions d'une variable réelle t :

$$x_i = x_i(t) \tag{2.1}$$

⁸ Cf. par exemple l'appendice A de mon livre « Les différentielles » (Masson 1996).

⁹ Il est dommage que ce terme soit tombé en désuétude. Dans notre contexte il exprime l'idée très importante que l'interprétation des variables dx_1, \dots, dx_n doit être « attachée » à celle des variables x_1, \dots, x_n .

En supposant que ces fonctions soient différentiables, « l'interprétation dérivée » (cf. § 1.2) peut s'écrire :

$$dx_i = v_i(t) dt \quad (2.1)'$$

où les $v_i(t)$ sont les dérivées des fonctions (2.1) :

$$v_i(t) = \frac{dx_i}{dt}(t)$$

Pour toute valeur de t , le vecteur $\vec{v}(t)$ de coordonnées $(v_1(t), \dots, v_n(t))$ est appelé **vecteur vitesse du point mobile au temps t** , et l'on écrit :

$$\vec{v}(t) = \frac{d\vec{x}}{dt}(t)$$

Dans la représentation graphique usuelle des vecteurs par des flèches, il est suggestif de prendre le point $x(t)$ comme origine de la flèche $\vec{v}(t)$ (cf. fig.3), ce qui revient à considérer le vecteur vitesse au temps t comme un *vecteur lié d'origine $x(t)$* (cf. fin du § 2.0).

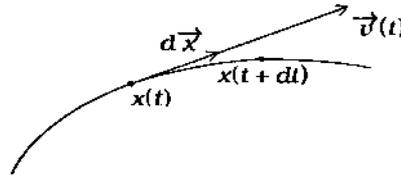


Figure 3. Vecteur vitesse d'un point mobile.

Soit maintenant f une fonction différentiable des variables x_1, \dots, x_n . Si dans l'expression de sa différentielle :

$$df = p_1(x_1, \dots, x_n) dx_1 + \dots + p_n(x_1, \dots, x_n) dx_n$$

nous substituons aux x_i , dx_i leurs interprétations (2.1), (2.1)', nous obtenons la différentielle de la fonction composée $f(x_1(t), \dots, x_n(t))$:

$$df = p_1(x_1(t), \dots, x_n(t)) v_1(t) dt + \dots + p_n(x_1(t), \dots, x_n(t)) v_n(t) dt$$

La dérivée de cette fonction composée est donc donnée par la formule :

$$\frac{df}{dt} = p_1(x_1(t), \dots, x_n(t)) v_1(t) + \dots + p_n(x_1(t), \dots, x_n(t)) v_n(t) \quad (2.2)$$

Pour toute valeur de t on reconnaît au second membre de (2.2) le nombre $df_{x(t)}(\vec{v}(t))$, dérivée de f au point $x(t)$ le long du vecteur $\vec{v}(t)$ (cf. § 1.1).

2.2. Interprétation géométrique de l'implication différentielle

La donnée d'une relation R entre n variables x_1, \dots, x_n équivaut à la donnée d'un sous-ensemble de \mathbb{R}^n , le *graphe* de la relation²⁰. De même la donnée d'une relation

²⁰ Cf. *Commentaire linguistique*, en fin de chapitre.

différentielle Σ entre les variables $x_1, \dots, x_n, dx_1, \dots, dx_n$ équivaut à la donnée d'un sous-ensemble de $\mathbb{R}^n \times \overrightarrow{\mathbb{R}^n}$: la notation $\overrightarrow{\mathbb{R}^n}$ est utilisée ici pour suggérer que les variables dx_1, \dots, dx_n soient interprétées comme coordonnées d'un vecteur, alors que x_1, \dots, x_n sont interprétées comme coordonnées d'un point ; autrement dit on interprétera les variables $x_1, \dots, x_n, dx_1, \dots, dx_n$ comme coordonnées d'un vecteur lié de \mathbb{R}^n (cf. § 2.0).

Étant donné un sous-ensemble $S \subset \mathbb{R}^n$, on appellera *point mobile sur S* un point mobile $x(t)$ dans \mathbb{R}^n tel que $x(t) \in S$ pour tout t . En lisant ce qui suit on se souviendra que le vecteur vitesse d'un point mobile est à considérer comme un vecteur lié.

PROPOSITION

L'implication différentielle $R \stackrel{d}{\Rightarrow} \Sigma$ (cf. § 1.2) admet la traduction géométrique suivante :

pour tout point mobile sur le graphe de R , les coordonnées du vecteur vitesse de ce point mobile à tout instant vérifient la relation Σ .

Preuve

Cela revient à dire qu'on ne restreint pas la généralité en ne considérant, dans la définition de $\stackrel{d}{\Rightarrow}$, que les interprétations φ de x_1, \dots, x_n comme fonctions différentiables d'une seule variable réelle t .

Notant $\stackrel{d}{\Rightarrow}$ cette version apparemment plus faible de l'implication différentielle, supposons donc que l'on ait $R \stackrel{d}{\Rightarrow} \Sigma$. Il s'agit de montrer que l'on a $R \stackrel{d}{\Rightarrow} \Sigma$, c'est-à-dire que pour toute interprétation φ de x_1, \dots, x_n comme fonctions différentiables de ℓ variables t_1, \dots, t_ℓ ,

$$R^\varphi \text{ identiquement vraie} \Rightarrow \Sigma^\varphi \text{ identiquement vraie}$$

Soit λ une interprétation de t_1, \dots, t_ℓ comme fonctions différentiables d'une variable s . Alors $\varphi \circ \lambda$ est une interprétation de x_1, \dots, x_n comme fonctions différentiables de s , et comme $R^{\varphi \circ \lambda} = (R^\varphi)^\lambda$ est identiquement vraie, notre hypothèse ($R \stackrel{d}{\Rightarrow} \Sigma$) implique que $\Sigma^{\varphi \circ \lambda}$ est identiquement vraie. En résumé,

pour toute interprétation λ de t_1, \dots, t_ℓ comme fonctions différentiables de s , la relation $(\Sigma^\varphi)^\lambda$ est identiquement vraie.

Il en résulte que Σ^φ est identiquement vraie²¹. □

Corollaire Dire qu'une relation différentielle a la propriété universelle (cf. début du § 1.3), c'est dire qu'elle est équivalente à l'énoncé suivant :

$$\text{« } x_1, \dots, x_n, dx_1, \dots, dx_n \text{ sont les coordonnées du vecteur vitesse d'un point mobile sur le graphe de } R \text{ »} \quad (\star_R)$$

¹ Pour le voir, remarquez que toute valeur du 2ℓ -uplet de variables $(t_1, \dots, t_\ell, dt_1, \dots, dt_\ell)$ peut s'interpréter comme vecteur vitesse d'un point mobile λ . Le fait que $(\Sigma^\varphi)^\lambda$ soit identiquement vraie implique que les coordonnées de ce vecteur vitesse vérifient Σ^φ .

Preuve

En effet cet énoncé (\star_R) « a la propriété universelle », car pour toute relation Σ dérivant de R , la proposition ci-dessus nous dit que $(\star_R) \Rightarrow \Sigma$. \square

Exemple (non lisse !) Soit R la relation $x^3 - y^2 = 0$ entre les deux variables réelles x et y . Alors la relation :

$$DR : \begin{cases} x^3 - y^2 = 0 \\ 3x^2 dx - 2y dy = 0 \\ \text{si } (x, y) = (0, 0), \text{ alors } (dx, dy) = (0, 0) \end{cases}$$

a la propriété universelle.

Preuve

La relation R est lisse en tout point, *sauf* en $(0, 0)$: en effet (cf. fig. 4),

i) elle est de type $\text{Dif}(x|y)$ en dehors de $y = 0$, donc au voisinage de tout point du graphe *sauf* $(0, 0)$ (la fonction $y^{2/3}$ n'est pas différentiable en 0) ;

ii) elle n'est pas de type $\text{Dif}(y|x)$ au voisinage de $(0, 0)$ (même restreinte à de petits intervalles ouverts contenant 0 elle n'est pas fonctionnelle en y).

La formule pour DR en dehors de $(0, 0)$ se démontre comme dans l'exemple 1.6. Pour justifier la formule en $(0, 0)$, il suffit d'après le corollaire 2.1 de faire l'exercice ci-après. \square

EXERCICE 2.1.— Montrer que sur la courbe plane d'équation $x^3 - y^2 = 0$ tout point mobile passant par l'origine (disons, à l'instant $t = 0$) a une vitesse nulle à cet instant (idée : raisonner sur les développements limités à l'ordre 1 des fonctions $x(t)$ et $y(t)$).

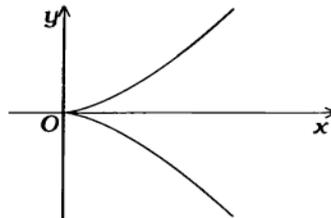


Figure 4. Graphe de la relation $x^3 - y^2 = 0$.

2.3. Les sous-variétés et leurs espaces tangents

Un sous-ensemble $S \subset \mathbb{R}^n$ est dit **lisse** (en un point a) si c'est le graphe d'une relation R qui est lisse (au point a). Au lieu de dire « ensemble lisse » (en tout point) on dit aussi **sous-variété**, en précisant le cas échéant **de classe C^f** .

À la notion de *dimension* (ou de *codimension*) d'une relation lisse correspond celle de dimension ou codimension d'une sous-variété. En particulier un sous-ensemble lisse de dimension 1 est appelé **courbe lisse** ; un sous-ensemble lisse de dimension 2 est appelé **surface lisse** ; un sous-ensemble lisse de codimension 1 est appelé

hypersurface lisse (de sorte que la notion d'hypersurface lisse coïncide dans \mathbb{R}^2 avec celle de courbe lisse et dans \mathbb{R}^3 avec celle de surface lisse).

CONVENTIONS Une sous-variété de dimension 0 est un ensemble de points isolés ; une sous-variété de codimension 0 est un ouvert de \mathbb{R}^n .

Le graphe de la différentielle DR est appelé **fibré tangent à la sous-variété** S , et noté TS . Ses éléments peuvent s'interpréter comme des *vecteurs liés* (cf. fin du § 2.0) appelés **vecteurs tangents à S** (on dira « vecteur tangent à S en a » lorsqu'on voudra préciser que le vecteur lié considéré a pour origine le point a). Comme nous l'avons vu au § 2.2, TS est l'ensemble des vecteurs vitesses de points mobiles sur S .

L'origine a d'un vecteur tangent à S est nécessairement un point de S . Etant donné $a \in S$, l'ensemble des vecteurs tangents à S en a est naturellement muni d'une structure d'espace vectoriel, dont la dimension m n'est autre que la dimension de S (le point a étant fixé, les opérations sur les couples $(a, \vec{u}) \in \{a\} \times \vec{\mathbb{R}}^n$ sont celles de l'espace vectoriel $\vec{\mathbb{R}}^n$).

Cet espace vectoriel est noté $T_a S$, et appelé **espace tangent à S en a** . On précise parfois *espace vectoriel tangent à S en a* , pour éviter de le confondre avec l'*espace affine tangent à S en a* , ensemble des points de \mathbb{R}^n qui sont les extrémités de vecteurs tangents à S en a (points de la forme $a + \vec{u}$, avec $(a, \vec{u}) \in T_a S$).

Cas particulier $T_a \mathbb{R}^n$ est l'espace vectoriel à n dimensions des vecteurs liés d'origine a . Les expressions « vecteur tangent à \mathbb{R}^n » et « vecteur lié de \mathbb{R}^n » sont donc synonymes. Plus généralement, pour tout ouvert $U \subset \mathbb{R}^n$, un « vecteur tangent à U » est un vecteur lié de \mathbb{R}^n dont l'origine est un point de U .

Autre interprétation de l'espace tangent $T_a S$ On peut aussi interpréter $T_a S$ comme « ce qu'on voit lorsqu'on regarde S avec une loupe centrée en a , de grossissement infini ». Plus précisément, introduisons un petit paramètre $\varepsilon > 0$, et appelons « zoom de grossissement $\frac{1}{\varepsilon}$, centré en a », le changement de variables $x_i := a_i + \varepsilon dx_i$ ($i = 1, \dots, n$). Soit $S_{a,\varepsilon}$ le sous-ensemble de la boule de rayon 1 ($\|dx\| < 1$), image de S par ce « zoom ». Alors quand ε tend vers 0 ce sous-ensemble $S_{a,\varepsilon}$ « tend vers » $\vec{T}_a S$, image de $T_a S$ par l'isomorphisme évident $\{a\} \times \vec{\mathbb{R}}^n \approx \vec{\mathbb{R}}^n$ (pour les détails, cf. mon livre « Les différentielles » (Masson éd., 1996), chap.6 § 1.3).

2.4. Systèmes de coordonnées locales sur une sous-variété

Point de vue restreint

Considérons à nouveau « l'interprétation canonique » des variables x_1, \dots, x_n , (cf. § 2.0), mais en la restreignant à une sous-variété $S \subset \mathbb{R}^n$, de dimension m : autrement dit la variable x_i est maintenant interprétée comme fonction sur S :

$$x_i = \text{pr}_i|_S : S \rightarrow \mathbb{R}, \quad \text{où } \text{pr}_i : \mathbb{R}^n \rightarrow \mathbb{R} \text{ est la } i\text{-ème projection canonique.}$$

Soit $a \in S$, et soit $U \subset \mathbb{R}^n$ un voisinage de a dans lequel S est le graphe d'une relation ayant la propriété $\text{Diff}(x_{j_1}, \dots, x_{j_p} | x_{i_1}, \dots, x_{i_m})$ (cf. § 1.4)²². Pour se donner un point de $S \cap U$ il suffit de se donner une valeur du m -uplet $(x_{i_1}, \dots, x_{i_m})$, valeur qui peut être choisie arbitrairement dans un voisinage \underline{U} de $(a_{i_1}, \dots, a_{i_m})$. Autrement dit, le m -uplet de variables $(x_{i_1}, \dots, x_{i_m})$ définit, dans l'interprétation ci-dessus, une bijection :

$$x = (x_{i_1}, \dots, x_{i_m}) : S \cap U \rightarrow \underline{U} \subset \mathbb{R}^m$$

DÉFINITION Ainsi interprété, le m -uplet de variables $(x_{i_1}, \dots, x_{i_m})$ est appelé **système de coordonnées sur $S \cap U$** , ou **système de coordonnées locales sur S au voisinage de a** . La bijection x est appelée **carte de $S \cap U$** , ou **carte locale de S au voisinage de a** .

Grâce à la bijection $x : S \cap U \rightarrow \underline{U}$, la donnée d'une fonction f sur $S \cap U$ équivaut à la donnée d'une fonction \underline{f} sur \underline{U} ²³, c'est-à-dire une fonction des m variables réelles x_{i_1}, \dots, x_{i_m} , définie dans le voisinage \underline{U} du point $\underline{a} = x(a)$.

DÉFINITION On dit alors que la fonction \underline{f} (des variables x_{i_1}, \dots, x_{i_m}) est l'**expression de la fonction f dans le système de coordonnées $(x_{i_1}, \dots, x_{i_m})$** . Une fonction f sur S est dite **différentiable au voisinage de a** si son expression dans un système de coordonnées locales $(x_{i_1}, \dots, x_{i_m})$ est différentiable.

Remarquons que cette définition n'est pas affectée par la liberté dont nous disposons dans le choix du système de coordonnées locales $(x_{i_1}, \dots, x_{i_m})$ (en effet les changements de coordonnées locales sont différentiables, de la classe C^r donnée par le contexte).

Point de vue élargi

Pour tout changement de variables différentiable $(x_{i_1}, \dots, x_{i_m}) \leftrightarrow (t_1, \dots, t_m)$, transformant l'ouvert \underline{U} en un ouvert $\tilde{\underline{U}}$ de \mathbb{R}^m , l'interprétation ci-dessus des x_i fournit une interprétation de t_1, \dots, t_m comme fonctions sur $S \cap U$, définissant une bijection :

$$t = (t_1, \dots, t_m) : S \cap U \rightarrow \tilde{\underline{U}} \subset \mathbb{R}^m$$

et tout ce qui a été dit de $(x_{i_1}, \dots, x_{i_m})$ peut être répété avec (t_1, \dots, t_m) .

DÉFINITION Ainsi interprété, le m -uplet de variables (t_1, \dots, t_m) est appelé **système de coordonnées sur $S \cap U$** , ou **système de coordonnées locales sur S au voisinage de a** . La bijection t est appelée **carte de $S \cap U$** , ou **carte locale de S au voisinage de a** .

À l'aide de la carte locale t on définit comme ci-dessus la notion d'expression d'une fonction f (sur S) dans le système de coordonnées locales (t_1, \dots, t_m) . Une fonction f sur

²² Dans toute cette section la « différentiabilité » pourra s'entendre au sens de différentiabilité de classe C^r , r étant un entier ≥ 1 fixé une fois pour toutes.

²³ La correspondance entre ces fonctions s'écrit $\underline{f} = f \circ x^{-1}$, et inversement $f = \underline{f} \circ x$; autrement dit la fonction f se déduit de la fonction de m variables réelles $\underline{f}(x_{i_1}, \dots, x_{i_m})$ en réinterprétant les variables x_{i_1}, \dots, x_{i_m} comme fonctions sur $S \cap U$.

S est différentiable au voisinage de a si et seulement si son expression dans un système de coordonnées locales est différentiable.

Paramétrages locaux

Revenons maintenant aux variables x_1, \dots, x_n , dans leur interprétation comme fonctions sur S . Ces fonctions sont évidemment différentiables au sens précédent (relire, au § 1.4, la première phrase de la preuve de la première proposition). Leurs expressions dans les coordonnées t_1, \dots, t_m ci-dessus sont donc des fonctions différentiables des variables t_1, \dots, t_m :

$$\begin{cases} x_1 = x_1(t_1, \dots, t_m) \\ \dots \\ x_n = x_n(t_1, \dots, t_m) \end{cases}$$

DÉFINITION Une telle interprétation des variables x_1, \dots, x_n (comme fonctions différentiables des variables t_1, \dots, t_m) est appelée **paramétrage local** de S (au voisinage de a).

Exemple ($n = 2$) Considérons le cercle S d'équation $x^2 + y^2 = 1$, au voisinage du point a de coordonnées $(x = 1, y = 0)$. Interprétée comme fonction sur S , la variable y est une coordonnée locale²⁴ sur S au voisinage de a : en effet la relation $x^2 + y^2 = 1$, restreinte à la « bande » $U : x > 0, -1 < y < 1$ est fonctionnelle différentiable (C^∞) en x , de sorte que l'application :

$$y : S \cap U \rightarrow]-1, +1[\subset \mathbb{R}$$

est bijective. La variable y ainsi interprétée est donc bien une coordonnée (C^∞) sur $S \cap U$ au sens restreint.

Un autre exemple de coordonnée C^∞ sur $S \cap U$ est fourni par la coordonnée angulaire θ , déduite de y par le changement de variable différentiable (C^∞) $\theta = \arcsin y$. Elle fournit une bijection :

$$\theta : S \cap U \rightarrow]-\frac{\pi}{4}, +\frac{\pi}{4}[$$

Il s'agit cette fois d'une coordonnée locale au sens élargi. Le paramétrage de S qui lui correspond est évidemment :

$$\begin{cases} x = \cos \theta \\ y = \sin \theta \end{cases}$$

Remarquons qu'en l'occurrence ce paramétrage est défini pour tout $\theta \in \mathbb{R}$ (et pas seulement pour $-\frac{\pi}{4} < \theta < +\frac{\pi}{4}$), et que la variable θ restreinte à un quelconque intervalle de longueur $< 2\pi$ peut ainsi être interprétée comme une coordonnée sur l'arc de cercle, image de cet intervalle par ce paramétrage.

²⁴ Comme la dimension m est 1, on ne dit pas « système de coordonnées » mais simplement « coordonnée ».

2.5. Calcul différentiel sur une sous-variété

La donnée d'un système de coordonnées locales sur une sous-variété S de dimension m permet de ramener localement l'étude des fonctions différentiables sur S à l'étude des fonctions différentiables de m variables, et de faire ainsi du *calcul différentiel sur S* . Nous allons montrer ici qu'on peut aussi présenter le calcul différentiel sur S comme restriction à S du *calcul différentiel dans l'espace ambiant \mathbb{R}^n* .

Lemme de restriction La restriction à S d'une fonction différentiable²⁵ dans un ouvert U de \mathbb{R}^n est une fonction différentiable sur $S \cap U$.

Réciproquement toute fonction différentiable sur S est localement²⁶ la restriction d'une fonction différentiable dans un ouvert de \mathbb{R}^n : pour toute fonction f différentiable sur S au voisinage d'un point a , il existe une fonction \check{f} différentiable dans un ouvert U de \mathbb{R}^n contenant a , telle qu'au voisinage de a on ait $f = \check{f}|_{S \cap U}$.

Preuve

Pour la première affirmation, on pourra utiliser un paramétrage local de S . Pour la réciproque on utilisera un système de coordonnées locales (x_1, \dots, x_m) (au sens restreint) ; toute fonction f sur S s'écrit localement sous la forme $\underline{f}(x_1, \dots, x_m)$, que l'on peut aussi interpréter comme une fonction \check{f} définie dans un ouvert U de \mathbb{R}^n (interpréter x_i comme la i -ème coordonnée canonique dans \mathbb{R}^n !) : la fonction \check{f} ainsi définie est indépendante des « autres » coordonnées x_{j_1}, \dots, x_{j_p} ; elle est différentiable si f (donc \underline{f}) l'était ; son ouvert U de définition est l'ensemble des valeurs de (x_1, \dots, x_n) pour lesquelles $(x_1, \dots, x_m) \in \underline{U}$, ouvert de définition de la fonction \underline{f} . \square

Notons qu'on dispose d'une grande liberté dans le choix de la fonction \check{f} : on peut lui rajouter n'importe quelle fonction différentiable \check{g} dont la restriction à S s'annule, et il y en a beaucoup !

Considérant ainsi les fonctions différentiables sur S comme restrictions de fonctions différentiables dans l'espace ambiant, nous allons maintenant définir les *formes différentielles sur S* comme restrictions de formes différentielles dans l'espace ambiant. Rappelons (cf. § 1.1) qu'une forme différentielle de n variables x_1, \dots, x_n est une expression formelle du type (1.1). Si l'on interprète $(x_1, \dots, x_n, dx_1, \dots, dx_n)$ comme coordonnées de vecteurs tangents à \mathbb{R}^n , toute forme différentielle ω s'interprète comme une fonction sur l'espace tangent $T\mathbb{R}^n$, et il est naturel de s'intéresser à la restriction de cette fonction à TS , restriction qu'au lieu de noter $\omega|_{TS}$ on a coutume de noter simplement $\omega|_S$.

Lemme de différentiation $\check{f}|_S = 0 \Rightarrow (d\check{f})|_S = 0$.

Preuve

Très semblable au lemme "d" (§ 1.2), ce lemme est une conséquence directe du principe de substitution. À titre d'« exercice de style », détaillons la preuve dans le

²⁵ Dans ce qui suit, l'adjectif « différentiable » peut s'entendre comme *différentiable de classe C^r* ($r \geq 1$ fixé une fois pour toutes).

²⁶ Nous verrons plus tard (chap. 2 § 3.3) que c'est même vrai *globalement*.

langage du § 1.2, en notant R la relation dont S est le graphe (concrètement, on pourra imaginer R sous la forme d'un système d'équations de S).

L'hypothèse $\check{f}|_S = 0$ se traduit évidemment par l'implication ($R \Rightarrow (\check{f} = 0)$). Or (lemme « d ») ($\check{f} = 0$) $\stackrel{d}{\Rightarrow} d\check{f} = 0$. Donc²⁷ $R \stackrel{d}{\Rightarrow} (d\check{f} = 0)$, de sorte que $DR \Rightarrow (d\check{f} = 0)$ (propriété universelle de DR), ce qui veut bien dire que $(d\check{f})|_{TS} = 0$. \square

Ce lemme montre que $(d\check{f})|_{TS}$ ne dépend que de $f = \check{f}|_S$, ce qui permet d'énoncer la définition suivante.

DÉFINITION (différentielle d'une fonction sur S) Soit f une fonction différentiable sur S . On appelle **différentielle** de f la fonction sur TS définie localement²⁸ par :

$$df = (d\check{f})|_{TS} (= d\check{f}|_S \text{ dans les notations consacrées})$$

où \check{f} est une fonction telle que $f = \check{f}|_S$.

Pour toute valeur fixée a de (x_1, \dots, x_n) (avec $a \in S$) cette fonction df définit une fonction linéaire :

$$df_a : T_a S \rightarrow \mathbb{R}$$

appelée **différentielle de f au point a** .

Énonçons encore une définition, qui sera exploitée dans l'étude 1 :

DÉFINITION (point critique d'une fonction sur S) On appelle **point critique** d'une fonction différentiable f un point a où $df_a = 0$.

3. Théorème des fonctions implicites

On se propose ici de donner des conditions suffisantes simples pour qu'une relation entre n variables x_1, \dots, x_n soit, localement au voisinage d'un point a de son graphe, fonctionnelle différentiable (cf. § 1.4) en certaines de ces variables.

3.1. Fonctions implicites en codimension 1

Théorème des fonctions implicites (codimension 1)

Soit R la relation $s(x_1, \dots, x_n) = 0$, où s est une fonction de classe C^r ($r \geq 1$). Supposons qu'en un point $a = (a_1, \dots, a_n)$ du graphe de R la forme linéaire :

$$ds_a = p_1(a) dx_1 + \dots + p_n(a) dx_n$$

ne soit pas nulle. Alors la relation R est lisse en a de classe C^r , de codimension 1. Plus précisément si i est un indice tel que $p_i(a) \neq 0$, il existe un voisinage de a dans lequel la relation R est fonctionnelle différentiable en x_i , de classe C^r . \square

²⁷ Il résulte immédiatement de la définition de $\stackrel{d}{\Rightarrow}$ que $[(R \Rightarrow R' \text{ et } R' \stackrel{d}{\Rightarrow} \Sigma) \text{ implique } (R \stackrel{d}{\Rightarrow} \Sigma)]$
²⁸ (dans la mesure où notre construction ci-dessus n'est que locale)

EXERCICE 3.1.—

- i) Démontrer que l'exemple de l'exercice 0.1 est une courbe lisse.
- ii) Discuter la lissité, en ses différents points, du sous-ensemble de \mathbb{R}^2 donné par l'équation $y^2 - x^k = 0$, où k est un entier ≥ 1 . Le dessiner pour diverses valeurs de k .
- iii) Même question pour $y^2 + x^k = 0$.
- iv) Même question pour $y^4 - 2xy^2 + x^2 = 0$.
- v) Même question pour le sous-ensemble de \mathbb{R}^3 d'équation $x^2 + y^2 - z^2 = 0$.

EXERCICE 3.2.— Une fonction f de n variables x_1, \dots, x_n est dite *homogène de degré* k (k réel > 0) si elle vérifie l'identité :

$$f(tx_1, \dots, tx_n) = t^k f(x_1, \dots, x_n) \quad (\text{pour tout } t \text{ réel } > 0) \quad (\star)$$

- i) Dédurre de (\star) l'identité d'Euler :

$$\sum_{i=1}^n x_i \frac{\partial f}{\partial x_i} = kf$$

(idée : en interprétant les x_i comme des constantes, les deux membres de (\star) définissent une fonction \tilde{f} de la variable t ; on dérivera cette fonction, puis on fera $t = 1$).

- ii) Soit f un polynôme *homogène de degré* k ($k \in \mathbb{N}$) en les variables x_1, \dots, x_n . Montrer que ses ensembles de niveau (définis par l'équation $f(x_1, \dots, x_n) = c$) sont, pour $c \neq 0$, des *hypersurfaces lisses de classe* C^∞ (pour un exemple de cette situation, cf. exercice 0.3).
Que pouvez-vous dire du cas $c = 0$? (pour un exemple de cette situation, cf. exercice 3.1v).
- iii) Dédurre de ii) que l'ensemble $S\ell(n)$ des matrices $n \times n$ de déterminant 1 est une sous-variété de \mathbb{R}^{n^2} .

3.2. Fonctions implicites en codimension p

La généralisation du théorème précédent nous conduira à considérer un système de p équations linéaires homogènes à n inconnues dx_1, \dots, dx_n .

Rappel d'algèbre linéaire Considérons le système linéaire suivant, à n inconnues (rebaptisées X_1, \dots, X_n) :

$$\begin{cases} a_{11}X_1 + \dots + a_{1n}X_n = 0 \\ \dots \\ a_{p1}X_1 + \dots + a_{pn}X_n = 0 \end{cases} \quad (3.0)$$

Si le rang de ce système est égal au nombre p d'équations, il existe un partage des inconnues X_1, \dots, X_n en « inconnues principales » X_{j_1}, \dots, X_{j_p} et « inconnues secondaires »

X_{i_1}, \dots, X_{i_m} ($m = n - p$) tel que le système linéaire (3.0) puisse être résolu en les inconnues principales, donnant celles-ci comme fonctions linéaires des inconnues secondaires : plus précisément un tel partage des inconnues convient si et seule-

ment si la matrice $\begin{pmatrix} a_{1j_1} & \dots & a_{1j_p} \\ \dots & \dots & \dots \\ a_{pj_1} & \dots & a_{pj_p} \end{pmatrix}$ est inversible.

On résumera cette situation en disant que **la relation linéaire (3.0) est fonctionnelle (linéaire) en X_{j_1}, \dots, X_{j_p}** .

La situation qui va nous intéresser ici sera la suivante. On considère un système d'équations à n inconnues x_1, \dots, x_n :

$$\begin{cases} s_1(x_1, \dots, x_n) = 0 \\ \dots \\ s_p(x_1, \dots, x_n) = 0 \end{cases} \quad (3.1)$$

où les s_i sont des fonctions de classe C^r ($r \geq 1$). En remplaçant les s_i par leurs différentielles on obtient un système linéaire homogène en dx_1, \dots, dx_n , à coefficients fonctions de x_1, \dots, x_n :

$$\begin{cases} ds_1 = 0 \\ \dots \\ ds_p = 0 \end{cases} \quad (3.1)'$$

En donnant à (x_1, \dots, x_n) dans (3.1)' la valeur $a = (a_1, \dots, a_n)$ on obtient un système linéaire homogène à n inconnues dx_1, \dots, dx_n , à coefficients constants :

$$\begin{cases} ds_{1a} = 0 \\ \dots \\ ds_{pa} = 0 \end{cases} \quad (3.1)''_a$$

Théorème des fonctions implicites (cas général) *Supposons que les formes linéaires ds_{1a}, \dots, ds_{pa} soient linéairement indépendantes, c'est-à-dire que le système linéaire (3.1)''_a soit de rang p . Alors la relation (3.1) est lisse en a de classe C^r , de codimension p . Plus précisément si la relation linéaire homogène (3.1)''_a est fonctionnelle en $dx_{j_1}, \dots, dx_{j_p}$, il existe un voisinage de a dans lequel la relation (3.1) est fonctionnelle différentiable en x_{j_1}, \dots, x_{j_p} , de classe C^r . \square*

Corollaire *Notons R la relation définie par la formule (3.1), et Σ la relation définie par ((3.1) et (3.1)'). Alors Σ n'est autre que la différentielle DR de R , au sens du § 1.4.*

Preuve

Cette relation Σ vérifie évidemment les propriétés caractéristiques 1 et 2bis du § 1.4. \square

EXERCICE 3.3.— Soit $S \subset \mathbb{R}^3$ la sphère unité $\langle x, x \rangle = 1$, où $\langle \cdot, \cdot \rangle$ désigne le produit scalaire euclidien. Montrer que :

$$V = \{ (x, y) \in S \times S \mid \langle x, y \rangle = 0 \}$$

est une sous-variété de \mathbb{R}^6 .

EXERCICE 3.4.— Dans l'espace M des matrices 2×3 à coefficients réels, espace identifié à \mathbb{R}^6 par la convention :

$$\begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix} \leftrightarrow (x_1, x_2, x_3, y_1, y_2, y_3)$$

on note S_r le sous-ensemble des matrices de rang exactement égal à r . Montrer que S_r ($r = 0, 1, 2$) est une sous-variété dont on déterminera la codimension (indication pour S_1 : dans l'ouvert U_i ($i = 1, 2, 3$) des points de M où $\begin{pmatrix} x_i \\ y_i \end{pmatrix}$, S_1 peut être défini par l'annulation de deux mineurs convenablement choisis).

Cas particulier : changements de variables locaux en dimension p

Considérons une relation de la forme :

$$\begin{cases} y_1 = g_1(x_1, \dots, x_p) \\ \dots \\ y_p = g_p(x_1, \dots, x_p) \end{cases} \quad (3.2)$$

donnant les variables y_1, \dots, y_p comme fonctions différentiables (de classe C^r) des variables x_1, \dots, x_p .

Théorème d'inversion locale Supposons qu'en $a = (a_1, \dots, a_p)$ les différentielles dg_{1a}, \dots, dg_{pa} soient linéairement indépendantes. Alors il existe dans \mathbb{R}^p un voisinage U de a et un voisinage V de $b = g(a)$ tels que dans $U \times V$ la relation (3.2) soit un changement de variables différentiable, de classe C^r . \square

On résume cette situation en disant que la relation (3.2) est un *changement de variables local* (au voisinage de $x = a$, $y = b$) de classe C^r .

N.B. Dans le cas $p = 1$ cela revient à affirmer simplement que si une fonction g d'une variable, de classe C^r , vérifie $g'(a) \neq 0$, alors elle admet localement une fonction réciproque de classe C^r .

3.3. Déformations des racines simples d'un système d'équations

Rappelons qu'une *racine simple* d'une équation à une inconnue $s(x) = 0$ est une valeur a de x pour laquelle $s(a) = 0$, $s'(a) \neq 0$. Plus généralement on appelle *racine simple d'un système de p équations à p inconnues* :

$$\begin{cases} s_1(x_1, \dots, x_p) = 0 \\ \dots \\ s_p(x_1, \dots, x_p) = 0 \end{cases} \quad (3.3)$$

une valeur a de (x_1, \dots, x_p) vérifiant (3.3), et telle que les différentielles ds_{1a}, \dots, ds_{pa} soient linéairement indépendantes (ce qui revient à dire que la matrice des dérivées partielles est inversible).

Permettons maintenant aux fonctions s_i de dépendre aussi de paramètres t_1, \dots, t_m , en étant fonctions différentiables de classe C^r ($r \geq 1$) de $(x_1, \dots, x_p, t_1, \dots, t_m)$. Autrement dit au lieu de considérer un système comme (3.3) on en considère une infinité paramétrée par $t = (t_1, \dots, t_m)$:

$$\begin{cases} s_1(x_1, \dots, x_p, t_1, \dots, t_m) = 0 \\ \dots \\ s_p(x_1, \dots, x_p, t_1, \dots, t_m) = 0 \end{cases} \quad (3.4)$$

Théorème de stabilité des racines simples *Supposons que pour $t = (0, \dots, 0)$ (disons) le système (3.4) ait $a = (a_1, \dots, a_n)$ comme racine simple. Alors il existe un voisinage U de a dans \mathbb{R}^p tel que pour tout t assez voisin de 0 (disons $\|t\| < \varepsilon$) le système d'équations (3.4) admette dans U une unique racine $a(t)$, qui de plus est simple, et est fonction de classe C^r de t .*

Preuve

Ce n'est là qu'une paraphrase du théorème des fonctions implicites (x_1, \dots, x_p jouant le rôle de x_{j_1}, \dots, x_{j_p} , et t_1, \dots, t_m celui de x_{i_1}, \dots, x_{i_m}). Le fait que $a(t)$ reste une racine simple pour t assez proche de 0 se voit en remarquant que si une matrice carrée, dépendant continûment de t , est inversible pour $t = 0$, elle l'est aussi pour tout t assez proche de 0 (en effet son déterminant est fonction continue de t). \square

EXERCICE 3.5.— Discuter, en fonction des paramètres réels p et q , le nombre et la multiplicité des racines de l'équation générale du troisième degré (à une inconnue) $x^3 + px + q = 0$ (idée : raisonner sur le graphe de la fonction $y = x^3 + px$). On résumera le résultat de la discussion par un dessin dans le plan des (p, q) , et l'on essaiera de définir dans ce plan des ouverts, aussi grands que possible, dans lesquels une racine convenablement choisie est fonction C^∞ de (p, q) ²⁹.

Nous reviendrons sur cet exemple dans l'étude 2 (exercice 1). Nous rencontrerons d'autres exemples dans les études 1 et 1^{bis}.

ÉTUDE 1

Points critiques et extrema locaux des fonctions

Rappelons qu'on appelle *point critique* d'une fonction différentiable un point a tel que $df_a = 0$ (la forme linéaire nulle). L'un des intérêts pratiques de cette notion tient dans le lemme suivant.

Lemme 1.— *Si f , différentiable en a , admet un extremum local³⁰ en a , alors a est un point critique de f .*

Preuve

La preuve, qui n'utilise l'hypothèse de différentiabilité qu'au seul point a , est facile dans le cas $n = 1$ ³¹. Le cas général s'y ramène facilement : supposant que f admette un extremum local en a , il s'agit de montrer que $df_a = 0$, c'est-à-dire que $df_a(\vec{v}) = 0$ pour tout vecteur \vec{v} ; or tout vecteur \vec{v} est le vecteur vitesse d'un point mobile

²⁹ Pour une réponse à ces questions, cf. mon livre « Les différentielles », chap. 5 § 1.2, où l'on trouvera d'autres exercices analogues.

³⁰ Rappelons que par « extremum local » on entend soit un maximum local ($f(x) \leq f(a)$ pour tout x dans un voisinage convenable de a), soit un minimum local ($f(x) \geq f(a)$ pour tout x dans un voisinage convenable de a).

³¹ Même dans ce cas elle recèle néanmoins une subtilité : cf. « Les différentielles », chap. 0, exercices 4.1 et 4.2.

$x := \lambda(t)$ (par exemple en mouvement rectiligne uniforme !), et l'on a :

$$df_a(\vec{v}) = (f \circ \lambda)'(0)$$

(forme condensée de la formule (2.2)) ; or $f \circ \lambda$ a évidemment un extremum local en 0, de sorte que $(f \circ \lambda)'(0) = 0$. \square

Remarque La réciproque du lemme est évidemment fautive (pensez par exemple à la fonction d'une seule variable $f(x) = x^3$, ou à la fonction de deux variables $f(x_1, x_2) = x_1^2 - x_2^2$).

Forme hessienne d'un point critique

Au voisinage d'un point critique a de la fonction f (supposée deux fois différentiable au moins), celle-ci admet un développement limité à l'ordre 2 qui s'écrit :

$$f(a + dx) = f(a) + \frac{1}{2} d^2 f_a + o(\|dx\|^2)$$

où $d^2 f_a$ est la *forme quadratique* (polynôme homogène de degré 2) en dx_1, \dots, dx_n dont nous rappelons ici l'expression (cf. formule (1.3)) :

$$d^2 f_a = \sum_{i,j} p_{ij}(a) dx_i dx_j, \quad \text{avec } p_{ij}(a) = \frac{\partial^2 f}{\partial x_i \partial x_j}(a)$$

DÉFINITION 1.— Cette forme quadratique (la différentielle seconde de f en a) est appelée la **hessienne** du point critique³². Le point critique est dit **quadratique non dégénéré** si le rang de sa hessienne est égal au nombre de variables.

L'hypothèse que a était un point critique a une conséquence remarquable. Soit \tilde{f} la fonction des variables u_1, \dots, u_ℓ déduite de f par les substitutions $x_i := \varphi_i(u_1, \dots, u_\ell)$, où les φ_i sont des fonctions deux fois différentiables définies au voisinage de $u = 0$, telles que $a = (\varphi_1(0, \dots, 0), \dots, \varphi_n(0, \dots, 0))$. Cette fonction \tilde{f} a évidemment l'origine pour point critique, et sa hessienne est donnée par le lemme suivant.

Lemme 2.— La hessienne de \tilde{f} en 0 se déduit de celle de f en a par les substitutions « évidentes » $dx_i := d\varphi_{i0}$.

Preuve

Il suffit d'appliquer la formule (1.5) (§ 1.3.3), en remarquant que dans le cas présent $p_i(a) = 0$! \square

Il en résulte en particulier que la question

comment la hessienne se transforme-t-elle par un changement de variables différentiable ?

se réduit à la question bien connue d'algèbre linéaire

³² Sa matrice $(p_{ij}(a))$ est appelée *matrice hessienne* du point critique

comment une forme quadratique se transforme-t-elle par un changement de variables linéaire ?

Rappelons le résultat central de la classification des formes quadratiques :

Rappel sur les formes quadratiques Soit Q une forme quadratique à n indéterminées X_1, \dots, X_n ; alors il existe un changement de variables linéaire $(X_1, \dots, X_n) \leftrightarrow (U_1, \dots, U_n)$ qui transforme Q en :

$$\tilde{Q} = \sum_{i=1}^r \varepsilon_i U_i^2$$

où les ε_i sont des scalaires³³ non nuls, et r est le rang de la forme quadratique (égal à n dans le cas non dégénéré). Si le corps des scalaires est \mathbb{C} on peut toujours se ramener (par extraction de racines carrées) au cas où tous les ε_i valent 1. Si le corps des scalaires est \mathbb{R} on peut se ramener au cas où $\varepsilon_i = \pm 1$; la forme quadratique Q est alors caractérisée, à congruence près³⁴, par sa signature (n_+, n_-) , où n_+ (resp. n_-) désigne le nombre d'indices i pour lesquels $\varepsilon_i = +1$ (resp. -1).

Corollaire 1.— Il existe, au voisinage de tout point critique a de f , un changement de variables affine $(x_1, \dots, x_n) \leftrightarrow (u_1, \dots, u_n)$ tel que la hessienne de la fonction transformée \tilde{f} s'écrive :

$$d^2 \tilde{f}_0 = \sum_{i=1}^r \varepsilon_i u_i^2$$

où r est le rang de la hessienne, et les ε_i sont comme dans le rappel ci-dessus.

Preuve

Avec les notations du « rappel », il suffit de poser $u_i(x_1, \dots, x_n) = U_i(X_1, \dots, X_n)$, avec $X_i = x_i - a_i$. □

La figure 5 représente, dans le cas $n = r = 2$, les courbes de niveau d'une telle fonction \tilde{f} (en interprétant (u_1, u_2) comme un système de coordonnées orthonormé dans le plan de la feuille). De tels dessins peuvent être considérés comme des modèles « redressés » de l'allure des courbes de niveau sur une carte d'état-major, au voisinage d'un sommet, d'une cuvette, ou d'un col.

³³ L'énoncé est vrai quel que soit le corps des scalaires, pourvu que la « caractéristique » de ce corps soit $\neq 2$.

³⁴ Deux formes quadratiques sont dites *congruentes* si l'on passe de l'une à l'autre par un changement de variables linéaire.

Signature	Modèle local	Dénomination	Dessin des lignes de niveau
(2, 0)	$u_1^2 + u_2^2$	Minimum quadratique ou « cuvette »	
(0, 2)	$-u_1^2 - u_2^2$	Maximum quadratique ou « sommet »	
(1, 1)	$u_1^2 - u_2^2$	Point selle ou « col »	

Figure 5.

Modèles locaux de points critiques quadratiques non dégénérés en dimension 2.

EXERCICE 1.— Étudiez les lignes de niveau de la fonction $f(x, y) = y^2 - x^3 + 3x$: calculez les points critiques et précisez leur nature ; en plaçant quelques points à tangentes remarquables, tracez l'allure des courbes de niveau passant par les points critiques ; complétez enfin la figure en esquissant l'allure des courbes de niveau non critiques.

EXERCICE 2.— Discutez en fonction du paramètre p les points critiques de la fonction de deux variables $f(x, y) = y^2 - x^3 - px$: quand sont-ils quadratiques non dégénérés, et que pouvez-vous dire alors de leur dépendance en p ?

EXERCICE 3'— Généralisez : une fonction f de $(x_1, \dots, x_n, t_1, \dots, t_\ell)$, de classe \mathcal{C}^2 , est considérée comme fonction de (x_1, \dots, x_n) dépendant des paramètres $t = (t_1, \dots, t_\ell)$. Que pouvez-vous dire de la façon dont varie avec t les points critiques quadratiques non dégénérés de cette fonction ? (Utiliser le § 3.3.)

EXERCICE 4.— Soit f une fonction de classe \mathcal{C}^2 de n variables, non nulle au point a . Montrer que a est un point critique quadratique non dégénéré pour f si et seulement s'il l'est pour f^2 .

Extrema liés et points critiques en restriction à une sous-variété

Un problème d'*extrema liés* est un problème de recherche d'*extrema de la restriction d'une fonction à une sous-variété* $S \subset \mathbb{R}^n$. Pour une telle fonction f (différentiable dans un ouvert de \mathbb{R}^n contenant S) on a l'exacte réplique du lemme 1 :

Lemme 1^{bis}.— Si $f|_S$ a un *extremum local* en a , alors a est un *point critique* de $f|_S$ ³⁵.

Preuve

Elle est calquée sur celle du lemme 1. Mais puisque seule la *restriction* de f à S est supposée avoir un *extremum local* en a , le raisonnement doit se restreindre

³⁵ (cf. § 2.5, en prenant garde au fait que notre actuel $f|_S$ était alors noté f)

aux points mobiles sur S ; mais ce sont justement ceux dont les vecteurs vitesses forment l'espace tangent $T_a S$. \square

Les considérations du § 2.5 s'avèrent particulièrement utiles lorsque S n'est pas donnée par un paramétrage (local) mais par un système d'équations ($s_1 = 0, \dots, s_p = 0$ du type (3.1)).

Lemme 2.— Pour que a soit un point critique de $f|S$ il faut et il suffit que df_a soit combinaison linéaire de ds_{1a}, \dots, ds_{pa} .

Preuve

Dire que a est un point critique de $f|S$, c'est dire que $df_a|T_a S = 0$. Cela revient donc à dire que $(3.1)'_a \Rightarrow df_a = 0$. Or on sait que pour les systèmes linéaires homogènes la relation « d'implication logique » (toute solution du premier système est solution du second) équivaut à « l'implication linéaire »³⁶, c'est-à-dire dans notre cas au fait que df_a soit combinaison linéaire des premiers membres de $(3.1)'_a$. \square

Remarque Comme les formes linéaires ds_{1a}, \dots, ds_{pa} sont linéairement indépendantes, l'éventuelle expression de df_a comme combinaison linéaire de ds_{1a}, \dots, ds_{pa} :

$$df_a = \lambda_1 ds_{1a} + \dots + \lambda_p ds_{pa}$$

est unique. Ses coefficients $\lambda_1, \dots, \lambda_p$ sont appelés les *multiplicateurs de Lagrange* du point critique.

EXERCICE 5.— Dans \mathbb{R}^3 muni des coordonnées (x, y, z) on note S le cylindre d'équation $(x-1)^2 + y^2 = 1$. On se propose d'étudier et de dessiner les courbes de niveau de la fonction $f|S$, restriction à S de $f(x, y, z) = x^2 + y^2 + z^2$ (carré de la distance à l'origine).

- Étudiez avec soin les points critiques de $f|S$ (position ? développement limité à l'ordre 2, dans un système de coordonnées locales convenables ?)
- Dessinez, après les avoir déterminées par le calcul, les images des courbes de niveau en projection orthogonale sur le plan des (x, z) .
- Par des raisonnements géométriques, et en plaçant quelques points remarquables avec leurs tangentes, dessinez les images des courbes de niveau en projection orthogonale sur le plan des (y, z) .

PROBLÈME : points critiques, en restriction à une hypersurface, de la fonction distance à un point extérieur Dans ce problème l'espace ambiant \mathbb{R}^{m+1} est interprété comme un espace euclidien, muni de la « distance euclidienne » usuelle.

Soit S une hypersurface lisse de \mathbb{R}^{m+1} ($m \geq 1$), de classe C^∞ . Pour tout $b \in \mathbb{R}^{m+1}$ on désigne par $f_b : S \rightarrow \mathbb{R}^+$ la fonction qui à $a \in S$ associe $f_b(a) = \|\vec{ab}\|^2$ (carré de la distance euclidienne de a à b)³⁷.

³⁶ Cf. mon livre avec Hervé DILLINGER, « Découvrir l'algèbre linéaire », Diderot éd. 1997.

³⁷ REMARQUE On pourrait aussi prendre pour f_b la distance euclidienne et non son carré. Rien de ce qui suit ne s'en trouverait changé (cf. exercice 4), si ce n'est que pour avoir une fonction partout différentiable il faudrait interdire à b d'être dans S .

- 1) Montrer que a_0 est un point critique de f_b si et seulement si la normale à S en a passe par b .
- 2) En utilisant des coordonnées cartésiennes (orthonormées) (x_1, \dots, x_m, y) centrées en a_0 , telles que l'axe des y soit normal à S en a_0 (c'est-à-dire tel que $T_{a_0}S$ soit donné localement par l'équation $dy = 0$), on peut représenter S localement par une équation $y = g(x_1, \dots, x_m)$, où g est de classe C^∞ (pourquoi ?).
À l'aide d'un développement limité de g , caractériser les b pour lesquels a_0 est un point critique quadratique non dégénéré de f_b .
- 3) Montrer que moyennant des hypothèses locales assez générales sur S , a_0 , et sur le point b_0 choisi sur la normale à S en a_0 , la propriété suivante est vraie : par tout point b assez voisin de b_0 dans \mathbb{R}^{m+1} passe une unique droite normale à S en un point a voisin de a_0 ; les coordonnées de ce point sont fonctions différentiables de b (de classe C^∞).
- 4) Discuter, selon la position du point b dans \mathbb{R}^2 , les normales (passant par b) à une courbe lisse de \mathbb{R}^2 que vous vous donnerez explicitement : un cas particulièrement intéressant, et dont la solution peut être complètement explicitée, est le cas de la parabole $y = x^2$.
- 5) Discuter, en fonction de b , les normales issues de b à la surface lisse de \mathbb{R}^3 d'équation $y = x_1^2 - x_2^2$.

Cette question, malgré son apparence simple, conduit à un système couplé de deux équations polynomiales à deux inconnues x_1, x_2 , qui a de quoi décourager toute velléité de résolution explicite. On pourra se contenter de résultats partiels : que se passe-t-il quand b est voisin de l'axe des y (appliquer la question 3) : pour $a_0 = (0, 0, 0)$ tracer, pour b choisi sur l'axe des y , l'allure des courbes de niveau de la fonction f_b en projection sur le plan des x_1, x_2 (discuter en fonction de b) ; leur aspect est-il différent quand b est hors de l'axe des y , tout en restant voisin de cet axe ? Que se passe-t-il pour b voisin de $(0, 0, 1)$? Etc.

ÉTUDE 1 (bis) Lemme de Morse

On va s'intéresser ici à une fonction f de n variables x_1, \dots, x_n , différentiable de classe C^r ($r \geq 2$) au voisinage d'un point a que pour simplifier les notations on supposera être l'origine de \mathbb{R}^n . Énoncé par le mathématicien américain Marston MORSE en 1939, le lemme suivant annonce le grand essor de la topologie différentielle au milieu du siècle.

Lemme de Morse Supposons que f ait l'origine comme point critique quadratique non dégénéré. Alors il existe un changement de variables local de classe C^{r-2} :

$$(x_1, \dots, x_n) \leftrightarrow (u_1, \dots, u_n)$$

envoyant 0 sur 0 et transformant la fonction f en :

$$\tilde{f} = \sum_{i=1}^n \pm u_i^2$$

où les nombres respectifs de signes + et - sont donnés par la signature de la hessienne de f en a .

Commentaire D'après l'étude 1 (cor.1) on sait déjà qu'il existe un changement de variables qui transforme f en :

$$\sum_{i=1}^n \pm u_i^2 \text{ mod. } o(\|u\|^2)$$

La force du lemme de Morse est de permettre de nous débarrasser des termes correctifs en $o(\|u\|^2)$.

Nous allons esquisser, sous forme de problème, une preuve du lemme de Morse qui passe par le lemme (de 30 ans postérieur) que voici.

Lemme de Gromoll-Meyer Supposons que l'on ait :

$$f(0) = 0, \quad \frac{\partial f}{\partial x_1}(0) = 0, \quad \frac{\partial^2 f}{\partial x_1^2}(0) \neq 0$$

Alors il existe un changement de variables local de classe C^{r-2}

$$(x_1, \dots, x_n) \leftrightarrow (\tilde{x}_1, \dots, \tilde{x}_n)$$

envoyant 0 sur 0 et transformant la fonction f en

$$\tilde{f} = \pm \tilde{x}_1^2 + h(\tilde{x}_2, \dots, \tilde{x}_n)$$

où h est une fonction de classe C^r .

PROBLÈME

- 1) En utilisant l'exercice 3 de l'étude 1, montrer qu'un changement de variables local convenable, de la forme $(x_1, x_2, \dots, x_n) \leftrightarrow (\tilde{x}_1, x_2, \dots, x_n)$, avec $\tilde{x}_1 = x_1 - a(x_2, \dots, x_n)$, transforme f en une famille de fonctions de \tilde{x}_1 (paramétrée par $t = (x_2, \dots, x_n)$) ayant pour tout t assez petit un point critique quadratique en $\tilde{x}_1 = 0$.
- 2) En retranchant à la famille de fonctions de \tilde{x}_1 ainsi obtenue sa valeur en $\tilde{x}_1 = 0$ (valeur qui dépend de t) on obtient une famille de fonctions ayant $\tilde{x}_1 = 0$ comme racine double (pour t assez petit).

En déduire le lemme de Gromoll-Meyer, en admettant le lemme suivant (qui se déduit facilement de la formule de Taylor avec reste intégral) :

toute (famille de) fonction(s) g de classe C^r ($r \geq m$) d'une variable réelle \tilde{x}_1 , ayant l'origine comme zéro d'ordre m , peut s'écrire au voisinage de 0

$$g(\tilde{x}_1, t) = \tilde{x}_1^m k(\tilde{x}_1, t)$$

où k est une fonction de classe C^{r-m} non nulle au voisinage de 0.

- 3) Démontrer le lemme de Morse par applications successives du lemme de Gromoll-Meyer.
- 4) (exemple d'une situation dégénérée) On considère une fonction C^∞ de deux variables, en un point critique où la hessienne est de rang 1. Montrer qu'il existe un changement de variables local de classe C^∞ qui transforme cette fonction en $\tilde{f} = \pm \tilde{x}_1^2 \pm \tilde{x}_2^m$ (m entier > 2) ou bien en $\tilde{f} = \pm \tilde{x}_1^2 + k(\tilde{x}_2)$, où k est une fonction « plate en 0 », c'est-à-dire s'annulant en 0 ainsi que toutes ses dérivées successives.

Commentaire La démonstration ci-dessus (qui m'a été suggérée en 1980 par Nguyễn Tiên Dai) est l'analogue dans le cadre différentiel de la *méthode de Lagrange* pour décomposer une forme quadratique en combinaison linéaire de carrés (cf. étude 1). Il est instructif de la relire dans cet esprit : en remplaçant la fonction de départ par une forme quadratique (de sorte que ses dérivées partielles sont des formes linéaires) on obtient la construction de Lagrange.

QUESTION : qu'est devenu dans ce contexte notre recours au théorème des fonctions implicites ?

Commentaire linguistique

Trois mots-clefs du discours mathématique

J'ai donné à la fin du § 0 (« petit test linguistique ») deux échantillons de tournures typiques du langage de ce chapitre, tournures un peu désuètes qui ont pu paraître « abusives » (voire « incorrectes ») à certains.

Dans la première, contenant les deux mots-clefs « variable » et « fonction », le mot « fonction » est utilisé dans son sens primitif, qu'un logicien d'aujourd'hui qualifierait de *métamathématique*. Les « maths modernes » répugnent à ce type de discours « méta » où les mathématiques sont regardées *de l'extérieur*. C'est pourquoi elles considèrent avec suspicion tout discours sur les *variables* : n'étant que des *noms* par lesquels on « désigne » des « objets mathématiques », les variables ne méritent pas d'être des objets d'attention ; pas question de leur reconnaître un rôle stratégique dans le discours, à la jonction de la *forme* et du *sens*, comme nous l'avons fait à plusieurs reprises dans ce chapitre en attribuant à une même variable diverses *interprétations* !

La seconde phrase fait la distinction entre une *relation* et le *graphe* de cette relation. Les « maths modernes » ont décrété qu'une telle distinction (qui allait de soi pour les mathématiciens classiques) devait être abolie, *une relation ne pouvant pas être*

définie autrement que comme graphe : dire qu'une relation est « ce qu'une formule signifie » (ce qui résume le point de vue « à l'ancienne ») renvoie au problème du sens, que les « maths modernes » récusent sous le prétexte imparable qu'on ne dispose pas d'une définition formelle du sens !

« Variables », « relations », « fonctions » : trois mots-clés du langage mathématique que nous ont légué nos pères. Sous l'éclairage de la logique moderne, nous allons essayer d'en comprendre toute la richesse, en résistant aux tentations « réductionnistes ».

Relations

Un texte mathématique s'articule autour de *formules*, assemblages de symboles obéissant à une certaine syntaxe. Parmi ces symboles figurent certaines lettres (comme x , y , etc.) appelées *variables*, qui dans les mathématiques élémentaires sont censées représenter des nombres.

Considérons par exemple la formule (1) du § 0.1. Elle exprime une *relation* entre les deux variables x et y . Le mot « relation » est essentiellement synonyme de ce que les logiciens appellent un « énoncé ouvert », énoncé dont l'expression contient des « variables libres » x_1, \dots, x_n , de sorte que la « valeur de vérité » – véricité ou fausseté – de l'énoncé dépend des valeurs qu'il nous plaira d'attribuer à ces variables : par exemple tout système d'équations ou d'inéquations polynomiales portant sur n variables x_1, \dots, x_n définit une relation entre ces variables³⁸.

On se gardera de confondre la *relation* et la *formule* qui en est l'expression : par exemple « $y^3 - x = 0$ » et « $y^3 = x$ » sont deux formules *différentes*, qui expriment la même relation entre les variables x et y . De façon générale deux formules contenant les mêmes variables libres x_1, \dots, x_n « expriment la même relation » si et seulement si elles sont équivalentes, c'est-à-dire si elles sont vérifiées pour exactement les mêmes valeurs du n -uplet (x_1, \dots, x_n) .

La notion de *relation* (*énoncé ouvert*) n'est donc pas une notion *syntactique* (c'est-à-dire se référant au langage dans lequel la relation est formulée) mais une notion *sémantique* (c'est-à-dire se référant au sens).

Rôle des variables dans une relation On reconnaît qu'un énoncé est ouvert au fait que sa formulation contient ce qu'on appelle des *variables libres*³⁹.

D'un point de vue *syntactique*, les variables libres renvoient au contexte d'où la formule est extraite ; par exemple dans un fragment de texte mathématique comme

« ... en utilisant la relation $x < y$, on déduit de ce qui précède que ... »

les lettres x et y renvoient à d'autres endroits du texte, où elles sont censées faire référence aux mêmes nombres que dans le fragment concerné.

³⁸ La connotation courante du mot « relation » suggère qu'au moins deux variables soient en jeu. Néanmoins il est mathématiquement peu commode d'exclure le cas $n = 1$ de considérations qui sont valides aussi dans ce cas. C'est pourquoi certains auteurs parlent de relations « unaires », comme on dit « relations binaires », « relation ternaires », etc.

³⁹ On distingue en logique mathématique deux types d'occurrences d'une variable dans une formule : les occurrences *libres* et les occurrences *liées* ; une occurrence d'une variable x est *liée* si elle est « dans le champ d'un quantificateur » : cf. exemple ci-après, où toutes les occurrences des trois variables sont dans le champ d'un quantificateur \forall .

Mais les variables libres peuvent aussi avoir un rôle *sémantique* : lorsque dans un calcul de thermodynamique apparaît la formule $PV = RT$, les variables P (pression), V (volume), T (température) ont une *signification physique* : elles renvoient au contexte de la réalité physique que le calcul prétend modéliser.

Ceci dit, il est souvent essentiel pour le mathématicien de *s'affranchir du contexte*, en faisant abstraction des variables (abstraire est l'acte mathématique par excellence !). Par exemple un énoncé comme

« la relation $<$ est transitive »

est vrai *indépendamment de tout contexte* (ce qu'atteste ici l'absence de variables dans sa formulation). Dans sa traduction en langage formalisé :

$$\forall x \forall y \forall z ((x < y \text{ et } y < z) \Rightarrow x < z)$$

apparaissent pourtant des variables, mais d'une façon qui ne renvoie à aucun contexte (on pourrait remplacer chacune d'entre elles par une lettre quelconque sans rien changer au reste du texte dans lequel la formule est insérée) : dans le langage de la logique, on dit qu'il s'agit de *variables liées*, ou plus précisément que toutes les occurrences des variables x, y, z dans la formule ci-dessus sont liées⁴⁰. En « argot » mathématique on dit qu'il s'agit de variables « muettes ».

Ces considérations expliquent pourquoi le mot « relation » peut être utilisé de deux façons dans un texte mathématique :

1. « la relation $x < y$ » ;
2. « la relation $<$ ».

Dans le premier cas la partie encadrée est l'expression d'un *énoncé ouvert*, à deux variables. Dans le deuxième cas la partie encadrée est l'expression de ce que les logiciens appellent un *prédicat à deux places*, déduit de l'énoncé précédent *en faisant abstraction des variables*.

Graphe d'une relation (point de vue de la théorie des ensembles) Considérons une relation entre n variables x_1, \dots, x_n , dont les domaines de variation seront notés X_1, \dots, X_n (par exemple $X_i = \mathbb{R}$ si x_i est une variable réelle sans restriction, $X_i = \mathbb{R}^+$ si x_i est une variable réelle ≥ 0 , etc.).

On appelle *graphe* de la relation le sous-ensemble $G \subset X_1 \times \dots \times X_n$, ensemble des valeurs du n -uplet (x_1, \dots, x_n) pour lesquelles la relation est vérifiée⁴¹.

Avec cette terminologie notre notion « sémantique » de relation peut se traduire ainsi :

deux relations entre les mêmes variables sont identiques (on peut dire aussi égales⁴¹) si et seulement si elles ont même graphe.

⁴⁰ Cf. note précédente. On se gardera de confondre cette notion (syntactique) d'« occurrence liée d'une variable » avec l'usage courant (sémantique) de l'adjectif « liée », comme dans l'introduction de cette première partie quand j'ai parlé des variables P, V, T « liées par la relation $PV = RT$ » !

⁴¹ Le mot « graphe » trouve son origine dans le fait qu'il est commode de représenter graphiquement des sous-ensembles de \mathbb{R}^2 par des *dessins* dans le plan, en interprétant x_1 et x_2 comme les coordonnées cartésiennes de points du plan.

⁴¹ ou *équivalentes* ! mais il est plus correct de réserver cet adjectif aux *formulations* des relations (deux formules sont équivalentes si et seulement si elles définissent la même relation).

C'est pourquoi en théorie des ensembles on *définit* la notion de relation en l'identifiant à celle de *graphe* : plus précisément on appelle *relation* (ou *correspondance*)⁴² entre les ensembles X_1, \dots, X_n un $(n + 1)$ -uplet (X_1, \dots, X_n, G) , où $G \subset X_1 \times \dots \times X_n$.

Remarquez comment cette définition « fait abstraction des variables », et préfère insister sur leurs domaines de variation. Remarquez aussi que cet « effacement des variables » a nécessité de les avoir préalablement rangées dans un certain ordre, ce qui relève d'une décision arbitraire (si nos variables s'appellent θ et x , il n'y a pas d'ordre alphabétique pour nous dicter comment les ranger !).

Fonctions

La notion de *fonction* est sans doute l'une des plus importantes et des plus subtiles de toutes les mathématiques. Son sens a beaucoup évolué au cours des siècles, et reste quelque peu fluctuant. Vers les années soixante, lors de la grande remise en ordre des mathématiques connue du grand public sous le nom de « maths modernes », on a pu croire que l'on avait enfin dégagé LA bonne notion de fonction. Mais certains développements récents des mathématiques (motivés notamment par les problèmes de calcul effectif sur ordinateur) conduisent actuellement à remettre en question cette conviction, et à réhabiliter des points de vue qu'on avait pu croire surannés⁴³.

Les premières fonctions que rencontre un débutant sont données par des expressions explicites, comme par exemple :

« la fonction x^2 »

Mais une fonction peut aussi être donnée *implicitement*, comme solution d'une équation : par exemple, l'équation :

$$x - y^3 = 0 \tag{1}$$

admet pour toute valeur réelle de x une seule solution y réelle, qui définit la fonction « racine cubique de x ». On dit que la relation (1) est *fonctionnelle en y* .

DÉFINITION 1.— Une relation $R(x, y)$ est dite **fonctionnelle en y** si pour toute valeur de x il existe une et une seule⁴⁴ valeur de y pour laquelle $R(x, y)$ est vraie⁴⁵.

DÉFINITION 2.— On appelle **fonction** une relation, fonctionnelle en y , entre deux variables x et y respectivement marquées du qualificatif de « variable de départ » et « variable d'arrivée ». Si cette relation est aussi fonctionnelle en x on dit que c'est un **changement de variable** ; en considérant y comme variable de départ et x comme

² « Correspondance » est le mot qu'utilise Bourbaki (Théorie des Ensembles Chap. I). Je n'utiliserai ici le mot « relation » que dans le sens *métamathématique* introduit plus haut. C'est aussi ce que fait Bourbaki, sauf que pour lui « relation » = « formule » (contenant des variables libres).

³ Cf. par exemple le joli article « Fonctions, programmes et démonstrations », de Jean-Louis KRIVINE, dans *La Gazette des Mathématiciens* numéro 60 (avril 1994).

⁴ Certains auteurs disent « au plus une ». Nous reviendrons sur ce point un peu plus loin.

⁵ On définirait de façon analogue le fait pour une relation $R(x_1, \dots, x_n, y_1, \dots, y_p)$ d'être *fonctionnelle en (y_1, \dots, y_p)* . Si l'on accepte que les variables puissent être autre chose que des nombres, ce cas se ramène au précédent en considérant que $x \in \mathbb{R}^n$ et $y \in \mathbb{R}^p$.

variable d'arrivée on obtient alors une fonction appelée fonction réciproque de la fonction précédente.

Exemple La formule (1) définit un changement de variable.

Attention aux domaines des variables ! Dans les exemples précédents nous avons supposé que les lettres x , y désignaient des nombres réels *quelconques*. Mais on est souvent amené à assigner aux variables des domaines de variation plus restreints, qu'il faut alors *déclarer* au moment où l'on introduit les lettres x , y (en disant par exemple : « x et y sont des variables réelles > 0 »). Pour comprendre l'importance de telles déclarations, considérons par exemple la relation :

$$x - y^2 = 0$$

Si le domaine des variables x et y est \mathbb{R} tout entier, cette relation n'est pas fonctionnelle en y , car pour $x > 0$ il y a deux valeurs de y (opposées l'une de l'autre) qui vérifient la relation (et pour $x < 0$ il n'y en a aucune). Mais si l'on a déclaré x et y comme des variables réelles *positives* cette relation devient fonctionnelle en y (elle définit la fonction « racine carrée »).

Comment se donner (et nommer) une fonction Lorsqu'une fonction est donnée par une formule explicite :

$$y = f(x) \quad (\text{exemple : } y = x^2) \quad (2)$$

l'usage classique est de considérer l'expression $f(x)$ figurant au membre de droite de (2) comme un *nom* de cette fonction (exemple : « la fonction x^2 »). Dans le langage de la logique, une telle expression est ce qu'on appelle un « terme » (c'est-à-dire une expression formelle désignant un objet du même type que les variables : ici un nombre réel). On peut alors réinterpréter la lettre y (membre de gauche de (2)) comme un *autre nom* de l'objet (« fonction de x ») que désigne le terme $f(x)$.

Considérons maintenant le cas d'une fonction donnée *implicitement*, comme par exemple la fonction « racine cubique de x », solution en l'inconnue y de l'équation (1) ci-avant. Pour ré-écrire cette fonction sous forme « explicite », il suffit de lui affecter un *symbole de fonction* qui lui soit propre : en l'occurrence le symbole traditionnel est $\sqrt[3]{}$, de sorte que la relation (1) peut se ré-écrire :

$$y = \sqrt[3]{x} \quad (3)$$

C'est ainsi que certaines fonctions remarquables se sont vues affecter par la tradition des symboles universellement reconnus : la fonction $\sqrt{}$, la fonction \ln , la fonction \exp , la fonction \sin , etc., etc. (remarquez comment là encore la notation fait abstraction des variables !).

Bien entendu il s'agit là d'un simple jeu d'écriture, qui facilite la façon de désigner la fonction mais n'a pas la vertu magique d'en faciliter le calcul !

La question de savoir quelles fonctions sont « exprimables explicitement » est donc fortement dépendantes des conventions d'écriture. Toutefois on peut la voir comme un cas particulier de la question, plus fondamentale, de la « calculabilité » des fonctions : lorsque les mathématiciens savent définir une fonction (fût-ce implicitement,

par une relation fonctionnelle telle que (1)), c'est presque toujours parce qu'ils savent en fournir un *algorithme de calcul*, qu'ils sont en principe capables d'écrire (par exemple sous forme d'un *programme exécutable* par un ordinateur), et il est naturel de considérer le texte d'un tel programme comme une « expression explicite » de la fonction, au même titre que l'expression « x^2 » de la formule (2) (qui après tout n'est rien d'autre qu'une écriture du programme simple « multipliez la variable x par elle-même »).

Prise en ce sens plus général, la notion de « fonction exprimable explicitement » peut être dotée d'un statut mathématique précis (cf. l'article de J.L. KRIVINE précédemment cité). La façon de parler classique

« la fonction $f(x)$ »

se voit ainsi réhabilitée, la notation « $f(x)$ » pouvant être considérée comme le *nom* d'un *programme de calcul de la fonction*.

Domaine de définition d'une fonction Il est fréquent que la valeur d'une « expression » $f(x)$ ne soit définie que pour certaines valeurs de x . L'ensemble de ces valeurs de x est ce qu'on appelle le *domaine de définition de la fonction f* . Si l'on prend le mot « expression » au sens général de « programme de calcul », le domaine de définition de la fonction est donc l'ensemble des valeurs de x pour lesquelles le programme aboutira. Déterminer ce domaine est souvent une tâche insurmontable. C'est pourquoi beaucoup d'auteurs préfèrent, dans la définition 1 d'une « relation fonctionnelle », remplacer le membre de phrase « il existe *une et une seule* valeur de y » par « il existe *au plus une* valeur de y » (un programme correctement écrit aboutit nécessairement à un résultat bien déterminé, lorsqu'il aboutit !). Autrement dit, ce qu'ils appellent « relation fonctionnelle » est une relation *qui deviendrait fonctionnelle au sens de la définition 1, si l'on restreignait convenablement le domaine de variation de x* .

Substitutions (composition des fonctions) La notion de « composition des fonctions » se comprend très simplement en termes de programmes de calcul : l'écriture $y = f(g(t)) = (f \circ g)(t)$ signifie que pour calculer y à partir de t on fournit comme valeur d'entrée x du programme $f(x)$ la valeur de sortie du programme $g(t)$:

$$y = f(x), \quad x = g(t)$$

Le domaine de définition de $f \circ g$ est bien sûr l'ensemble des valeurs de t pour lesquelles $g(t)$ appartient au domaine de définition de f .

Les fonctions en théorie des ensembles Nous avons déjà signalé comment la théorie des ensembles traduit l'idée de « relation » en terme de « graphe ». Plus précisément :

1. On appelle **graphe** (binaire) un ensemble G dont tous les éléments sont des couples (x, y) . Un graphe G est dit **fonctionnel** si :

$$\forall x \forall y \forall y', (x, y) \in G \text{ et } (x, y') \in G \Rightarrow y = y'$$

2. Lorsque $G \subset X \times Y$, le triplet (X, Y, G) est appelé **correspondance** entre les ensembles X et Y : l'ensemble X resp. Y est appelé **ensemble de départ** resp. **ensemble d'arrivée**⁴⁶ de la correspondance ; l'ensemble G est appelé **graphe** de la correspondance.

On appelle **application de X dans Y** une correspondance (X, Y, G) telle que $\forall x \in X$ il existe un et un seul $y \in Y$ pour lequel $(x, y) \in G$.

Le mot *fonction* en théorie des ensembles peut être utilisé dans deux sens différents (cf. Bourbaki, Théorie des Ensembles Chap. I) :

1. **fonction = graphe fonctionnel** ;
2. **fonction (sur X à valeurs dans Y) = application de X dans Y .**

L'usage 1 permet de parler de fonctions sans préciser ni domaine de définition ni ensemble d'arrivée. L'usage 2 s'impose au contraire lorsqu'on juge important de donner de telles précisions.

Remarquez que c'est l'ordre dans lequel on se donne les ensembles X et Y qui détermine lequel est ensemble de départ et lequel est ensemble d'arrivée (ce qui dispense du « marquage » dont il était question dans la définition 2).

Remarquez aussi à nouveau comment le point de vue ensembliste fait abstraction des variables x, y : dans les énoncés écrits quelques lignes plus haut ces variables sont « muettes », et il en est de même lorsqu'on écrit par exemple :

$$\text{« la fonction } \mathbb{R} \rightarrow \mathbb{R} \text{ » , « la fonction } \mathbb{R}^+ \rightarrow \mathbb{R} \text{ » , etc.}$$

$$x \mapsto x^2 \qquad x \mapsto \sqrt{x}$$

C'est ce qui fait la grande différence avec la façon de s'exprimer classique, où parler de :

$$\text{« la fonction } x^2 \text{ » , « la fonction } \sqrt{x} \text{ » etc.}$$

n'a de sens que si l'on a déclaré x comme *variable*, déclaration de nature *métamathématique* qui se réfère à un *contexte*.

⁴⁶ Les géomètres disent aussi *espace source* resp. *espace but*

Applications différentiables et sous-variétés

Nous abordons ici le point de vue moderne sur le calcul différentiel

1. Applications différentiables entre ouverts d'espaces numériques

Ce paragraphe résume ce qu'un licencié en mathématiques est censé savoir aujourd'hui sur les applications différentiables.

1.1. Différentielle d'une application

Une application $f : U \rightarrow V$ d'un ouvert U de \mathbb{R}^n dans un ouvert V de \mathbb{R}^p est dite **différentiable au point** $a \in U$ s'il existe une application linéaire $Df_a : \mathbb{R}^n \rightarrow \mathbb{R}^p$ ¹ telle que :

$$f(a + u) = f(a) + (Df_a)(u) \text{ mod. } o(\|u\|)$$

Cette application linéaire est alors unique. On l'appelle la **différentielle de f en a** .

La différentielle comme fonction du point

L'application f est dite **différentiable** si elle l'est en tout point a de son ouvert U de définition. On appelle alors **différentielle de f** l'application :

$$\begin{aligned} Df : U &\rightarrow L(n, p) \\ a &\mapsto Df_a \end{aligned}$$

à valeurs dans l'espace $L(n, p)$ des applications linéaires de \mathbb{R}^n dans \mathbb{R}^p . On dit que f est **continûment différentiable** (ou **de classe C^1**) si cette application Df est continue.

Différentielles d'ordre supérieur

En identifiant $L(n, p)$ à l'espace des matrices $p \times n$ (évidemment isomorphe à \mathbb{R}^{np}) on peut considérer Df comme une application à valeurs dans \mathbb{R}^{np} , et se poser à nouveau la question de sa différentiabilité. On définit ainsi la **différentielle seconde** $D^2f_a = D(Df)_a$, et l'on peut continuer, définissant la notion **d'application de classe C^2** , etc., **de classe C^r** ,...

¹ La lettre d minuscule est parfois employée au lieu du D majuscule. Dans ce livre le d minuscule sera réservé au cas où f est une *fonction numérique*, c'est-à-dire le cas où l'espace but est \mathbb{R} .

Lien avec les notations du chap. 1 En notant x_1, \dots, x_n les coordonnées canoniques de l'espace source \mathbb{R}^n , et y_1, \dots, y_p celles de l'espace but \mathbb{R}^p , l'application f donne y_1, \dots, y_p comme fonctions de (x_1, \dots, x_n) :

$$\begin{cases} y_1 = f_1(x_1, \dots, x_n) \\ \dots \\ y_p = f_p(x_1, \dots, x_n) \end{cases} \quad (1.1)$$

La différentielle Df_a est l'application linéaire qui au vecteur de \mathbb{R}^n de coordonnées (dx_1, \dots, dx_n) associe le vecteur de \mathbb{R}^p de coordonnées (dy_1, \dots, dy_p) données par :

$$\begin{cases} dy_1 = df_{1a} = \frac{\partial f_1}{\partial x_1}(a) dx_1 + \dots + \frac{\partial f_1}{\partial x_n}(a) dx_n \\ \dots \\ dy_p = df_{pa} = \frac{\partial f_p}{\partial x_1}(a) dx_1 + \dots + \frac{\partial f_p}{\partial x_n}(a) dx_n \end{cases} \quad (1.1)'$$

Autrement dit Df_a est l'application linéaire de \mathbb{R}^n dans \mathbb{R}^p définie par la matrice $(\frac{\partial f_i}{\partial x_j})_{i,j}$, appelée **matrice jacobienne de l'application f au point a** .

L'application f est de classe C^1 si et seulement si sa matrice jacobienne est fonction continue du point.

Si f est deux fois différentiable en a sa différentielle seconde D^2f_a est une *application bilinéaire symétrique* :

$$D^2f_a : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^p$$

$$(u, v) \mapsto \left(\sum_{k=1}^n \frac{\partial^2 f_k}{\partial x_i \partial x_j}(a) u_i v_j \right)_{k=1, \dots, p}$$

Dans le cas d'une fonction numérique ($p = 1$), ce qu'au chap. 1 § 1.3.1 nous avons appelé la *différentielle seconde de f (en a)* n'est pas autre chose que la *forme quadratique associée à cette forme bilinéaire* (restriction de celle-ci à la diagonale $u = v$). De même $D^r f_a$, différentielle d'ordre r en a , est une *application r -linéaire symétrique* de $\mathbb{R}^n \times \dots \times \mathbb{R}^n$ (r fois) dans \mathbb{R}^p , dont les coefficients sont les dérivées partielles d'ordre r de f au point a .

On retiendra de tout ceci que l'application $f = (f_1, \dots, f_p)$ est de classe C^r (ou de classe C^∞ , ou de classe C^ω c'est-à-dire *analytique*) si et seulement si chacune de ses fonctions composantes f_1, \dots, f_p l'est (au sens du chap. 1 § 1.3.1).

Remarque « Être de classe C^r » est une propriété locale. On entend par là que :

1. Être de classe C^r n'a de sens (par définition) que dans un ouvert².
2. Si une fonction (application) f définie dans un ouvert U est de classe C^r dans un voisinage U_a de tout point $a \in U$, elle est de classe C^r dans U .

² Il n'en est pas de même de la propriété d'être différentiable : par exemple une fonction peut être différentiable en un point a et seulement en ce point. Il en est d'ailleurs de même de la propriété de continuité ; mais une fonction continue en un point a et seulement en ce point ne sera pas dite « de classe C^0 en a » : qui dit « classe C^0 » dit continuité dans un ouvert.

1.2. Composition des différentielles

Considérons la situation suivante, où la notation $\mathbb{R}^n, a \cdots \xrightarrow{f}$ désigne une application définie dans un voisinage du point a dans \mathbb{R}^n , voisinage qu'on n'a pas jugé utile de préciser :

$$\mathbb{R}^n, a \cdots \xrightarrow{f} \mathbb{R}^p, b \cdots \xrightarrow{g} \mathbb{R}^q {}^3$$

Si f est différentiable en a , et g différentiable en b , $g \circ f$ est différentiable en a , et $D(g \circ f) = Dg_b \circ Df_a$.

De plus, si f est de classe C^r dans U , et g de classe C^r dans V , avec $f(U) \subset V$, l'application composée $g \circ f$ est de classe C^r dans U .

1.3. Difféomorphismes

Un **difféomorphisme** (de classe C^r) d'un ouvert U de \mathbb{R}^n dans un ouvert V de \mathbb{R}^n est une application bijective $\phi : U \rightarrow V$ qui est différentiable (de classe C^r) de même que son application réciproque ϕ^{-1} .

L'étude d'objets géométriques à *difféomorphismes près* constitue ce qu'on appelle la *topologie différentielle*. Par exemple les deux dessins de la figure 1, déduits l'un de l'autre par un difféomorphisme du plan, sont équivalents du point de vue de la topologie différentielle.

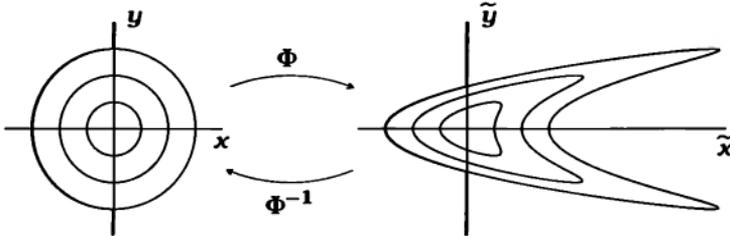


Figure 1. Familles de cercles concentriques $x^2 + y^2 = 0$, et leurs images par le difféomorphisme $\phi : (x, y) \mapsto (x + y^2, y)$.

Lien avec les notions de « changement de variables » et de « changement de coordonnées »

Pour faire la figure 1 nous avons introduit le *changement de variables* $(x, y) \leftrightarrow (\tilde{x}, \tilde{y})$ défini par $(\tilde{x} = x + y^2, \tilde{y} = y)$ (ou inversement $(x = \tilde{x} - \tilde{y}^2, y = \tilde{y})$), et interprété chacun des deux couples de variables comme un *système de coordonnées cartésiennes* dans un plan muni d'un repère orthonormé.

Le point de vue des « maths modernes » ne fait aucune distinction entre « changement de variables différentiable » et « difféomorphisme » de \mathbb{R}^n , les identifiant d'ailleurs aussi avec « changement de coordonnées » ! Ce n'est pourtant pas sans raison que la tradition nous a légué trois termes différents. Considérons par exemple la relation :

$$R : \begin{cases} x = \rho \cos \theta \\ y = \rho \sin \theta \end{cases}$$

³ Ce type de notation sous-entend aussi que l'application envoie le « point marqué à la source » sur le « point marqué au but » : par exemple ici $f(a) = b$.

entre les variables x, y, ρ, θ dont on aura restreint les domaines de variation de façon telle que R définisse un *changement de variables différentiable* $(\rho, \theta) \leftrightarrow (x, y)$: disons, pour fixer les idées, $\rho > 0, -\frac{\pi}{2} < \theta < \frac{\pi}{2}, x > 0$. Qualifier ce changement de variables de *changement de coordonnées*, c'est dire qu'on interprète les deux couples de variables comme deux façons de repérer une même point du plan (c'est, dans le cas présent, l'interprétation la plus traditionnelle : $(x, y) =$ « coordonnées cartésiennes », $(\rho, \theta) =$ « coordonnées polaires »). Mais dire que ce changement de variables définit un *difféomorphisme* entre les ouverts $U = \mathbb{R}^+ \times]-\frac{\pi}{2}, \frac{\pi}{2}[$ et $V = \mathbb{R}^+ \times \mathbb{R}$:

$$\begin{array}{ccc} U & \xrightarrow{\phi} & V \\ (\rho, \theta) & \mapsto & (\rho \cos \theta, \rho \sin \theta) \end{array} \quad 4$$

c'est proposer une autre interprétation des deux couples de variables, comme « coordonnées canoniques » dans \mathbb{R}^2 (cf. chap. 1, § 2.0)⁵; il s'agit alors bien sûr des coordonnées de points différents, images l'un de l'autre par le difféomorphisme. Plus concrètement, on pourra aussi interpréter les deux couples de variables comme des coordonnées cartésiennes « dans le plan de la feuille », ce qui permet de donner du difféomorphisme ϕ une représentation imagée analogue à celle de la figure 1⁶.

Exercice En utilisant cette dernière interprétation, donnez-vous dans le demi-plan V un quadrillage parallèle aux axes de coordonnées, et dessinez dans la « demi-bande » U l'image réciproque de ce quadrillage par le difféomorphisme ϕ ci-dessus.

En résumé, si les « maths modernes » ne voient aucune différence entre « changements de variables », « changements de coordonnées », et « difféomorphismes » de \mathbb{R}^n , c'est à cause de leur parti-pris formaliste qui refuse de s'intéresser aux questions d'interprétation.

1.4. Difféomorphismes locaux et applications différentiables de rang maximal

Il résulte de 1.2 que la différentielle d'un difféomorphisme est une application linéaire inversible : $D\phi_a$ a pour inverse $(D\phi_b)^{-1}$ (avec $b = \phi(a)$). C'est d'ailleurs pour cette raison que source et but ont été supposés de même dimension. Le théorème que voici dit que la réciproque est vraie localement.

Théorème d'inversion locale Soit $\phi : U \rightarrow V$ une application différentiable de classe C^r ($r \geq 1$) dont la différentielle $D\phi_a$ en un certain point $a \in U$ est **inversible**. Alors ϕ est un **difféomorphisme local**, c'est-à-dire qu'il existe dans U un voisinage U_a de a , et dans V un voisinage V_b de $b = \phi(a)$, tels que $\phi|_{U_a} : U_a \rightarrow V_b$ soit un difféomorphisme de classe C^r . □

⁴ Remarquez qu'ici les variables ρ, θ sont « muettes » : on aurait pu les remplacer par deux lettres quelconques.

⁵ Ainsi comprise la notion de difféomorphisme est la version ensembliste de la notion de changement de variables ; on peut y voir une illustration du processus « d'abstraction des variables », déjà illustré dans le commentaire linguistique en fin de chapitre 1.

⁶ Le processus « d'abstraction des variables » nous fait ainsi quitter le domaine du calcul pour entrer dans celui de la géométrie.

Nous avons déjà énoncé au chap. 1 § 3.2 la version « à l'ancienne » de ce théorème.

Attention ! Une application différentiable peut être, en tout point de l'espace source, un difféomorphisme local, sans pour cela être un difféomorphisme, même si elle est surjective : un bon exemple à méditer est l'application :

$$f : \mathbb{R}^+ \times \mathbb{R} \rightarrow \mathbb{R}^2 \setminus \{(0, 0)\}$$

$$(\rho, \theta) \mapsto (\rho \cos \theta, \rho \sin \theta)$$

Comme corollaire du théorème d'inversion locale nous allons caractériser localement, à difféomorphisme près, les applications différentiables $f : (\mathbb{R}^n, a) \cdots \rightarrow (\mathbb{R}^p, b)$ dont la différentielle Df_a est de rang maximal.

1.4.1. Submersions (rang $Df_a = \dim \text{but}$)

PROPOSITION

Les deux propriétés suivantes sont équivalentes :

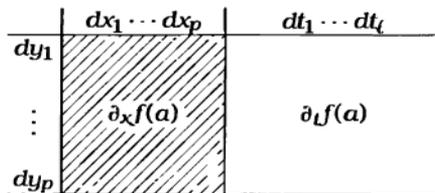
- i) Df_a est surjective, c'est-à-dire de rang égal à la dimension du but ;
- ii) la composition avec un difféomorphisme local convenable de l'espace source (au voisinage de a) transforme f en la projection canonique $\text{pr}_1 : \mathbb{R}^p \times \mathbb{R}^\ell \rightarrow \mathbb{R}^p$.

Une application différentiable f vérifiant ces propriétés est appelée **submersion** (locale en a). On dit *submersion* tout court si f est une submersion locale en tout point de son ouvert de définition.

Preuve

L'implication ii) \Rightarrow i) résulte immédiatement de 1.2. Montrons la réciproque.

Dire que Df_a est surjective, c'est dire que l'un au moins des mineurs $p \times p$ de la matrice jacobienne de f en a est inversible, et quitte à permuter les coordonnées à la source on peut toujours supposer que c'est le cas du mineur de gauche, noté $\partial_x f(a)$ ci-dessous :



On a noté $x = (x_1, \dots, x_p)$ les p « premières » coordonnées canoniques de l'espace source et $t = (t_1, \dots, t_\ell)$ les $\ell = n - p$ « dernières » (après permutation). En notant $y = (y_1, \dots, y_p)$ les coordonnées canoniques de l'espace but, données en fonction des précédentes par $y = f(x, t)$, on conclut en remarquant que l'application f se factorise en :

$$f : (x, t) \xrightarrow{\phi} (y, t) \xrightarrow{\text{pr}_1} y$$

où ϕ vérifie les hypothèses du théorème d'inversion locale, car sa matrice jacobienne s'écrit :

	$dx_1 \cdots dx_p$	$dt_1 \cdots dt_\ell$
dy_1	$\partial_x f(a)$	$\partial_t f(a)$
\vdots		
dy_p		
dt_1	0	II
\vdots		
dt_ℓ		

□

1.4.2. Immersions (rang $Df_a = \dim$ source)

PROPOSITION

Les deux propriétés suivantes sont équivalentes :

- i) Df_a est injective, c'est-à-dire de rang égal à la dimension de la source ;
- ii) la composition avec un difféomorphisme local convenable de l'espace but (au voisinage de $b = f(a)$) transforme f en l'injection canonique :

$$\mathbb{R}^n \times \{0\} \rightarrow \mathbb{R}^p = \mathbb{R}^n \times \mathbb{R}^k$$

Une application différentiable f vérifiant ces propriétés est appelée **immersion** (locale en a). On dit *immersion* tout court si f est une immersion locale en tout point de son ouvert de définition.

Preuve

L'implication ii) \Rightarrow i) résulte immédiatement de 1.2. Montrons la réciproque.

Dire que Df_a est injective, c'est dire que l'un au moins des mineurs $n \times n$ de la matrice jacobienne de f en a est inversible, et quitte à permuter les coordonnées au but on peut toujours supposer que c'est le cas du mineur du haut, noté $\partial_x g(a)$ ci-dessous :

	$dx_1 \cdots dx_n$
dy_1	$\partial_x g(a)$
\vdots	
dy_n	
dz_1	$\partial_x h(a)$
\vdots	
dz_k	

On a noté $x = (x_1, \dots, x_n)$ les coordonnées canoniques de l'espace source, $y = (y_1, \dots, y_n)$ les n « premières » coordonnées canoniques de l'espace but et

$z = (z_1, \dots, z_k)$ les $k (= p - n)$ « dernières » (après permutation). Les lettres g, h désignent les deux composantes de l'application

$$f : x \mapsto (y = g(x), z = h(x))$$

On remarque que l'application f se factorise en

$$f : x \xrightarrow{(\mathbb{I}, h)} (x, z) \xrightarrow{\phi} (y, z)$$

où l'application $\phi = g \times \mathbb{I}$ vérifie les hypothèses du théorème d'inversion locale en a .

On conclut en remarquant que l'application (\mathbb{I}, h) se factorise à son tour en $(\mathbb{I}, h) = \psi \circ j$, où j est l'inclusion canonique $x \mapsto (x, 0)$, tandis que ψ est le difféomorphisme évident $(x, z) \mapsto (x, z + h(x))$. \square

2. Point de vue géométrique sur les sous-variétés

2.0. Reformulation géométrique de la notion de lissité

Une propriété importante de la notion de lissité est d'être *invariante par difféomorphismes* : l'image d'une sous-variété par un difféomorphisme est une sous-variété (la fig. 1 en est une illustration).

Ce fait, qui n'est pas tout à fait évident sur la définition du chap. 1, se verra tout de suite sur la définition équivalente que voici.

DÉFINITION Une sous-ensemble $S \subset \mathbb{R}^n$ est dit **lisse au point** a ($a \in S$), de classe C^r , de **dimension** m (ou de **codimension** $p = n - m$) s'il existe un difféomorphisme de classe C^r :

$$\rho : U_a \rightarrow V$$

d'un voisinage U_a de $a \in \mathbb{R}^n$ dans un ouvert V de $\mathbb{R}^m \times \mathbb{R}^p = \mathbb{R}^n$, tel que :

$$\rho(S \cap U_a) = V \cap (\mathbb{R}^n \times 0) \quad (\text{cf. Fig. 2}).$$

Un tel difféomorphisme sera appelé **redressement local** de S au voisinage de a .

Un sous-ensemble qui est lisse en chacun de ses points est appelé **sous-variété**.

Équivalence avec la définition du chap. 1 (§ 2.3)

Tout sous-ensemble lisse au sens du chap. 1 peut évidemment être « redressé » par le difféomorphisme :

$$(x_1, \dots, x_n) \mapsto (x_1, \dots, x_m, x_{j_1} - g_1(x_1), \dots, (x_{j_p} - g_p(x_1)))$$

(les notations sont celles du § 1.4, formule (1.6), et l'on a posé $x_j = (x_{j_1}, \dots, x_{j_m})$).

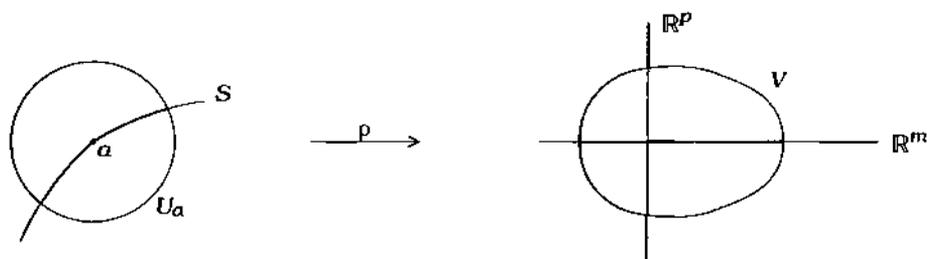


Figure 2. Redressement local d'une sous-variété.

Inversement, tout sous-ensemble S lisse au sens de la définition ci-dessus peut être défini localement par le système d'équations $(s_1(x_1, \dots, x_n) = 0, \dots, s_p(x_1, \dots, x_n) = 0)$, où l'on a noté s_1, \dots, s_p les p dernières composantes $s_1 = \rho_{m+1}, \dots, s_p = \rho_m$ d'un difféomorphisme $\rho = (\rho_1, \dots, \rho_n)$ redressant S : qu'il soit lisse au sens du chap. 1 résulte alors du théorème des fonctions implicites (chap. 1, § 3).

DÉFINITION Un tel système d'équations de $S \cap U_a$ est appelé **système d'équations locales de S au voisinage de a** . Par abus de langage il nous arrivera parfois d'appeler « système d'équations locales » de S le p -uplet de fonctions (s_1, \dots, s_p) constitué par les **premiers membres** d'un système d'équations locales.

Remarque Il résulte facilement de l'une comme de l'autre des définitions que si S est lisse en a , il l'est aussi en tous ses points suffisamment voisins de a .

Par ailleurs, comme le difféomorphisme de redressement transforme $S \cap U_a$ en un fermé de V , toute sous-variété S de \mathbb{R}^n est *localement fermée*, c'est-à-dire intersection d'un ouvert et d'un fermé.

2.0'. Sous-variétés de classe C^0

Si dans la définition ci-dessus on remplace le difféomorphisme ρ par un homéomorphisme, on obtient la notion d'ensemble lisse de classe C^0 [resp. sous-variété de classe C^0 , ou **sous-variété topologique**].

Exemple Le « demi-cône quadratique » (cornet de glace) :

$$S = \left\{ (x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 - z^2 = 0, z \geq 0 \right\}$$

est une sous-variété de classe C^0 (exercice : en construire un redressement de classe C^0).

Question On pourrait aussi imaginer une variante de la définition du chap. 1 où C^r serait remplacé par C^0 (c'est-à-dire que g serait supposée *continue* seulement). Quelle relation y a-t-il entre cette définition et celle que nous venons de donner ? Que pouvez-vous dire du demi-cône quadratique (ayant les trois axes de coordonnées pour génératrices) :

$$S = \left\{ (x, y, z) \in \mathbb{R}^3 \mid xy + yz + zx = 0, x + y + z \geq 0 \right\} ?$$

Dans toute la suite du paragraphe, la classe de différentiabilité (des sous-variétés et des applications) pourra s'entendre comme n'importe laquelle des classes pour

lesquelles le théorème des fonctions implicites est vrai (classe C^r avec $r \geq 1$, classe C^∞ , classe analytique C^ω).

2.1. Submersions et systèmes d'équations locales

PROPOSITION

L'image réciproque de 0 par une submersion $s : U \rightarrow \mathbb{R}^p$ (U ouvert de \mathbb{R}^n) est une sous-variété de codimension p .

Preuve

Conséquence immédiate de la proposition du § 1.4.1, cet énoncé peut être considéré comme une paraphrase du théorème des fonctions implicites (chap. 1, § 3.2). \square

Réciproque locale Inversement, tout sous-ensemble lisse en a peut être défini dans un voisinage U_a de a comme l'image réciproque de 0 par une submersion $s : U_a \rightarrow \mathbb{R}^p$.

Preuve

Prendre pour $s = (s_1, \dots, s_p)$ les premiers membres d'un système d'équations locales de S , comme nous en avons construit au § 2.0. \square

2.2. Immersions et paramétrages locaux

PROPOSITION

Tout sous-ensemble $S \subset \mathbb{R}^n$, lisse en a de dimension m , peut s'écrire localement au voisinage de a comme l'image d'une immersion :

$$\alpha : (\mathbb{R}^m, 0) \cdots \rightarrow (\mathbb{R}^n, a)$$

Preuve

Poser $\alpha = \rho^{-1} \circ j$, où ρ est un redressement local de S , et $j : \mathbb{R}^m = \mathbb{R}^m \times \{0\} \rightarrow \mathbb{R}^m \times \mathbb{R}^p$ l'inclusion canonique. \square

Lien avec la notion de paramétrage local (chap. 1, § 2.4)

En notant x_1, \dots, x_n les coordonnées canoniques de l'espace but \mathbb{R}^n , se donner l'application α équivaut à se donner une interprétation des variables x_i comme fonctions de m variables (t_1, \dots, t_m) (les coordonnées canoniques de l'espace source \mathbb{R}^m) :

$$x_1 = \alpha_1(t_1, \dots, t_m), \dots, x_n = \alpha_n(t_1, \dots, t_m) \quad (2.1)$$

Dire que α est une immersion équivaut à dire que de la famille $(d\alpha_1, \dots, d\alpha_n)$ on peut extraire une sous-famille libre de longueur m , $(d\alpha_{i_1}, \dots, d\alpha_{i_m})$. Le théorème d'inversion

locale (dans sa version du chap. 1 § 3) implique donc que dans l'interprétation (2.1) des x_i la correspondance entre (t_1, \dots, t_m) et (x_1, \dots, x_m) est un *changement de variables différentiable*. Autrement dit la formule (2.1) définit un *paramétrage local* de S , au sens du chap. 1 § 2.4. Réciproquement il est facile de vérifier que tout paramétrage local de S définit une immersion.

Réciproque locale de la proposition Pour toute immersion $\alpha : (\mathbb{R}^m, t_0) \cdots \rightarrow (\mathbb{R}^n, a)$ il existe un voisinage W de t_0 dans \mathbb{R}^m tel que $\alpha(W)$ soit lisse en a de dimension m .

Preuve

C'est une conséquence immédiate de la proposition du § 1.4.2. □

Attention Cette « réciproque locale » n'implique nullement que l'image globale d'une immersion soit toujours lisse !

2.3. Exemples d'immersions dont l'image n'est pas lisse

- i) $\alpha : \mathbb{R} \rightarrow \mathbb{R}^2, t \mapsto x = t^2, y = t - t^3$ (cf. Fig. 3i)).
- ii) $\beta = \alpha|_{]-\infty, 1[}$ (α comme en i)) (cf. Fig. 3ii)).

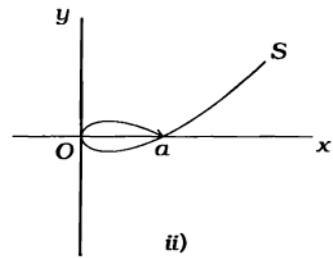
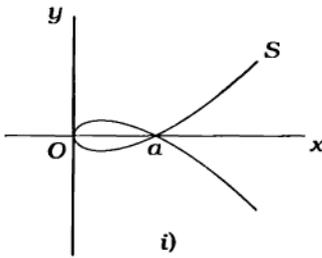


Figure 3. Images non lisses d'immersions.

La figure 4 montre comment restreindre les domaines de définition des immersions précédentes, pour que l'image soit lisse au voisinage de $a = (1, 0)$.

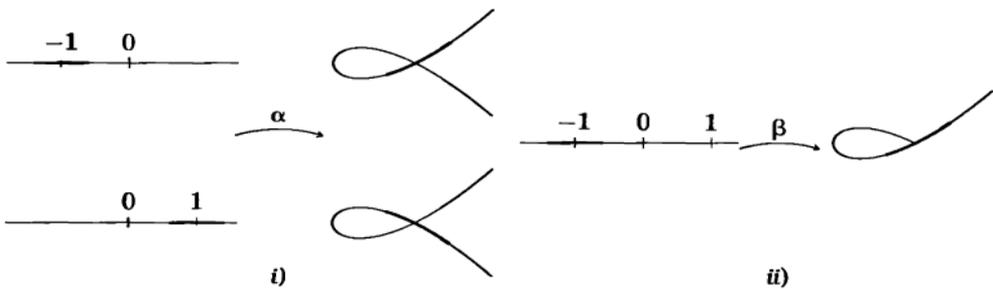


Figure 4. Lissité en a des images locales des immersions de la figure 3.

Réflexion sur les exemples précédents

Dans l'exemple i), la non-lissité de l'image en a s'explique par le fait que l'immersion n'est pas injective ($\alpha(+1) = \alpha(-1) = a$)⁷.

Dans l'exemple ii), l'immersion est injective mais elle n'est pas propre : le point $\beta(t)$ ne tend pas vers l'infini quand t tend vers le bord $+1$ de l'intervalle de définition de β .

Rappel Une application est dite **propre** si l'image réciproque de tout compact est compacte. En appelant « suite tendant vers l'infini » une suite qui sort de tout compact, on peut paraphraser cette définition en disant que toute suite tendant vers l'infini à la source doit avoir pour image une suite tendant vers l'infini au but.

PROPOSITION

Toute immersion injective et propre a pour image une sous-variété.

Preuve

Il s'agit de montrer que pour tout t_0 dans l'espace source de α , l'image de α est lisse en $a = \alpha(t_0)$. Soit $K_\varepsilon = \alpha^{-1}(\overline{B}_\varepsilon)$, où $\overline{B}_\varepsilon \subset \mathbb{R}^n$ désigne la boule fermée de centre a et de rayon ε . K_ε est compact puisque α est propre. De plus $\bigcap_{\varepsilon>0} K_\varepsilon = \alpha^{-1}(a) = \{t_0\}$ puisque α est injective. Le diamètre du compact K_ε tend donc vers 0 quand $\varepsilon \rightarrow 0$, de sorte que pour ε assez petit, K_ε est inclus dans un ouvert W du type de la « réciproque locale » du § 2.2. \square

3. Applications différentiables entre sous-variétés

Dans tout ce paragraphe la classe de différentiabilité C^r des sous-variétés sera fixée une fois pour toutes à une valeur $r \geq 1$ (éventuellement $r = \infty$ ou $r = \omega$). Par contre les dimensions des espaces ambiants ne seront pas fixées, c'est-à-dire que les diverses « sous-variétés » seront des sous-variétés de \mathbb{R}^n pour diverses valeurs de n (que d'ailleurs nous omettrons souvent de préciser).

Nous supposerons connue la notion de *système de coordonnées locales sur une sous-variété* (cf. chap. 1, § 1.4). Abstraction faite de l'espace ambiant, qui a servi à la définir, nous en retiendrons surtout les propriétés suivantes.

3.0. Rappel sur les systèmes de coordonnées locales

Soit (x_1, \dots, x_m) un système de coordonnées locales sur la sous-variété S ⁸. Les variables x_1, \dots, x_m sont interprétées comme des fonctions sur S , ayant comme domaine

⁷ Pour d'autres exemples analogues. cf. exercice 0.2 du chap. 1.

⁸ Il s'agit donc d'un m -uplet comme ceux notés (x_1, \dots, x_m) ou ceux notés (t_1, \dots, t_m) au chap. 1, § 2.4. Les lettres x_i n'ont donc plus nécessairement de rapport direct avec les coordonnées canoniques de l'espace ambiant.

de définition commun un ouvert U de S appelé *ouvert de définition* du système de coordonnées locales⁹. Cette interprétation définit ce qu'on appelle une *carte locale*¹⁰ de S , bijection $x : U \rightarrow \underline{U}$ entre U et un ouvert \underline{U} de \mathbb{R}^m (« l'ouvert des valeurs » du système de coordonnées locales).

Soit $\tilde{x} : \tilde{U} \rightarrow \tilde{\underline{U}}$ une autre carte locale de S dont le domaine de définition \tilde{U} a une intersection non vide avec U . Pour simplifier les notations on supposera même que $\tilde{U} = U$ (cas auquel on peut toujours se ramener en restreignant x et \tilde{x} à $U \cap \tilde{U}$, renomé U). Alors la bijection composée $\tilde{x} \circ x^{-1} : \underline{U} \rightarrow \tilde{\underline{U}}$ est un *difféomorphisme* (de classe C^r) appelé *difféomorphisme de changement de carte*. Autrement dit le changement de coordonnées $(x_1, \dots, x_m) \leftrightarrow (\tilde{x}_1, \dots, \tilde{x}_m)$ est un *changement de variables* de classe C^r .

3.1. Applications différentiables d'une sous-variété dans une autre

Soient S une sous-variété de dimension m , et S' une sous-variété de dimension m' . Soit (x_1, \dots, x_m) un système de coordonnées locales de S , défini sur un ouvert U de S , et soit $(y_1, \dots, y_{m'})$ un système de coordonnées locales de S' , défini sur un ouvert U' de S' . Les cartes locales correspondantes seront notées $x : U \rightarrow \underline{U}$ et $y : U' \rightarrow \underline{U}'$. Puisqu'il s'agit de bijections, se donner une application $f : U \rightarrow U'$ équivaut à se donner une application $\underline{f} : \underline{U} \rightarrow \underline{U}'$. La correspondance entre f et \underline{f} est donnée par le « diagramme commutatif¹¹ » suivant :

$$\begin{array}{ccc} U & \xrightarrow{f} & U' \\ \downarrow x & & \downarrow y \\ \mathbb{R}^m \supset \underline{U} & \xrightarrow{\underline{f}} & \underline{U}' \subset \mathbb{R}^{m'} \end{array} \quad (3.1)$$

Autrement dit, se donner f équivaut à se donner une interprétation des variables $y_1, \dots, y_{m'}$ comme fonctions des variables x_1, \dots, x_m :

$$\begin{cases} y_1 & := & \underline{f}_1(x_1, \dots, x_m) \\ \dots & & \\ y_{m'} & := & \underline{f}_{m'}(x_1, \dots, x_m) \end{cases} \quad (3.1)^{\star 12}$$

où $\underline{f} = (\underline{f}_1, \dots, \underline{f}_{m'})$ est une application de \underline{U} dans \underline{U}' .

On résume cette situation en disant que l'application $\underline{f} = y \circ f \circ x^{-1}$ est *l'expression de l'application f dans les systèmes de coordonnées* $(x_1, \dots, x_m), (y_1, \dots, y_{m'})$.

⁹ Cet ouvert était noté $S \cap U$ au § 2.4 du chap. 1.

¹⁰ Le point de vue des « maths modernes » conduit à considérer « carte » et « système de coordonnées » comme synonymes. Selon le point de vue un peu plus fin exposé à la fin du chap.1 (Commentaire linguistique), « système de coordonnées » est une notion *métamathématique*, tandis que « carte » est la notion mathématique qui s'en déduit « en faisant abstraction des variables ».

¹¹ (façon suggestive de dire que $y \circ f = \underline{f} \circ x$)

¹² On passe de (3.1)^{*} à (3.1) en ré-interprétant les y_i comme fonctions sur U , par les substitutions $y_i := y_i \circ f$.

Effet d'un changement de coordonnées

Soit $\tilde{f} = \tilde{y} \circ f \circ \tilde{x}^{-1}$ l'expression de f dans d'autres cartes \tilde{x}, \tilde{y} (que pour simplifier les notations nous supposons définies dans les mêmes ouverts respectifs que x et y). Alors on passe évidemment de f à \tilde{f} par la formule de composition :

$$\tilde{f} = (\tilde{y} \circ y^{-1}) \circ f \circ (\tilde{x} \circ x^{-1})^{-1}$$

où l'on reconnaît dans les parenthèses les *difféomorphismes de changement de carte à la source et au but*. Il en résulte que la propriété pour f d'avoir une expression \underline{f} différentiable de classe C^r (dans les cartes x, y) ne dépend pas du choix des cartes. Nous pouvons donc énoncer la définition suivante.

DÉFINITION Une application $f : S \rightarrow S'$ est dite **différentiable en un point** $a \in S$ (resp. **différentiable de classe C^r au voisinage de** $a \in S$) si son expression $\underline{f} = y \circ f \circ x^{-1}$ est différentiable en a (resp. différentiable de classe C^r dans un voisinage U de a), où :

- x est une (quelconque) carte locale de S définie dans un voisinage U de a ;
- y est une (quelconque) carte locale de S' définie dans un voisinage V de $b = f(a)$.

L'application f est dite **différentiable de classe C^r** si elle l'est au voisinage de tout point $a \in S$.

L'affirmation suivante résulte immédiatement de cette définition.

Principe de différentiabilité des applications composées

La composée de deux applications différentiables (de classe C^r) est différentiable (de classe C^r).

Remarque Considérons le cas particulier où chacune des sous-variétés S resp. S' est une sous-variété de codimension 0, c'est-à-dire un ouvert de \mathbb{R}^n resp. $\mathbb{R}^{n'}$. On peut alors prendre comme carte x resp. y l'application identique de cet ouvert dans lui-même, ce qui montre que la notion d'application différentiable (de classe C^r) entre sous-variétés coïncide dans ce cas avec la notion, définie au § 1. d'application différentiable entre ouverts d'espaces numériques.

Images réciproques des fonctions différentiables par une application différentiable

Étant donné une sous-variété S , de classe C^r , l'ensemble des fonctions différentiables de classe C^r sur S est un anneau¹³ que l'on note $C^r(S)$. Étant donné une application différentiable $f : S \rightarrow S'$ de classe C^r , et une fonction $\varphi \in C^r(S')$, la fonction composée $\varphi \circ f$ appartient à $C^r(S)$ en vertu du principe de différentiabilité des applications composées. L'application de $C^r(S')$ dans $C^r(S)$ qui à φ associe $\varphi \circ f$ est notée :

$$\begin{aligned} f^* : C^r(S') &\rightarrow C^r(S) \\ \varphi &\mapsto \varphi \circ f \end{aligned}$$

La fonction $f^*(\varphi) = \varphi \circ f$ est appelée **image réciproque de la fonction φ par l'application f** . Il est immédiat de vérifier que l'application f^* est un *homomorphisme d'anneaux*. On l'appelle homomorphisme « image réciproque par f ».

¹³ Autrement dit les sommes et produits de fonctions de classe C^r sur S sont des fonctions de classe C^r sur S . Pour le voir, on se ramène (à l'aide d'une carte locale) au cas des fonctions sur un ouvert U de \mathbb{R}^m .

Comme la différentiabilité est une propriété locale, tout ce qui précède peut être ré-énoncé au voisinage d'un point $a \in S$, $b = f(a) \in S'$: étant donné une application $f : S, a \cdots \rightarrow S', b$ ¹⁴ de classe C^r au voisinage de a , on peut voir l'homomorphisme « image réciproque par f » comme un homomorphisme $f^* : C_{S',b}^r \rightarrow C_{S,a}^r$ de l'anneau des fonctions de classe C^r sur S' au voisinage de b dans l'anneau des fonctions de classe C^r sur S au voisinage de a ¹⁵.

3.2. Le langage des « catégories »

Par leur propriété générale de « stabilité par composition », les applications différentiables de sous-variétés sont un exemple d'une structure très fréquente en mathématiques, la structure de *catégorie*.

Nous avons défini une classe d'**objets** (les sous-variétés)¹⁶. Pour tout couple de tels objets S, S' nous venons de définir un ensemble de **flèches** $f : S \rightarrow S'$ entre ces objets (les applications différentiables de S dans S').

De plus, nous savons composer les flèches : notant $\mathcal{F}(S, S')$ l'ensemble des flèches de S dans S' , nous savons associer à tout couple de flèches $f \in \mathcal{F}(S, S')$, $g \in \mathcal{F}(S', S'')$ une flèche $g \circ f \in \mathcal{F}(S, S'')$.

La loi de composition des flèches vérifie les axiomes d'*associativité* :

$$h \circ (g \circ f) = (h \circ g) \circ f$$

et d'*existence d'éléments neutres* :

il existe dans chaque ensemble $\mathcal{F}(S, S)$ un élément privilégié noté \mathbb{I}_S , tel que :

$$\mathbb{I}_S \circ f = f \quad [\text{resp. } g \circ \mathbb{I}_S = g]$$

pour toute flèche f ayant S comme objet « but » [resp. toute flèche g ayant S comme objet « source »].

On résume ces propriétés en disant que la classe d'objets considérée, avec les ensembles de flèches entre ces objets, forme une **catégorie**.

Vous connaissez déjà beaucoup d'exemples de catégories : la *catégorie des espaces vectoriels* (dont les objets sont les espaces vectoriels et les flèches les applications linéaires) ; la *catégorie des groupes* (dont les objets sont les groupes et les flèches les homomorphismes de groupes) ; la *catégorie des espaces topologiques* (dont les objets sont les espaces topologiques et les flèches les applications continues)...

Nous venons de définir la **catégorie des sous-variétés de classe C^r** , dont les flèches sont les applications différentiables de classe C^r entre sous-variétés. Cette catégorie contient une sous-catégorie remarquable, dont les objets sont les ouverts

¹⁴ La notation $\cdots \rightarrow$ indique, comme au § 1.2, que l'application considérée n'a besoin d'être définie qu'au voisinage du point considéré.

¹⁵ De façon précise, la notation $C_{S,a}^r$ désigne l'ensemble des germes en a de fonctions de classe C^r sur S , c'est-à-dire les fonctions (de classe C^r) définies dans un voisinage de a que l'on n'a pas jugé utile de préciser : nous reviendrons de façon plus précise sur cette notion de « germe » au début du chap. 5 (note de bas de page).

¹⁶ En réalité un objet de la catégorie des sous-variétés est le couple formé par la sous-variété et son espace ambiant. La notation abrégée S sous-entend ici la donnée de l'espace ambiant.

d'espaces numériques (sous-variétés de codimension 0), les flèches entre deux tels objets étant les applications différentiables de classe C^r au sens du § 1.1. Nous avons déjà remarqué au § 3.1 qu'étant donné un couple d'objets (U, U') de cette sous-catégorie, l'ensemble $\mathcal{F}(U, U')$ des flèches de U vers U' au sens de la « grande » catégorie (c'est-à-dire au sens du § 3.1) coïncide avec l'ensemble des flèches de U vers U' au sens de la « petite » catégorie (c'est-à-dire au sens du § 1) : on dit que cette dernière est une *sous-catégorie pleine* de la « grande » catégorie.

À la place du mot « flèche » (pourtant joli et imagé) on emploie plus couramment le mot « **morphisme** », qui fait plus savant et permet des combinaisons avec d'autres racines grecques.

Un morphisme $f \in \mathcal{F}(S, S')$ est appelé **isomorphisme** s'il admet un « morphisme inverse », c'est-à-dire s'il existe une flèche $g \in \mathcal{F}(S', S)$ telle que $g \circ f = \mathbb{I}_S$, $f \circ g = \mathbb{I}_{S'}$.

Exemples Les isomorphismes de la catégorie des espaces vectoriels, ou de la catégorie des groupes, sont ce que tous les cours d'algèbre enseignent sous ce nom. Les isomorphismes de la catégorie des espaces topologiques sont ce qu'on appelle en topologie les homéomorphismes.

Les isomorphismes de la catégorie des sous-variétés différentiables (de classe C^r) sont appelés **difféomorphismes** (de classe C^r).

Exemples de difféomorphismes : une *carte locale* de S n'est rien d'autre qu'un difféomorphisme x entre un ouvert U de S (le domaine de la carte) et un ouvert \underline{U} de \mathbb{R}^m . Si $\tilde{x} : U \rightarrow \tilde{U}$ est une autre carte ayant même domaine de définition, le difféomorphisme composé $\tilde{x} \circ x^{-1} : \underline{U} \rightarrow \tilde{U}$ est ce qu'au § 3.0 nous avons appelé le difféomorphisme de *changement de carte*.

3.3. Retour à l'espace ambiant

Plongement canonique et rétractions

Soit S une sous-variété de \mathbb{R}^n , et soit U un ouvert de \mathbb{R}^n contenant S .

L'injection canonique $i : S \rightarrow U$ (encore appelée **plongement canonique**) est évidemment un morphisme de sous-variétés.

On appelle **rétraction** (différentiable) de U sur S un morphisme $r : U \rightarrow S$ tel que $r \circ i = \mathbb{I}_S$.

Lemme de rétraction (version globale) Pour toute sous-variété $S \subset \mathbb{R}^n$ il existe un ouvert U contenant S et une rétraction $r : U \rightarrow S$.

Voici une variante locale du lemme, plus facile à démontrer et se prêtant donc mieux à certaines généralisations (par exemple elle reste vraie dans la catégorie des *variétés analytiques complexes*, ce qui n'est pas le cas de la version globale).

Lemme de rétraction (version locale) Pour tout point a d'une sous-variété $S \subset \mathbb{R}^n$ il existe un ouvert U_a de \mathbb{R}^n contenant a et une rétraction $r : U_a \rightarrow S \cap U_a$.

Preuve

Dans la situation « redressée » où $S = \mathbb{R}^m \times \{0\}$ la projection :

$$\pi : (x_1, \dots, x_m, x_{m+1}, \dots, x_n) \mapsto (x_1, \dots, x_m, 0, \dots, 0)$$

répond évidemment à la question. Dans le cas général il suffira de poser $r = \rho^{-1} \circ \pi \circ \rho$, où ρ est un difféomorphisme de redressement local de S (cf. § 2.0). \square

Problème En utilisant la question 3 du problème de l'étude 1, démontrez la version globale du lemme de rétraction lorsque S est une *hypersurface* (idée : définir $r(x)$ comme le « pied de la normale à S issue du point x »). Généralisez en codimension quelconque.

Extensions ambiantes d'applications différentiables

Étant donné une sous-variété S de \mathbb{R}^n , et une sous-variété S' de $\mathbb{R}^{n'}$, considérons un diagramme commutatif :

$$\begin{array}{ccc} S & \xrightarrow{f} & S' \\ \downarrow i & & \downarrow i' \\ U & \xrightarrow{\check{f}} & U' \end{array} \quad (3.2)$$

où i est le plongement canonique de S dans un ouvert U de \mathbb{R}^n ;

et i' est le plongement canonique de S' dans un ouvert U' de $\mathbb{R}^{n'}$.

On résumera cette situation en disant que l'application \check{f} est une **extension ambiante** de f , ou que f est la **restriction** de \check{f} . La donnée de \check{f} détermine f sans ambiguïté, c'est pourquoi nous disons *la restriction*. Par contre une application f donnée peut avoir beaucoup d'extensions différentes.

PROPOSITION

Pour qu'une application $f : S \rightarrow S'$ soit différentiable, il faut et il suffit qu'elle admette une extension ambiante différentiable.

Preuve

Il suffit Quitte à remplacer U' par un ouvert plus petit¹⁷ on peut supposer qu'il existe une rétraction $r' : U' \rightarrow S'$. Alors $f = (r' \circ i') \circ f = r' \circ (i' \circ f) = r' \circ \check{f} \circ i$, qui est bien une application différentiable puisque composée d'applications différentiables¹⁸.

Il faut Cette fois on se donne f , et il s'agit de construire \check{f} (les ouverts U, U' n'étant pas non plus prédéterminés). On pourra par exemple choisir une rétraction $r : U \rightarrow S$, et poser $\check{f} = i' \circ f \circ r$ ¹⁹. \square

Cas particulier des fonctions Dans le cas particulier $S' = U' = \mathbb{R}$ nous venons de démontrer qu'une fonction sur S est différentiable si et seulement si elle est la restriction d'une fonction différentiable définie dans un ouvert U de l'espace ambiant. La version locale de ce résultat a déjà été démontrée au chap. 1 (§ 2.5, lemme de restriction). Il est

¹⁷ Pour tout ouvert V de $\mathbb{R}^{n'}$ contenant S' on obtient un diagramme analogue à (3.2) en remplaçant U' par $U' \cap V$ et U par $U \cap \check{f}^{-1}(V)$.

¹⁸ On pourrait éviter le recours au lemme de rétraction *global* en remarquant que la propriété à démontrer (différentiabilité de f) est une propriété *locale*.

¹⁹ Remarquez que l'extension \check{f} ainsi construite a une propriété très particulière : son image est incluse dans S' . Notre construction est donc loin de donner toutes les extensions possibles !

intéressant d'en relire la preuve, pour y reconnaître une *version déguisée* de la preuve ci-dessus : où était cachée la rétraction r ?

Cas des difféomorphismes

Étant donné un difféomorphisme f de sous-variétés, il n'est pas si facile (ni toujours possible) d'en construire une « extension ambiante » \tilde{f} qui soit un *difféomorphisme* (si l'on applique la construction ci-dessus à f puis à f^{-1} , les applications \tilde{f} et (\tilde{f}^{-1}) obtenues ne sont pas inverses l'une de l'autre !). Néanmoins on a le résultat *local* suivant.

PROPOSITION

Étant donné deux sous-variétés S et S' , dont les espaces ambiants ont même dimension, et un point $a \in S$, tout difféomorphisme f entre S et S' peut s'étendre localement en un difféomorphisme local \tilde{f} des espaces ambiants, faisant commuter le diagramme :

$$\begin{array}{ccc} S, a & \xrightarrow{\quad f \quad} & S', b \\ \downarrow \iota & & \downarrow \iota' \\ U, a & \xrightarrow{\quad \tilde{f} \quad} & U', b \end{array} \quad (3.3)$$

Preuve

En faisant des « redressements locaux » à la source et au but on se ramène à une situation facile (les détails sont laissés au lecteur). \square

4. Le « foncteur tangent »

4.1. Applications tangentés

Relisons la définition de la différentielle d'une application (§ 1.1), en l'interprétant géométriquement dans l'esprit du chap. 1 (§ 2). Dans les formules (1.1)', les variables dx_1, \dots, dx_n sont interprétées comme coordonnées d'un *vecteur tangent* $\vec{u} \in T_a \mathbb{R}^n$ (« vecteur lié » d'origine a), et les variables dy_1, \dots, dy_p comme coordonnées d'un vecteur tangent $\vec{v} \in T_b \mathbb{R}^p$ (où $b = f(a)$), image de \vec{u} par l'application linéaire Df_a . Ainsi interprétée, cette application linéaire est rebaptisée **application tangente à f** en a , et notée

$$T_a f : T_a \mathbb{R}^n \rightarrow T_b \mathbb{R}^p$$

Considérant \vec{u} comme vecteur vitesse d'un point mobile $\lambda : \mathbb{R}, 0 \cdots \rightarrow \mathbb{R}^n, a$ on voit, grâce au principe de composition des différentielles (§ 1.2), que le vecteur $\vec{v} = T_a f(\vec{u})$ n'est autre que le *vecteur vitesse du point mobile* $f \circ \lambda : \mathbb{R}, 0 \cdots \rightarrow \mathbb{R}^p, b$, image du point mobile λ par l'application f .

Cette façon de ré-interpréter la différentielle a l'avantage de « bien passer aux sous-variétés ». On a vu à la fin du § 3.3 que tout morphisme de sous-variétés $f : S \rightarrow S'$ peut être considéré localement au voisinage de tout point (et même globalement, mais peu importe) comme la restriction d'une application différentiable des espaces ambiants.

Considérons donc une telle application différentiable :

$$\check{f} : \mathbb{R}^n, a \cdots \rightarrow \mathbb{R}^p, b \quad (b = f(a))$$

envoyant S dans S' et faisant commuter le diagramme :

$$\begin{array}{ccc} S, a \cdots & \xrightarrow{f} & S', b \\ \cap & & \cap \\ \mathbb{R}^n, a \cdots & \xrightarrow{\check{f}} & \mathbb{R}^p, b. \end{array}$$

PROPOSITION-DÉFINITION

L'application linéaire tangente $T_a \check{f} : T_a \mathbb{R}^n \rightarrow T_b \mathbb{R}^p$ définit par restriction une application linéaire de $T_a S$ dans $T_b S'$ qui ne dépend que de $f : S, a \cdots \rightarrow S', b$. On la note :

$$T_a f : T_a S \rightarrow T_b S'$$

et on l'appelle **application linéaire tangente à f en a** .

Preuve

(cf. fig.5) Comme on l'a vu au chap. 1 (§ 2.3), tout vecteur $\vec{u} \in T_a S$ est le vecteur vitesse d'un point mobile $\lambda : \mathbb{R}, 0 \cdots \rightarrow \mathbb{R}^n, a$ tel que l'image de λ soit dans S ; il est alors clair que l'application $\check{f} \circ \lambda$ a son image dans S' , et ne dépend de \check{f} que via sa restriction f à S . \square

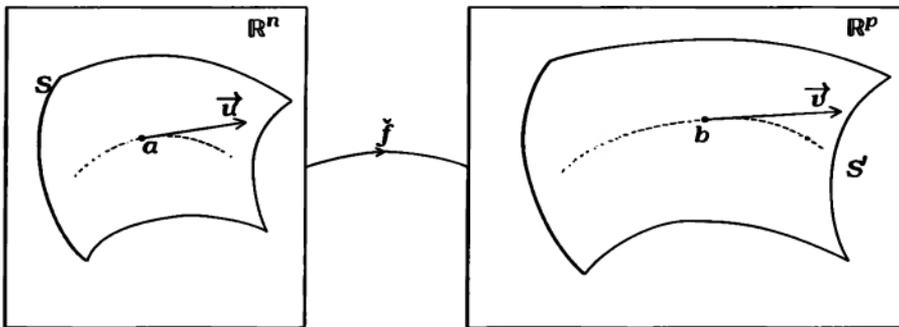


Figure 5. Application tangente.

Composition des applications tangentés

Soient $f : S, a \cdots \rightarrow S', b$ et $g : S', b \cdots \rightarrow S'', c$ deux morphismes de sous-variétés pointées. Alors $T_a(g \circ f) = T_b g \circ T_a f$.

Nous savons donc maintenant associer à toute sous-variété pointée un espace vectoriel (son espace tangent) et à tout morphisme de variétés pointées un homomorphisme d'espaces vectoriels (l'application linéaire tangente), d'une façon qui préserve la composition. Une telle correspondance est ce qu'on appelle un *foncteur* (de la catégorie des sous-variétés pointées dans la catégorie des espaces vectoriels).

4.2. Le langage des « foncteurs »

Considérons deux catégories \mathcal{C} et \mathcal{C}' , appelées respectivement catégorie « source » et catégorie « but ».

Un **foncteur covariant** T de \mathcal{C} dans \mathcal{C}' est une correspondance qui :

- i) à tout objet S de \mathcal{C} associe un objet TS de \mathcal{C}' ;
- ii) à toute flèche $f \in \mathcal{F}_{\mathcal{C}}(S, S')$ associe une flèche $Tf \in \mathcal{F}_{\mathcal{C}'}(TS, TS')$.

De plus, cette correspondance doit vérifier les axiomes suivants :

$$\begin{aligned} (0) \quad T(\mathbb{1}_S) &= \mathbb{1}_{TS} \\ (1) \quad T(g \circ f) &= Tg \circ Tf. \end{aligned}$$

Un **foncteur contravariant** de \mathcal{C} dans \mathcal{C}' est défini par des axiomes similaires :

- i) inchangé ;
- ii)* à toute flèche $f \in \mathcal{F}_{\mathcal{C}}(S, S')$ est associée une flèche $Tf \in \mathcal{F}_{\mathcal{C}'}(TS', TS)$ (la flèche image « va dans l'autre sens » !).

$$\begin{aligned} (0) \quad &\text{inchangé ;} \\ (i)* \quad &T(g \circ f) = Tf \circ Tg. \end{aligned}$$

Exemple de foncteur contravariant \mathcal{C} = catégorie des sous-variétés ; \mathcal{C}' = catégorie des anneaux (avec comme morphismes les homomorphismes d'anneaux).

À toute sous-variété S associons l'anneau $\mathcal{C}^\infty(S)$ des fonctions différentiables sur S (cf. § 3.1).

À tout morphisme $f : S \rightarrow S'$ associons l'homomorphisme d'anneaux :

$$\begin{aligned} f^* : \mathcal{C}^\infty(S') &\rightarrow \mathcal{C}^\infty(S) \\ \varphi &\mapsto f^*\varphi := \varphi \circ f \quad (\text{cf. § 3.1}).^{20} \end{aligned}$$

Il est immédiat de vérifier que la correspondance ainsi définie est un *foncteur contravariant de la catégorie des sous-variétés dans la catégorie des anneaux*.

Remarque Il résulte immédiatement des axiomes que tout foncteur (covariant ou contravariant) transforme les isomorphismes de \mathcal{C} en isomorphismes de \mathcal{C}' .

Exemples Un difféomorphisme de sous-variétés pointées induit un isomorphisme des espaces vectoriels tangents ; un difféomorphisme de sous-variétés induit un isomorphisme de leurs anneaux de fonctions différentiables.

Commentaire

Inspirée par le souci de clarifier et mettre en ordre les idées, l'invocation de grandes notions générales comporte un double danger : chez certains étudiants que rebute l'abstraction elle peut inspirer de l'effroi ; chez d'autres elle peut encourager une tendance malsaine à se gargariser de grands mots. J'aimerais que la référence fréquente aux « catégories » et « foncteurs » dans ce cours ne suscite ni l'une ni l'autre réaction, mais soit comprise comme une *incitation à se libérer l'esprit* : le calcul différentiel sur les variétés est un sujet si riche qu'il est facile de s'y perdre, et chaque fois qu'un raisonnement invoquera la « fonctorialité » de telle ou telle correspondance, le lecteur saura qu'il peut oublier les aspects trop fins du sujet (comme ceux qui sont liés au théorème des fonctions implicites).

²⁰ Nous avons déjà utilisé le symbole $:=$ comme une instruction de **substitution** (substituer l'expression écrite à droite à la variable écrite à gauche). Mais on l'utilise aussi (c'est le cas ici) pour une **définition** (ici la définition de f^*).

4.3. Représentation des vecteurs tangents et des applications tangentes dans des cartes

Par raison de functorialité, toute carte locale $x : S, a \cdots \rightarrow \mathbb{R}^m, \underline{a}$ d'une sous-variété S à m dimensions induit un *isomorphisme d'espaces vectoriels* :

$$T_a x : T_a S \xrightarrow{\cong} T_{\underline{a}} \mathbb{R}^m = \mathbb{R}^m$$

qui à chaque vecteur tangent $\vec{u} \in T_a S$ associe un m -uplet de nombres $(dx_1(\vec{u}), \dots, dx_m(\vec{u}))$, les **coordonnées du vecteur \vec{u} dans la carte x** .

Soit $\tilde{x} : S, a \cdots \rightarrow \mathbb{R}^m, \underline{\tilde{a}}$ une autre carte locale au voisinage du même point a . En appliquant le foncteur tangent au diagramme commutatif :

$$\begin{array}{ccc} & S, a & \\ \swarrow x & \cdots & \searrow \tilde{x} \\ \mathbb{R}^m, \underline{a} & \xrightarrow{\tilde{x} \circ x^{-1}} & \mathbb{R}^m, \underline{\tilde{a}} \end{array}$$

on voit que les coordonnées d'un vecteur tangent dans la carte \tilde{x} se déduisent de ses coordonnées dans la carte x par l'application linéaire tangente $T_{\underline{a}}(\tilde{x} \circ x^{-1})$:

$$d\tilde{x}_i = \sum_j \frac{\partial (\tilde{x} \circ x^{-1})_i}{\partial x_j}(\underline{a}) dx_j$$

Dans le langage du chap. 1, on peut traduire cette formule en disant que $d\tilde{x}_i$ se déduit de (dx_1, \dots, dx_m) en différentiant l'expression de \tilde{x}_i comme fonction des variables x_1, \dots, x_m .

Considérons maintenant un morphisme $f : S, a \cdots \rightarrow S', b$ de sous-variétés pointées de dimensions respectives m et m' , et soit x (resp. y) une carte locale de S au voisinage de a (resp. une carte locale de S' au voisinage de b). Notant, comme au § 3.1, $f : \mathbb{R}^m, \underline{a} \cdots \rightarrow \mathbb{R}^{m'}, \underline{b}$ l'expression de f dans les cartes x, y , considérons le diagramme commutatif :

$$\begin{array}{ccc} S, a & \cdots \xrightarrow{f} & S', b \\ \parallel \downarrow x & & \parallel \downarrow y \\ \mathbb{R}^m, \underline{a} & \cdots \xrightarrow{f} & \mathbb{R}^{m'}, \underline{b} \end{array}$$

En lui appliquant le foncteur tangent on en déduit un diagramme commutatif :

$$\begin{array}{ccc} T_a S & \xrightarrow{T_a f} & T_b S' \\ \parallel \downarrow T_a x & & \parallel \downarrow T_b y \\ \mathbb{R}^m & \xrightarrow{D_{\underline{a}} f} & \mathbb{R}^{m'} \end{array}$$

L'application linéaire $D_{\underline{a}}f$ donne les coordonnées $dy_1, \dots, dy_{m'}$ du vecteur $T_{\underline{a}}f(\vec{u})$ (dans la carte y) en fonction des coordonnées dx_1, \dots, dx_m du vecteur \vec{u} (dans la carte x). Son expression s'obtient tout simplement en différenciant au point \underline{a} l'expression (3.1)* des variables $y_1, \dots, y_{m'}$ comme fonctions des variables x_1, \dots, x_m .

4.4. Applications différentiables de rang maximal

ou : le retour du « théorème des fonctions implicites »

Transposé aux sous-variétés, le théorème d'inversion locale s'énonce ainsi :

Théorème Soit $f : S, a \cdots \rightarrow S', b$ un morphisme de sous-variétés pointées dont l'application tangente $T_{\underline{a}}f$ est un isomorphisme. Alors f est un **difféomorphisme local** en s (difféomorphisme entre un voisinage de a dans S et un voisinage de b dans S').

Preuve

En composant f à la source et au but avec une carte d'un voisinage de a [resp. de b], on est ramené au théorème d'inversion locale usuel dans \mathbb{R}^m . \square

Pour $S' = \mathbb{R}^m$ ce théorème fournit une caractérisation « différentielle » des cartes locales (ou systèmes de coordonnées locales) d'une sous-variété.

Caractérisation différentielle des systèmes de coordonnées locales

Sur une sous-variété S de dimension m , un m -uplet (x_1, \dots, x_m) de fonctions numériques $x_i : S, a \cdots \rightarrow \mathbb{R}$ définit un système de coordonnées locales au voisinage de a si (et seulement si) les différentielles dx_1, \dots, dx_m au point a sont linéairement indépendantes.

Corollaire Tout m' -uplet de fonctions différentiables sur S , dont les différentielles en a sont linéairement indépendantes, peut être complété en un système de coordonnées locales sur S au voisinage de a .

Preuve

Il suffit d'appliquer le théorème de la base incomplète à l'espace vectoriel des fonctions linéaires sur $T_a S$, et de remarquer que toute fonction linéaire sur $T_a S$ est de la forme $d\varphi_a$ avec $\varphi \in C_{S,a}^r$. \square

Considérons maintenant le cas où les variétés source et but sont de dimensions différentes : $\dim S = m, \dim S' = m'$.

Un morphisme $f : S, a \cdots \rightarrow S', b$ est appelé *submersion* [resp. *immersion*] si son application linéaire tangente est *surjective* [resp. *injective*].

PROPOSITION

f est une submersion [resp. une immersion] si et seulement si son expression dans des coordonnées locales convenables x_1, \dots, x_n de S , $y_1, \dots, y_{m'}$ de S' s'écrit :

i) (submersion : $m \geq m' = \text{rang } T_a f$) :

$$\begin{cases} y_1 & = & x_1 \\ & \dots & \\ y_{m'} & = & x_{m'} \end{cases}$$

ii) (immersion : $\text{rang } T_a f = m \leq m'$) :

$$\begin{cases} y_1 & = & x_1 \\ & \dots & \\ y_m & = & x_m \\ y_{m+1} & = & 0 \\ & \dots & \\ y_{m'} & = & 0. \end{cases}$$

Preuve

On pourrait là encore se ramener, par composition avec des cartes locales à la source et au but, aux cas déjà étudiés au § 1.4. Mais pour s'habituer au langage des coordonnées locales, il est intéressant de récrire la démonstration.

i) *Cas des submersions* Soit $y_1, \dots, y_{m'}$ un système de coordonnées locales sur S' au voisinage de b . Posons $x_1 = y_1 \circ f, \dots, x_{m'} = y_{m'} \circ f$. Puisque l'application $T_a f$ est de rang m' , les formes linéaires $dx_{1a}, \dots, dx_{m'a}$ sont linéairement indépendantes. On peut donc compléter $(x_1, \dots, x_{m'})$ en un système de coordonnées locales (x_1, \dots, x_m) de S au voisinage de a , et dans les cartes x, y ainsi construites l'application f a évidemment la forme indiquée.

ii) *Cas des immersions* Soit $y_1, \dots, y_{m'}$ un système de coordonnées locales sur S' au voisinage de b . Puisque l'application tangente est de rang m , on peut extraire du m' -uplet de formes linéaires $d(y_1 \circ f)_a, \dots, d(y_{m'} \circ f)_a$ m formes linéairement indépendantes et pas davantage. Quitte à permuter les y_i , on peut supposer que ce sont les m premières. Les fonctions $x_1 = y_1 \circ f, \dots, x_m = y_m \circ f$ définissent alors un système de coordonnées locales sur S , et l'expression de f dans ces coordonnées s'écrit :

$$\begin{aligned} y_1 &= x_1 \\ &\dots \\ y_m &= x_m \\ y_{m+1} &= h_1(x_1, \dots, x_m) \\ &\dots \\ y_{m'} &= h_{m'-m}(x_1, \dots, x_m), \end{aligned}$$

où $h = (h_1, \dots, h_{m'-m})$ est une application différentiable.

Pour que l'expression de f prenne la forme voulue il ne reste plus qu'à « redresser » le graphe de h par le changement de coordonnées évident :

$$\begin{aligned} y_1 &:= y_1, \dots, y_m := y_m, y_{m+1} := y_{m+1} - h_1(x_1, \dots, x_m), \dots \\ y_{m'} &:= y_{m'} - h_{m'-m}(x_1, \dots, x_m) \end{aligned}$$

□

5. Transversalité

La notion de « transversalité » fournit une interprétation géométrique, en termes d'espaces tangents, des diverses variantes et conséquences du théorème des fonctions implicites.

5.0. Transversalité de deux sous-espaces vectoriels

Soit E un espace vectoriel de dimension finie. Rappelons que, pour tout couple de sous-espaces vectoriels F et F' de E , on a :

$$\dim(F + F') = \dim F + \dim F' - \dim F \cap F'$$

DÉFINITION F et F' sont dits **transverses** dans E (notation abrégée : $F \pitchfork F'$) si $F + F' = E$.

D'après l'égalité rappelée ci-dessus, cette propriété équivaut à la suivante :

$$\text{codim}_E F \cap F' = \text{codim}_E F + \text{codim}_E F'$$

Vérification pratique de la transversalité

Si F [resp. F'] est donné dans E par un système d'équations linéaires en nombre égal à sa codimension (donc linéairement indépendantes) :

$$\begin{array}{l} F : \ell_1 = \dots = \ell_q = 0 \quad (q = \text{codim} F) \\ F' : \ell'_1 = \dots = \ell'_{q'} = 0 \quad (q' = \text{codim} F') \end{array}$$

la transversalité de F et F' équivaut à :

- i) l'indépendance linéaire de $\ell_1, \dots, \ell_q, \ell'_1, \dots, \ell'_{q'}$; ou à
- ii) l'indépendance linéaire de $\ell_1 | F', \dots, \ell_q | F'$.

5.1. Transversalité de deux sous-variétés de \mathbb{R}^n

Deux sous-variétés S, S' de \mathbb{R}^n sont dites *transverses en* $a \in S \cap S'$ (notation abrégée : $S \pitchfork_{\mathbb{R}^n, a} S'$) si $T_a S$ et $T_a S'$ sont transverses dans $T_a \mathbb{R}^n$.

Elles sont dites *transverses* (en abrégé : $S \pitchfork_{\mathbb{R}^n} S'$) si elles sont transverses en tout point $a \in S \cap S'$.

Remarque Si $\text{codim} S + \text{codim} S' > n$ (exemple : deux courbes de \mathbb{R}^3),

$$S \pitchfork_{\mathbb{R}^n} S' \iff S \cap S' = \emptyset.$$

PROPOSITION

L'intersection de deux sous-variétés transverses est une sous-variété, dont la codimension est la somme des codimensions ; plus précisément :

en un point a où S et S' sont lisses et transverses, $S \cap S'$ est lisse, et

$$T_a(S \cap S') = T_a S \cap T_a S'.$$

Preuve

Définir S et S' par des équations locales, et appliquer le lemme ci-après (critère i). \square

Lemme (vérification pratique de la transversalité) Si S [resp. S'] est donnée dans \mathbb{R}^n au voisinage de a par un système d'équations locales s_1, \dots, s_q [resp. s'_1, \dots, s'_q], la transversalité de S et S' en a équivaut à :

- i) l'indépendance linéaire de $ds_1, \dots, ds_q, ds'_1, \dots, ds'_q$, en a ; ou à
- ii) l'indépendance linéaire de $ds_1|_{T_a S}, \dots, ds_q|_{T_a S}$.

5.2. Transversalité d'une application à une sous-variété

Soit $g : U \rightarrow \mathbb{R}^n$ une application différentiable définie sur un ouvert U de \mathbb{R}^m , et soit S une sous-variété de \mathbb{R}^n . On dit que g est *transverse* à S en a (en abrégé : $g \pitchfork_a S$) si l'image de Dg_a est transverse en $b = g(a)$ à $T_b S$.

On dit que g est *transverse* à S si elle l'est en tout point $a \in U$.

Remarque Si g est une immersion locale en a , $g \pitchfork_a S \iff \text{Im } g_a \pitchfork_{\mathbb{R}^n, b} S$, où g_a désigne la restriction de g à un voisinage de a choisi assez petit pour que l'image de g_a soit une sous-variété (cf. § 2.2).

Dans le cas général on a la proposition suivante.

PROPOSITION

g est transverse [en a] à S si et seulement si son graphe $(g) \subset U \times \mathbb{R}^n$ est transverse [en $a \times b$] à $U \times S$.

Preuve

En posant $u = Dg_a$ on est ramené à démontrer le lemme d'algèbre linéaire ci-après. \square

Lemme Soit $u : E_1 \rightarrow E_2$ une application linéaire d'espaces vectoriels de dimensions finies, et soit F un sous-espace vectoriel de E_2 . Alors :

$$\text{Im } u \pitchfork_{E_1} F \iff (u) \pitchfork_{E_1 \times E_2} E_1 \times F$$

où (u) désigne le graphe de u .

Preuve du lemme

Si F est défini dans E_2 par l'annulation du système de formes linéaires $\ell = (\ell_1, \dots, \ell_q)$ ($q = \text{codim} F$), $E_1 \times F$ est défini dans $E_1 \times E_2$ par l'annulation du système $\ell \circ pr_2$, où $pr_2 : E_1 \times E_2 \rightarrow E_2$ est la deuxième projection. En appliquant le critère de transversalité du paragraphe 4.0 ii), on a donc d'une part :

$$\text{Im } u \pitchfork_{E_2} F \iff \ell | \text{Im } u \text{ est surjectif}$$

et d'autre part :

$$(u) \pitchfork_{E_1 \times E_2} E_1 \times F \iff \ell \circ pr_2 | (u) \text{ surjectif}$$

Ces deux conditions sont manifestement équivalentes puisque $\text{Im } u = \text{pr}_2((u))$. \square

Corollaire $g \nabla S \Rightarrow g^{-1}(S)$ est une sous-variété de U , de codimension égale à celle de S .

En effet, $g^{-1}(S)$ est définie au voisinage de $a \in U$ par le système d'équations locales $s \circ g$, où $s = (s_1, \dots, s_q)$ est un système d'équations locales de S en $b = g(a)$; que $s \circ g$ soit une submersion résulte de la preuve du lemme ci-dessus, avec $\ell = ds$.

EXERCICES

EXERCICE 1.— « Parapluie de Whitney », ou « surface cuspidale ».

- a) C'est la surface $x^2 - y^2z = 0$ de \mathbb{R}^3 . La dessiner avec soin (sections par $z = \text{Cte}$, ou $y = \text{Cte}$, ou $z = ay\dots$). Montrer qu'en dehors de l'axe des z elle est lisse, de codimension 1. Y a-t-il des points de cet axe où elle est visiblement lisse ? Montrer qu'au voisinage d'un point $(0, 0, z_0)$, $z_0 > 0$, un difféomorphisme local convenable la transforme en deux plans : $\tilde{x}^2 - \tilde{y}^2 = 0$.
- b) En quels points l'application $(u, v) \mapsto (uv, v, u^2)$ est-elle une immersion de \mathbb{R}^2 dans \mathbb{R}^3 ? Déterminer son image, et comparer avec a).

EXERCICE 2.— En utilisant comme ingrédient de base le paramétrage usuel du cercle trigonométrique par l'angle, construire une application simple de \mathbb{R}^2 dans \mathbb{R}^3 ayant pour image :

- a) un cylindre de révolution : l'application est-elle une immersion ? Est-elle injective ? Son image est-elle lisse ?
- b) une surface de révolution obtenue en faisant tourner un cercle autour d'un axe coplanaire au cercle.
- i) Dessiner cette surface, en distinguant différents cas selon la position relative de l'axe et du cercle. Dans chacun des cas, dire en quels points l'application est une immersion.
- ii) Cette surface est-elle une sous-variété ? Chercher d'abord à deviner la réponse sur le dessin. Aux points où vous pensez que la surface est lisse, le démontrer par une des méthodes du cours ; aux points où vous pensez qu'elle ne l'est pas, débrouillez-vous...

EXERCICE 3.—

- a) Le paramétrage du cercle par l'angle :

$$\theta \mapsto (x = \cos \theta, y = \sin \theta)$$

permet de repérer la position d'un point du cercle par une « coordonnée angulaire ». Préciser ce dont il s'agit, de façon à obtenir des cartes du cercle. Existe-t-il une carte ayant comme domaine tout le cercle ?

- b) Définir de façon analogue des cartes du cylindre (cf. exercice 2)a), puis du tore (surface de l'exercice 2)b), dans le cas où l'axe de révolution ne rencontre pas le cercle).

- c) Pour repérer une position sur la sphère terrestre, les géographes utilisent les coordonnées « latitude et longitude ». Expliquer comment elles sont définies, et comment on obtient ainsi des cartes de la sphère (on pourra encore se référer à l'exercice 2)b), cas où l'axe de révolution est un diamètre du cercle).

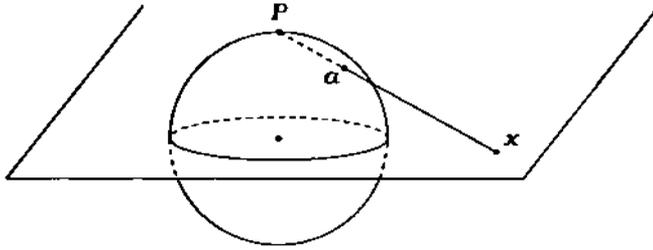


Figure 6.

- d) Montrer que la projection stéréographique (cf. Fig. 6) définit une carte de la sphère S^2 dont on précisera le domaine.

Thème d'étude suggéré Se documenter sur les différents systèmes de représentation de la sphère terrestre utilisés par les géographes (projection de Mercator, de Lambert, etc.) et montrer qu'il s'agit bien de cartes au sens du paragraphe 3.0.

RÉFÉRENCE: Marcel Berger, *Géométrie*, (Cedic-Nathan, 1977).

EXERCICE 4.— Une application dont le graphe est lisse est-elle toujours différentiable ? Réfléchissez au contre-exemple de la fig. 7, et proposez un énoncé modifié faisant intervenir la position de l'espace tangent.

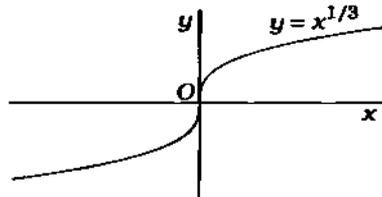


Figure 7.

EXERCICE 5.— Soit $f : (\mathbb{R}^n, 0) \cdots \rightarrow (\mathbb{R}^n, 0)$ une application différentiable. À quelle condition sur l'application linéaire tangente Df_0 le graphe de f sera-t-il transverse en 0 au graphe de l'application identique ?

EXERCICE 6.— Soit $F : (\mathbb{R}^n, a) \cdots \rightarrow \mathbb{R}^p$ une submersion, et soit $S \subset \mathbb{R}^n$ un sous-ensemble lisse en a .

Montrer que $f = F|_S$ est une submersion [resp. une immersion] si et seulement si $\text{Ker } DF_a$ est transverse à $T_a S$ [resp. $\text{Ker } DF_a \cap T_a S = 0$]. (N.B. On trouvera des exemples de ces situations dans les figures de l'étude 2.)

EXERCICE 7.— Applications idempotentes et rétractions différentiables. Soit U un ouvert de \mathbb{R}^n , et soit $u : U \rightarrow U$ une application différentiable *idempotente*, c'est-à-dire telle que $u \circ u = u$.

On se propose de démontrer qu'alors u se factorise en :

$$u : U \xrightarrow{r} S \xrightarrow{\iota_S} U$$

où S est une sous-variété fermée de U , ι_S est l'inclusion canonique, et r est une rétraction, c'est-à-dire une application telle que $r|_S = \mathbb{1}_S$.

o) Montrer que le résultat annoncé se ramène à l'affirmation suivante : l'ensemble des points fixes de u est une sous-variété $S \subset U$. Tout le problème est donc local au voisinage d'un point fixe $a \in U$.

i) Énoncer et démontrer le résultat analogue pour un endomorphisme linéaire d'un espace vectoriel.

ii) Que peut-on dire quand Du_a est de rang maximal (n) ?

iii) On suppose que $\text{rg } Du_a = m < n$, et l'on pose $p = n - m$. On se ramène (com-

ment ?) au cas où $Du_a = \mathbb{I} \times 0 : \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R}^n$, et l'on pose $u = f \times g$, c'est-à-dire $u(x, y) = (f(x, y), g(x, y)) \in \mathbb{R}^m \times \mathbb{R}^p$ ($x = (x_1, \dots, x_m), y = (y_1, \dots, y_p)$). Montrer que le système d'équations $y = g(x, y)$ définit localement une sous-variété Σ de dimension m , tangente en a à $\{y = 0\}$.

iv) En se ramenant (comment ?) au cas où Σ est définie par les équations $y = 0$, montrer que $u(\Sigma) \subset \Sigma$, et en déduire le résultat annoncé en appliquant ii) à l'application idempotente ν_Σ ainsi définie de Σ dans elle-même.

ÉTUDE 2

Contours apparents des surfaces

Pour dessiner sur une feuille de papier une courbe ou une surface S de l'espace à trois dimensions, il faut choisir une direction de projection (de l'espace sur le plan de la feuille). Si S est une courbe, la restriction à S de la projection est « en général » une immersion, sauf en certains points (en général isolés). Si S est une surface, la restriction à S de la projection est en général un difféomorphisme local, sauf aux points dits **critiques**²¹ où l'espace tangent à S ne se projette pas isomorphiquement sur la feuille de papier ; ces points forment en général une courbe de S , la **courbe critique**, dont l'image sur la feuille de papier est le **contour apparent** de S .

Les schémas de la figure 8 montrent les contours apparents de quelques surfaces simples, pour diverses directions de projection. On remarquera, sur le schéma du tore « vu de biais », que le contour apparent ne coïncide pas exactement avec le bord de la projection de S , mais « déborde un peu » à l'intérieur de cette projection : vous essaieriez d'expliquer pourquoi, et de définir ce qui distingue les parties en traits pleins des parties en tirets du contour apparent.

De façon générale, un **point critique** d'une application différentiable est un point de l'espace source où l'application tangente n'est pas surjective. Dans le cas particulier des fonctions numériques on retrouve la notion introduite dans l'étude 1.

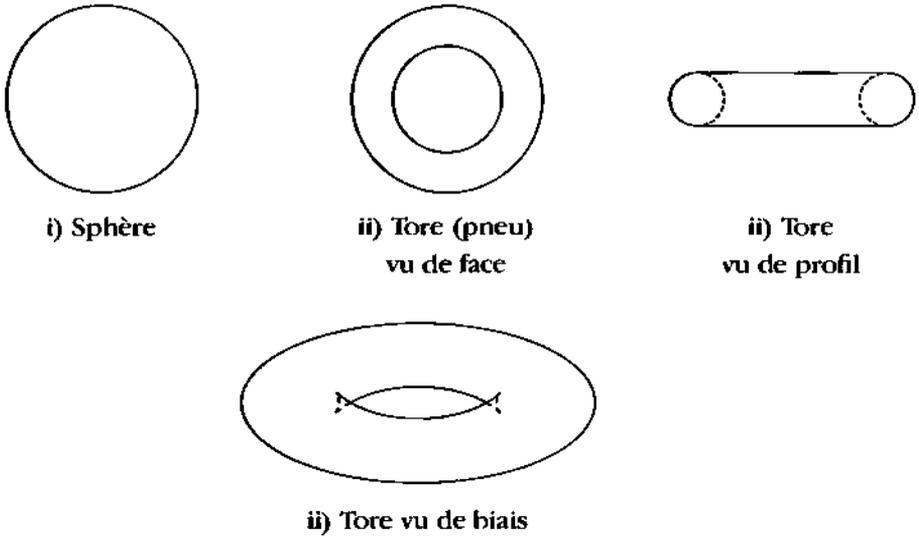


Figure 8.

Une variante intéressante, tout aussi utile pratiquement, est le problème de la représentation graphique de la paire constituée par une surface et une courbe de cette surface. Les figures 9 et 10 en montrent quelques exemples sur lesquels vous pouvez réfléchir.

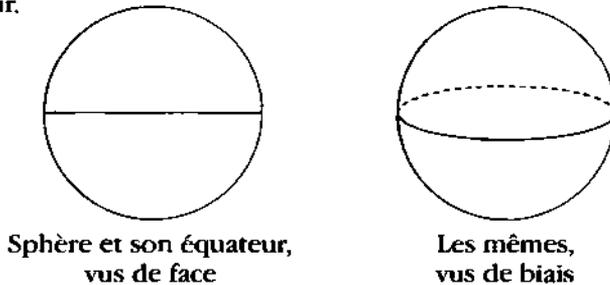


Figure 9.

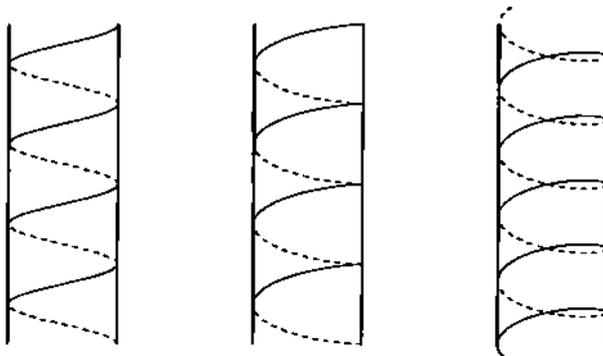


Figure 10. Cylindre avec spirale (ressort d'amortisseur de moto), pour différentes directions de projection ; pouvez-vous caractériser le choix de direction qui a conduit au schéma du milieu ?

EXERCICES SUR LE THÈME D'ÉTUDE 2

EXERCICE 1.— La « fronce ». On considère dans \mathbb{R}^3 la surface :

$$S = \{(x, p, q) \in \mathbb{R}^3 \mid x^3 + px + q = 0\}$$

- a) Montrer que S est une sous-variété, et déterminer son plan tangent en 0. Le couple (p, q) est-il un système de coordonnées locales de S au voisinage de 0 ? Et le couple (x, p) ? Et (x, q) ?
- b) Calculer l'ensemble critique C de l'application :

$$(p \mid S, q \mid S) : S \rightarrow \mathbb{R}^2$$

Montrer que C est une courbe lisse passant par l'origine, et déterminer sa tangente à l'origine.

La variable x est-elle une coordonnée locale de C au voisinage de 0 ? Et p ? Et q ?

- c) Dessiner le lieu des zéros, dans \mathbb{R}^2 , du « discriminant » $\Delta(p, q) = 4p^3 + 27q^2$ (ensemble des valeurs de (p, q) pour lesquels le polynôme du troisième degré a une racine multiple). Quel lien y a-t-il entre ce lieu et la courbe C ?

EXERCICE 2.— On se propose de calculer, pour diverses directions de projection, l'ensemble critique de la projection d'un tore sur un plan. Les coordonnées (x, y, z) de \mathbb{R}^3 sont choisies de façon que l'axe des z soit l'axe de révolution du tore. Pour $\varepsilon \in \mathbb{R}$, on note π_ε la restriction au tore de la projection $(x, y, z) \mapsto (x, z - \varepsilon y)$.

- a) En utilisant la représentation paramétrique du tore donné par l'exercice 2 de ce chapitre, déterminer l'ensemble critique C_ε de π_ε .
Vérifier que pour $\varepsilon = 0$ c'est une union de quatre cercles, et démontrer que pour $\varepsilon \neq 0$ c'est une courbe lisse qui a deux composantes connexes.

Confronter les résultats du calcul avec les figures du début du thème d'étude.

- b) En définissant le tore par une équation $s(x, y, z) = 0$ (à déterminer), essayer de retrouver les résultats précédents (le cas $\varepsilon \neq 0$ est réservé aux courageux !).

EXERCICE 3.— Soit S une surface de révolution de \mathbb{R}^3 , d'équation $x^2 + y^2 = g(z)$ (l'axe Oz étant l'axe « vertical »).

- a) Donner des hypothèses simples et générales sur la fonction g , garantissant que S soit une sous-variété. Calculer, sous ces hypothèses, l'ensemble critique C_0 de l'application :

$$\begin{aligned} \pi_0 : S &\rightarrow \mathbb{R}^2 \\ (x, y, z) &\mapsto (y, z) \end{aligned}$$

(projection orthogonale de S sur un plan vertical).

Montrer que C_0 est une courbe lisse.

- b) Préciser les hypothèses sur g pour que la surface lisse S soit compacte et connexe (cf. par exemple Fig. 11).

Sous ces hypothèses, on déforme la projection précédente en prenant comme direction de projection la direction $(1, 0, \varepsilon)$ (au lieu de $(1, 0, 0)$). Montrer que pour

tout ε assez petit l'ensemble critique C_ε de cette projection π_ε reste une courbe lisse, voisine de la précédente.

- c) La figure 11 représente la surface d'un œuf, éclairée par une source lumineuse située à l'infini dans une direction presque horizontale (parallèle au plan de la feuille). Expliquer comment les considérations précédentes permettent de traiter cette situation (facultatif : proposer un exemple explicite de fonction g telle que S ressemble *vraiment* à l'œuf de la figure).

Indiquer comment on calcule la courbe C séparant la zone éclairée de la zone d'ombre, et donner une construction géométrique simple permettant de « placer » les points A, A', B sur le dessin.

Considérant la figure comme une photographie « de profil » de l'œuf éclairé, dessiner à côté d'elle sa photographie « de face », c'est-à-dire après une rotation de $\frac{\pi}{2}$ de l'appareil de prise de vue autour de l'axe de l'œuf.

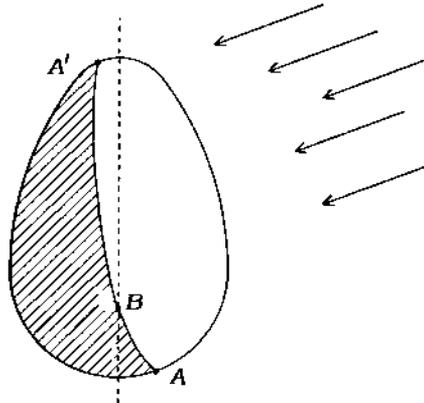


Figure 11.

Champs de vecteurs

Un **champ de vecteurs** v (de classe C^∞) sur un ouvert U de \mathbb{R}^n est une correspondance qui à tout $a \in U$ associe un vecteur¹ $v(a) \in T_a\mathbb{R}^n$ dont les composantes² $v_1(a), \dots, v_n(a)$ sont fonctions C^∞ de a .

Tous les champs de vecteurs seront désormais supposés C^∞ .

1. Intégration d'un champ de vecteurs et dérivation le long d'un champ de vecteurs

1.1. Courbes intégrales

Une **courbe intégrale** du champ de vecteurs v est une application différentiable $\lambda : I \rightarrow U$, définie sur un intervalle ouvert I de \mathbb{R} , telle que :

$$\forall t \in I, \quad \frac{d}{dt} \lambda(t) = v(\lambda(t))$$

Trouver une courbe intégrale d'un champ de vecteurs v , c'est donc résoudre un système de n équations différentielles à n inconnues $\lambda_1, \dots, \lambda_n$:

$$\begin{cases} \frac{d}{dt} \lambda_1(t) &= v_1(\lambda_1(t), \dots, \lambda_n(t)) \\ &\dots \\ \frac{d}{dt} \lambda_n(t) &= v_n(\lambda_1(t), \dots, \lambda_n(t)) \end{cases} \quad (1)$$

D'après le théorème d'existence locale et d'unicité des solutions d'équations différentielles, la donnée d'une « condition initiale » $\lambda(0) = a_0$ détermine une telle solution

¹ Nous omettons désormais les flèches sur les vecteurs tangents, en écrivant v au lieu de \vec{v} .

² Dans \mathbb{R}^n , il est courant d'appeler « composantes » d'un vecteur ses coordonnées dans la base canonique. De façon générale, la notion de « composantes » d'un vecteur est relative à une *décomposition* (de l'espace vectoriel considéré) en somme directe (de sous-espaces vectoriels) : en l'occurrence $\mathbb{R}^n = \mathbb{R} \oplus \dots \oplus \mathbb{R}$. La notion de « coordonnées » (d'un point ou d'un vecteur) est relative à un repère ou à une base (de l'espace affine ou vectoriel considéré).

sans ambiguïté (localement, c'est-à-dire pour t assez voisin de 0). De plus cette solution dépend différentiablement de la condition initiale.

Autrement dit, l'intégration du système différentiel (1) définit, localement au voisinage de $t = 0$, $a = a_0 \in U$, une application différentiable :

$$\begin{aligned} \phi : (\mathbb{R} \times U, 0 \times a_0) \cdots &\rightarrow U \\ t, a &\mapsto \phi(t, a) = \lambda_a(t). \end{aligned}$$

Cette application est appelée **flot local** du champ de vecteurs v en a_0 . Intuitivement, si l'on pense à v comme au champ de vitesses d'un fluide remplissant U , $\phi(t, a)$ nous donne la position au bout d'un temps t d'une particule de fluide initialement observée à la position a .

Trajectoires intégrales et portrait de phase Pour représenter graphiquement l'allure des courbes intégrales d'un champ de vecteurs (par exemple dans \mathbb{R}^2), on en dessine quelques « **trajectoires intégrales** » (images dans \mathbb{R}^2 des courbes intégrales), en indiquant par une flèche le sens de parcours de chaque trajectoire (cf. par exemple Fig. 5 de l'intermède ci-après). La collection de toutes les trajectoires intégrales ainsi orientées s'appelle le *portrait de phase* du champ de vecteurs.

1.2. Intégrales premières

On appelle **intégrale première** (ou **constante du mouvement**) d'un champ de vecteurs une fonction f sur le domaine de définition U de ce champ, qui reste constante le long de toute courbe intégrale.

PROPOSITION

Pour que f soit une intégrale première du champ de vecteurs v , il faut et il suffit que sa dérivée directionnelle le long de v soit identiquement nulle :

$$df_x(v(x)) = 0 \text{ pour tout } x \in U$$

Preuve

Il suffit d'appliquer à $f(\lambda(t))$, où λ est une courbe intégrale, le résultat selon lequel une fonction d'une variable t est constante sur un intervalle si et seulement si sa dérivée est identiquement nulle. \square

Exemple Les équations de la mécanique de Newton pour un point matériel (disons, de \mathbb{R}^3) dans un potentiel V peuvent s'écrire dans $\mathbb{R}^3 \times \mathbb{R}^3$ comme un système différentiel du premier ordre :

$$\begin{cases} \frac{dq}{dt} = p \\ \frac{dp}{dt} = -\text{grad } V(q) := -(\partial_{q_1} V, \partial_{q_2} V, \partial_{q_3} V) \end{cases}$$

Les solutions $t \mapsto (p(t), q(t)) \in \mathbb{R}^3 \times \mathbb{R}^3$ sont donc les courbes intégrales du champ de vecteurs v de $\mathbb{R}^3 \times \mathbb{R}^3$ dont les composantes respectives selon l'espace des p et l'espace des q s'écrivent :

$$\begin{cases} v_p &= -\text{grad } V(q) \\ v_q &= p \end{cases}$$

On vérifiera que la « fonction de Hamilton » $H = \frac{p^2}{2} + V(q)$ en est une intégrale première : c'est le théorème de conservation de l'énergie, la valeur constante de H s'interprétant comme l'énergie du point matériel (on a normalisé à 1 la masse, de sorte que $\frac{p^2}{2}$ est l'énergie cinétique).

Plus généralement, étant donnée une fonction H de $2n$ variables $p \in \mathbb{R}^n$, $q \in \mathbb{R}^n$, on appelle **système hamiltonien, de fonction de Hamilton** H , le système différentiel du premier ordre :

$$\begin{cases} \frac{dq}{dt} &= \partial_p H(p, q) \\ \frac{dp}{dt} &= -\partial_q H(p, q) \end{cases} \quad (\text{« équations de Hamilton »})$$

On vérifie immédiatement que la fonction H est une intégrale première de ce système, c'est-à-dire une intégrale première du champ de vecteurs ($v_p = -\partial_q H$, $v_q = \partial_p H$).

1.3. Les champs de vecteurs vus comme dérivations

La dérivée directionnelle d'une fonction f le long d'un vecteur $v(x) \in T_x \mathbb{R}^n$ est un nombre. En faisant varier le point x dans U (domaine de définition du champ de vecteurs v) on obtient une fonction, que l'on notera pour abrégier vf , et qu'on appellera **dérivée directionnelle de f le long du champ de vecteurs v** :

$$vf(x) = df_x(v(x))$$

Si f et v sont de classe C^∞ sur U , il en sera de même de vf . La correspondance $f \mapsto vf$ est ainsi un endomorphisme linéaire :

$$v : C^\infty(U) \rightarrow C^\infty(U)$$

de l'algèbre des fonctions C^∞ sur U . On vérifie immédiatement que cet endomorphisme est une *dérivation* de l'algèbre $C^\infty(U)$, c'est-à-dire qu'il a les mêmes propriétés algébriques que la dérivation usuelle

$$\begin{cases} v(fg) &= (vf)g + f(vg), \\ va &= 0 \text{ si } a \text{ est une fonction constante.} \end{cases}$$

Remarquons que le champ de vecteurs v est déterminé sans ambiguïté par la dérivation qu'il définit : en effet, sa i -ème composante v_i n'est autre que la fonction $v x_i$, obtenue en faisant agir la dérivation v sur la i -ème fonction coordonnée canonique

de \mathbb{R}^n . Il n'y a donc aucun inconvénient à désigner par la même lettre v un champ de vecteurs et la dérivation qu'il définit.

Exemple Le champ de vecteurs constant, de composantes $(0, \dots, 1, 0, \dots, 0)$ (avec 1 en i -ème position) peut être identifié à la dérivation ∂_i (i -ème dérivée partielle) ; c'est donc par la notation ∂_i qu'il est en général désigné (cette notation a l'avantage d'être canonique, alors que les notations pour le i -ème vecteur de base de \mathbb{R}^n varient d'un ouvrage à l'autre !).

Remarque Avec les notations ci-dessus, la propriété pour une fonction f d'être une intégrale première d'un champ de vecteurs v s'écrit simplement $vf = 0$ (Prop. 1.2). Ainsi, par exemple, dans \mathbb{R}^2 la relation évidente $(x_1 \partial_2 - x_2 \partial_1)(x_1^2 + x_2^2) = 0$ peut s'interpréter en disant que $x_1^2 + x_2^2$ est une intégrale première du champ de vecteurs $x_1 \partial_2 - x_2 \partial_1$ (dessiner ce champ !).

EXERCICE-TEST En s'inspirant de la remarque précédente, construire sur $\mathbb{R}^{2n} = \mathbb{R}^2 \times \dots \times \mathbb{R}^2$ un champ de vecteurs ne s'annulant qu'à l'origine et ayant $\sum_{i=1}^n x_i^2$

comme intégrale première.

Interprétation physique En notant $(p_1, q_1, \dots, p_n, q_n)$ les coordonnées canoniques dans \mathbb{R}^{2n} , la fonction :

$$H = \sum_{i=1}^n p_i^2 + q_i^2$$

est la fonction de Hamilton de l'oscillateur harmonique à n degrés de liberté (point matériel dans le « puits de potentiel » $V(q) = \sum_{i=1}^n q_i^2$). Une solution de l'exercice est donc donnée par le champ de vecteurs de Hamilton.

1.4. Champs de vecteurs sur une sous-variété

DÉFINITION Un **champ de vecteurs** (de classe C^∞) sur une sous-variété $S \subset \mathbb{R}^n$ est une correspondance qui à tout $a \in S$ associe un vecteur tangent à S en a :

$$v(a) \in T_a S$$

dont les composantes $v_1(a), \dots, v_n(a)$ dans l'espace ambiant \mathbb{R}^n sont des fonctions C^∞ de a sur S .

Exemple type Soit v un champ de vecteurs C^∞ sur un ouvert U de \mathbb{R}^n contenant S . Supposons-le tangent à S , c'est-à-dire que $v(a) \in T_a S$ pour tout $a \in S$. Alors la restriction de v à S définit un champ de vecteurs sur S .

Réciproquement, il résulte immédiatement du « lemme de restriction » du chap. 1 (§ 2.5) que tout champ de vecteurs sur S peut, au moins localement (au voisinage d'un point a donné), être considéré comme la restriction d'un champ de vecteurs de l'espace ambiant. D'après le § 3.3 du chap. 2, c'est même la restriction d'un champ de vecteurs défini sur tout un voisinage de S , de sorte que notre « exemple type » est en fait le cas général !

« EXERCICE » (évident) Dédurre de l'exercice du paragraphe 1.3 le résultat suivant :

sur toute sphère de dimension impaire il existe un champ de vecteurs C^∞ (et même analytique) ne s'annulant nulle part.

Tout ce que nous avons dit précédemment des champs de vecteurs de \mathbb{R}^n s'étend sans modification aux champs de vecteurs sur les sous-variétés : notions de *courbe intégrale* et de *trajectoire intégrale* ; notion d'*intégrale première* ; interprétation des champs de vecteurs C^∞ sur S comme *dérivations* $v : C^\infty(S) \rightarrow C^\infty(S)$, définies par la même formule qu'en 1.3 :

$$vf(x) = df_x(v(x))$$

formule dont le second membre s'interprète maintenant au sens du calcul différentiel sur S (chap. 1 § 2.5) ; là encore le champ de vecteurs v est déterminé sans ambiguïté par la dérivation de $C^\infty(S)$ qu'il définit : les composantes v_1, \dots, v_n de ce champ de vecteurs dans l'espace ambiant \mathbb{R}^n s'obtiennent là encore par la formule $v_i = vx_i$, où cette fois la fonction x_i sur laquelle on fait agir la dérivation v est la restriction à S de x_i , i -ème fonction coordonnée canonique de \mathbb{R}^n .

2. Transformation des champs de vecteurs par difféomorphismes

2.1. Image d'un champ de vecteurs par un difféomorphisme

La donnée d'un difféomorphisme entre ouverts de \mathbb{R}^n (ou entre sous-variétés) :

$$U \xrightarrow{\phi} \tilde{U}$$

met en bijection divers objets définis sur U et les objets correspondants sur \tilde{U} :

i) *points mobiles* (courbes paramétrées)

$$\lambda \leftrightarrow \tilde{\lambda} \quad (\tilde{\lambda} = \phi \circ \lambda)$$

ii) *fonctions* C^∞

$$\varphi \leftrightarrow \tilde{\varphi}$$

($\varphi = \phi^* \tilde{\varphi} := \tilde{\varphi} \circ \phi$; la notation $\phi^* : C^\infty(\tilde{U}) \rightarrow C^\infty(U)$ a été introduite au chap. 2 (fin du § 3.1).

Nous pouvons de même établir une correspondance entre les champs de vecteurs sur U et les champs de vecteurs sur \tilde{U} :

DÉFINITION Soit v un champ de vecteurs sur U . On appelle **image du champ de vecteurs v par le difféomorphisme ϕ** le champ de vecteurs \tilde{v} sur \tilde{U} défini par :

$$\tilde{v}(\tilde{a}) = (T_a \phi)(v(a)), \quad \text{où } a = \phi^{-1}(\tilde{a}) \quad (*)$$

(Fig. 1)

Remarque Cette formule n'aurait aucun sens si l'application ϕ n'était pas un difféomorphisme (on a dû faire intervenir ϕ^{-1}).

PROPOSITION

La correspondance $v \leftrightarrow \tilde{v}$ ainsi définie vérifie les deux propriétés suivantes, dont chacune suffit à la caractériser :

- i) le difféomorphisme ϕ transforme les courbes intégrales de v en les courbes intégrales de \tilde{v} ;
- ii) la dérivation v de $C^\infty(U)$ provient, par l'isomorphisme ϕ^* (correspondance ii)), de la dérivation \tilde{v} de $C^\infty(\tilde{U})$: autrement dit on a un diagramme commutatif :

$$\begin{array}{ccc}
 C^\infty(U) & \xleftarrow{\phi^*} & C^\infty(\tilde{U}) \\
 \downarrow v & & \downarrow \tilde{v} \\
 C^\infty(U) & \xleftarrow{\phi^*} & C^\infty(\tilde{U}).
 \end{array}$$

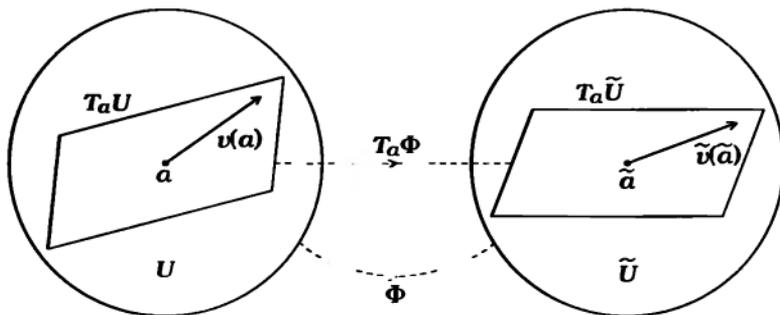


Figure 1.

Preuve

La propriété i) s'obtient par dérivation de l'application composée $\tilde{\lambda} = \phi \circ \lambda$, et la propriété ii) par dérivation de la fonction composée $\varphi = \tilde{\varphi} \circ \phi$. Les détails sont laissés au lecteur. □

2.2. Représentation d'un champ de vecteurs dans une carte

Appliquons la correspondance du paragraphe ci-dessus au cas où $\phi = x$, carte d'une sous-variété S à m dimensions. Notons $U \subset S$ le domaine de la carte, et $\underline{U} \subset \mathbb{R}^m$ son image :

$$U \xrightarrow{x} \underline{U}$$

À chaque champ de vecteurs v de U correspond ainsi un champ de vecteurs \underline{v} de \underline{U} et inversement. Nous appellerons \underline{v} le **représentant de v dans la carte x** .

Nous noterons ∂_{x_i} le champ de vecteurs de \underline{U} représenté dans la carte x par ∂_i (champ de vecteurs constant de \underline{U} , égal au i -ème vecteur de base de \mathbb{R}^m). En tant que dérivation de $C^\infty(\underline{U})$, ∂_{x_i} mérite d'être appelé opérateur de « i -ème dérivation partielle dans le système de coordonnées locales (x_1, \dots, x_m) ».

Attention ! Contrairement à ce que la notation ∂_{x_i} pourrait laisser croire (de même que la notation plus traditionnelle $\frac{\partial}{\partial x_i}$) la dérivation ∂_{x_i} ne dépend pas seulement de la fonction x_i mais du m -uplet de fonctions $x = (x_1, \dots, x_m)$. Autrement dit, notre notation ∂_{x_i} doit être comprise comme « ∂_x indexé par i » et non pas comme « ∂ indexé par x_i » !

EXERCICE TEST 1 En thermodynamique, la pression P , le volume V et la température T d'un gaz sont liés par une relation. Pour désigner les dérivées partielles d'une grandeur thermodynamique f (fonction de P, V, T) les livres de thermodynamique n'utilisent pas la notation $\frac{\partial f}{\partial P}, \frac{\partial f}{\partial V}, \frac{\partial f}{\partial T}$ mais éprouvent le besoin de préciser :

$$\left. \frac{\partial f}{\partial P} \right|_{V=Cte} \cdot \left. \frac{\partial f}{\partial P} \right|_{T=Cte} \cdot \text{etc.}$$

Pouvez-vous expliquer pourquoi, et traduire dans le langage introduit précédemment ?

Explicitiez les relations entre ces diverses dérivées partielles, dans le cas d'un gaz parfait où la relation liant P, V, T est $PV = RT$ ($R = Cte$).

EXERCICE TEST 2 Soit $S \subset \mathbb{R}^m \times \mathbb{R}^p$ le graphe d'une application $f : \mathbb{R}^m \rightarrow \mathbb{R}^p$. On note $x = (x_1, \dots, x_m)$ [resp. $y = (y_1, \dots, y_p)$] la projection canonique $S \rightarrow \mathbb{R}^m$ [resp. $S \rightarrow \mathbb{R}^p$].

a) Calculez les composantes, dans l'espace ambiant $\mathbb{R}^m \times \mathbb{R}^p$, du champ de vecteurs ∂_{x_i} sur S .

b) Pourquoi la même question à propos de « ∂_{y_i} » n'a-t-elle en général aucun sens ? Quelle hypothèse faut-il faire sur f pour qu'elle en ait un ? Explicitiez dans ce cas les relations entre les champs de vecteurs $\partial_{x_1}, \dots, \partial_{x_m}, \partial_{y_1}, \dots, \partial_{y_p}$.

2.3. Redressement local des champs de vecteurs non singuliers

DÉFINITION On appelle **point singulier d'un champ de vecteurs** v un point a où :

$$v(a) = 0$$

Théorème Soit v un champ de vecteurs non singulier au point a_0 . Alors v peut être transformé, au voisinage de a_0 , par un difféomorphisme convenable, en un champ de vecteurs constant.

Preuve

Exploitant l'hypothèse que $v(a_0) \neq 0$, donnons-nous une immersion locale :

$$\begin{array}{ccc} \alpha & : & (\mathbb{R}^{n-1}, 0) \cdots \rightarrow (\mathbb{R}^n, a_0) \\ & & u \quad \mapsto \alpha(u) \end{array}$$

transverse à $v(a_0)$ (c'est-à-dire telle que $v(a_0) \notin \text{Im } D\alpha_0$).

Pour u assez voisin de 0, donc $\alpha(u)$ assez voisin de a_0 , résolvons le système différentiel (1) (§ 1.1) avec $a = \alpha(u)$ comme condition initiale : on obtient ainsi une courbe intégrale $t \mapsto \lambda_{\alpha(u)}(t)$ définie pour t assez voisin de 0, et telle que $\lambda_{\alpha(u)}(t)$ dépende différemment de t et de u (dépendance des solutions de l'équation différentielle par rapport aux conditions initiales).

Autrement dit, nous venons de définir localement une application différentiable :

$$\phi : \begin{matrix} (\mathbb{R} \times \mathbb{R}^{n-1}, 0 \times 0) \\ t, u \end{matrix} \cdots \rightarrow (\mathbb{R}^n, a_0) \mapsto \lambda_{\alpha(u)}(t)$$

□

Lemme 1.— ϕ est un difféomorphisme local.

Preuve

Sa matrice jacobienne en $(0, 0)$ s'écrit :

$$\begin{matrix} & t & u_1 \cdots u_{n-1} \\ \begin{matrix} x_1 \\ \vdots \\ x_n \end{matrix} & \begin{matrix} v_1(a_0) \\ \vdots \\ v_n(a_0) \end{matrix} & \begin{matrix} \\ \\ D\alpha_0 \end{matrix} \end{matrix}$$

qui est bien de rang n d'après l'hypothèse de transversalité.

□

Lemme 2.— ϕ^{-1} transforme v en le champ de vecteurs constants ∂_1 (de composantes $(1, 0, \dots, 0)$).

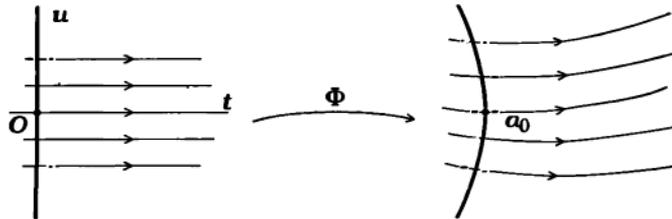


Figure 2. Redressement local d'un champ de vecteurs non singulier.

Preuve

On peut éviter tout calcul en raisonnant sur les courbes intégrales (Fig. 2), comme nous y autorise la proposition du paragraphe 2.1.

□

APPENDICE COMPLÉMENTS SUR LA NOTION DE « FLOT »

Nous avons défini en 1.1 le « flot local » d'un champ de vecteurs v comme l'application :

$$\phi : \begin{matrix} \mathbb{R} \times U, 0 \times a_0 \\ t, a \end{matrix} \cdots \rightarrow U \mapsto \phi(t, a) = \lambda_a(t)$$

définie en intégrant localement le système différentiel (1) avec la condition initiale $\lambda_a(0) = a$.

A1 Dans les bons cas (et notamment si v est à support compact) le flot est défini globalement sur $\mathbb{R} \times U$:

$$\phi : \mathbb{R} \times U \rightarrow U$$

et peut donc être considéré comme une famille, paramétrée différemment par $t \in \mathbb{R}$, d'applications différentiables :

$$\phi^t : U \rightarrow U \quad (\phi^t(a) = \phi(t, a))$$

On déduit immédiatement de la définition de ϕ (et l'idée de l'écoulement d'un fluide est particulièrement suggestive à cet égard) que cette famille (ϕ^t) obéit aux lois de composition :

$$\begin{aligned} \phi^0 &= \mathbb{I}_U \\ \phi^{t'} \circ \phi^t &= \phi^{t+t'} \end{aligned} \quad (2)$$

d'où il résulte que les ϕ^t sont des difféomorphismes de U dans lui-même ($(\phi^t)^{-1} = \phi^{-t}$).

Une telle famille (ϕ^t) de difféomorphismes, indexée sur \mathbb{R} et vérifiant (2), est appelée **groupe à un paramètre de difféomorphismes de U** .

Exemple Tout champ de vecteurs *linéaires* sur \mathbb{R}^n :

$$v(x) = Ax$$

(où A est une matrice $n \times n$) admet globalement sur \mathbb{R}^n un flot défini par :

$$\phi^t(x) = e^{tA}x = \left(1 + \frac{tA}{1!} + \frac{t^2 A^2}{2!} + \dots \right) x$$

EXERCICE (classique !) Comment un champ de vecteurs linéaires de \mathbb{R}^n se transforme-t-il par un difféomorphisme *linéaire* de \mathbb{R}^n ? En utilisant de tels difféomorphismes pour réduire la matrice A à une forme simple, dessiner dans le cas $n = 2$ tous les cas de portraits de phase qui peuvent se présenter (réponse : cf. Fig. 5 de l'intermédiaire ci-après).

A2 Dans les cas où le flot ne peut être défini que *localement*, il faut comprendre que les particules de fluide, ou tout au moins certaines d'entre elles, atteignent les confins de U au bout d'un temps fini (par « confins » nous entendons la frontière dans le cas d'un domaine borné, ou éventuellement l'infini dans le cas général).

L'intervalle de temps sur lequel on peut suivre l'évolution d'une particule dépend alors de la position initiale de celle-ci : pour tout ouvert $U_0 \subset U$ assez petit (et par conséquent pour tout ouvert U_0 relativement compact dans U) il existe un intervalle de temps $T = T(U_0)$ tel que $\phi(t, a)$ soit bien défini pour tout $t \in T(U_0)$, $a \in U_0$.

Au lieu d'un groupe à un paramètre de difféomorphismes de U , on a alors ce qui en est l'analogue local, à savoir une famille, indexée par l'ouvert U_0 et par $t \in T(U_0)$, de difféomorphismes entre ouverts relativement compacts³ de U :

$$\phi_{U_0}^t : U_0 \rightarrow U_t$$

³ Ouverts bornés non adhérents à la frontière de U .

(on a noté U_t l'image de U_0 par le flot au temps t).

Cette famille vérifie la loi de « pseudo-groupe » suivante :

$$\begin{aligned}\phi_{U_0}^0 &= \mathbb{1}_{U_0} \\ \phi_{U_t}^{t'} \circ \phi_{U_0}^t &= \phi_{U_0}^{t+t'} \quad (t, t' \in T(U_0); t' \in T(U_t)).\end{aligned}$$

Commentaire Il ne faut pas s'effrayer de la complexité des notations ; elle ne fait que traduire l'idée suivante, qui seule mérite d'être retenue : *l'intervalle de temps sur lequel on est sûr de pouvoir suivre l'évolution d'une particule dépend de l'ouvert dans lequel on sait qu'elle se trouvait initialement ; c'est pour être sûr que cet intervalle de temps ne soit pas vide qu'on a choisi un ouvert initial « relativement compact » dans U .*

EXERCICE TEST Calculer le flot du champ de vecteurs suivant défini sur \mathbb{R} , et dire si ce flot est global ou non :

- a) $v(x) = \sqrt{1+x^2}$
 b) $v(x) = 1+x^2$.

EXERCICES

EXERCICE 1.— Exemple de champ de vecteurs sur la sphère S^2

La projection stéréographique définit un difféomorphisme entre $S^2 \setminus P$ et \mathbb{R}^2 (cf. chap. 2, exercice 3.– d) ; le point P est le « pôle nord » de la sphère).

À un champ de vecteurs constant de \mathbb{R}^2 (par exemple le champ ∂_1) correspond par ce difféomorphisme un champ de vecteurs v non singulier sur $S^2 \setminus P$.

- a) Déterminer sans calcul la forme de ses trajectoires intégrales.
 b) Montrer par le calcul que v s'étend sur toute la sphère en un champ de vecteurs C^∞ (et même analytique), nul au point P .

On pourra procéder de deux façons, soit en calculant les composantes de v dans l'espace ambiant \mathbb{R}^3 , soit en calculant ses composantes dans une carte de S^2 contenant le point P , par exemple la carte que définit la projection stéréographique à partir du pôle sud.

EXERCICE 2.— Soit v un champ de vecteurs défini au voisinage de l'origine de \mathbb{R}^n , de la forme $v(x) = Ax + \alpha(\|x\|)$, où A est une matrice $n \times n$.

Montrer que le champ de vecteurs transformé de v par un difféomorphisme local $\phi : (\mathbb{R}^n, 0) \rightarrow (\mathbb{R}^n, 0)$ est de la forme $w(y) = By + \alpha(\|y\|)$, où B est une matrice que l'on calculera.

En déduire que pour tout champ de vecteurs v au voisinage de l'origine de \mathbb{R}^n , les valeurs propres de la matrice Dv_0 sont invariantes par tout difféomorphisme local de $(\mathbb{R}^n, 0)$.

EXERCICE 3.— Soit v un champ de vecteur défini au voisinage de l'origine de \mathbb{R} , de la forme $v(x) = \lambda x + \alpha(x)$, avec $\lambda \neq 0$.

Construire un difféomorphisme local $x = \phi(y)$ qui transforme v en le champ linéarisé $w(y) = \lambda y$.

EXERCICE 4.— Montrer qu'en dimension $n = 2$, il n'y a pas de résultat analogue à celui de l'exercice 3) pour le champ :

$$v_\varepsilon(x) = (x_1 \partial_2 - x_2 \partial_1) + \varepsilon (x_1^2 + x_2^2) (x_1 \partial_1 + x_2 \partial_2)$$

[Idée : tracer sur un même dessin les trajectoires intégrales du champ linéarisé v_0 , l'allure du champ v_ε et de ses trajectoires ; étudier le comportement de la fonction $x_1^2 + x_2^2$ le long des trajectoires de v_ε , et conclure.]

ÉTUDE 3 Champs de gradients

On dit que le champ de vecteurs v , défini sur un ouvert connexe U de \mathbb{R}^n , est un **champ de gradient**, et on écrit $v = \text{grad} f$, s'il existe sur U une fonction différentiable f telle que $v_i(x) = \partial_i f(x)$ ($i = 1, \dots, n$). La fonction f est alors unique à l'addition d'une constante près.

EXERCICE 0.— Démontrer que les trajectoires intégrales d'un champ de gradient $\text{grad} f$ sont orthogonales aux hypersurfaces de niveau de la fonction f (là où le gradient ne s'annule pas). Dessiner les unes et les autres dans le cas $f = x_1^2 + \dots + x_n^2$ (par exemple pour $n = 2$ ou 3). Comment le dessin est-il affecté par un difféomorphisme (par exemple linéaire) de \mathbb{R}^n ?

EXERCICE 1.— Une condition nécessaire pour que v soit un champ de gradient est évidemment $\partial_i v_j = \partial_j v_i$ identiquement sur U .

Il existe donc des champs qui ne sont pas de gradient, même localement (par exemple le champ $x_1 \partial_2 - x_2 \partial_1$ sur \mathbb{R}^2 ne vérifie *nulle part* cette condition). Cela ne paraît-il pas contradictoire avec le théorème de redressement local du paragraphe 2.3, au voisinage d'un point où v ne s'annule pas (un champ de vecteurs constant est un champ de gradient !)?

EXERCICE 2.— Démontrer que sur un pavé de \mathbb{R}^n (produit d'intervalles) la condition nécessaire E.1 est aussi suffisante pour que v soit un champ de gradient : supposant par exemple que U est un produit d'intervalles contenant l'origine, on pourra définir f par :

$$f(x_1, \dots, x_n) = \int_0^{x_1} v_1(t, 0, \dots, 0) dt + \int_0^{x_2} v_2(x_1, t, 0, \dots, 0) dt + \dots + \int_0^{x_n} v_n(x_1, \dots, x_{n-1}, t) dt.$$

Interprétation de cette formule (par exemple pour $n = 3$; cf. Fig. 3) :

$$f = \int_0^{A_1} v_1 dt + \int_{A_1}^{A_2} v_2 dt + \int_{A_2}^{A_3} v_3 dt$$

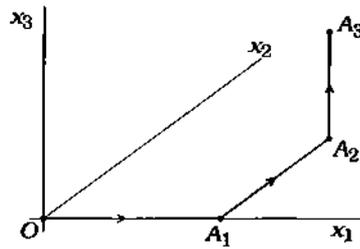


Figure 3.

EXERCICE 3.— Exemple d'ouvert U sur lequel la condition 1) n'est pas suffisante : $U = \mathbb{R}^2 \setminus \{0\}$.

La fonction définie pour $x_1 \neq 0$ par $\theta = \text{Arctan} \frac{x_2}{x_1}$ a son gradient donné par la formule :

$$v = \frac{x_1 \partial_2 - x_2 \partial_1}{x_1^2 + x_2^2},$$

qui définit sur tout $\mathbb{R}^2 \setminus \{0\}$ un champ de vecteurs vérifiant la condition de l'exercice 1. Voyez-vous pourquoi ce champ ne peut pas être globalement de gradient sur $\mathbb{R}^2 \setminus \{0\}$?

ÉTUDE 3bis

Mécanique de Newton et orbites planétaires

1. Exemple préliminaire : oscillateur à une dimension

Considérons un système mécanique à un degré de liberté, dont l'état est repéré par un paramètre q (exemples : élongation d'un ressort par rapport à sa position au repos ; écart angulaire d'un pendule par rapport à la verticale ; etc.). D'après Newton, l'évolution temporelle d'un tel système est régie par une loi de la forme :

$$m \ddot{q} = -V'(q) \quad (0)$$

où V est la fonction « potentiel », et m la « masse » du système (le point signifie la dérivée par rapport au temps, et le double point $\ddot{\cdot}$ la dérivée seconde).

Par exemple, dans le cas du pendule, le potentiel est de la forme $V(q) = -a \cos q$. Dans l'approximation des « petites oscillations », où l'on remplace ce potentiel par son développement limité à l'ordre 2, on a :

$$V(q) = aq^2/2$$

(on a éliminé la constante additive) : c'est ce qu'on appelle l'« oscillateur harmonique ».

L'équation différentielle (0), du deuxième ordre, équivaut à un système différentiel du premier ordre dans l'« espace des phases » dont les coordonnées sont q (« position ») et $p = m\dot{q}$ (« impulsion »). Ce système différentiel du premier ordre peut être écrit sous la forme de Hamilton (chap. 3, § 1.2) :

$$\begin{cases} \dot{q} = \partial_p H \\ \dot{p} = -\partial_q H \end{cases} \quad (1)$$

avec la fonction de Hamilton :

$$H(p, q) = \frac{p^2}{2m} + V(q)$$

On sait que pour tout système hamiltonien (1), la fonction de Hamilton est une constante du mouvement (qui s'interprète ici comme l'**énergie** E du système) ; autrement dit le champ de vecteurs hamiltonien :

$$v_H = \frac{p}{m} \partial_q - V'(q) \partial_p$$

est tangent aux courbes de niveau de H :

$$\frac{p^2}{2m} + V(q) = E. \quad (2)$$

Fixons-nous une valeur E de H , et supposons que la fonction $E - V(q)$ soit strictement positive sur un intervalle $]q_-, q_+[$, négative ou nulle en dehors, avec un zéro simple en q_- et en q_+ . Alors l'équation (2) définit dans le plan des phases une courbe C_E connexe, lisse et compacte (Fig. 4).

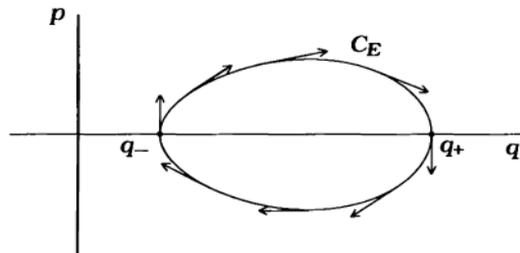


Figure 4. Courbe d'énergie constante et champ de vecteur v_H .

Cette sous-variété compacte connexe à une dimension est munie d'un champ de vecteurs *nulle part nul* (la restriction du champ de vecteurs v_H) dont l'intégration définit un *difféomorphisme local périodique* (paramétrage périodique de C_E) :

$$\lambda : \mathbb{R} \rightarrow C_E \quad (\lambda(t+T) = \lambda(t))$$

EXERCICE Vérifier cette affirmation, en intégrant l'équation différentielle $m\ddot{q} = \pm\sqrt{E - V(q)}$ par séparation des variables ; on montrera ainsi que la période T est donnée par la formule :

$$T = \oint \frac{dq}{\dot{q}} = 2 \int_{q_-}^{q_+} \frac{dq}{\sqrt{\frac{2}{m}(E - V(q))}}$$

2. Loïs de Kepler

Le mouvement des planètes autour du Soleil est décrit à une très bonne approximation par les « **loïs de Kepler** » (Johannes Kepler, 1571-1630) :

- *première loi* : les orbites sont des ellipses ayant le Soleil pour foyer ;
- *deuxième loi (loi des aires)* : l'orbite est parcourue de telle façon qu'un segment joignant le Soleil à la planète balaye des aires égales durant des temps égaux ;
- *troisième loi* : le carré de la période de révolution d'une planète autour du Soleil est proportionnel au cube des dimensions linéaires de l'orbite.

En fait, une description plus fine montre que les orbites, tout en restant planes, ne sont qu'approximativement elliptiques : le grand axe de l'ellipse tourne lentement autour du Soleil, comme indiqué sur la figure 5 (où l'on a exagéré le phénomène pour le rendre plus visible) ; c'est ce qu'on appelle la « précession du périhélie » (le « périhélie » est le point de l'orbite où la planète est le plus près du Soleil).

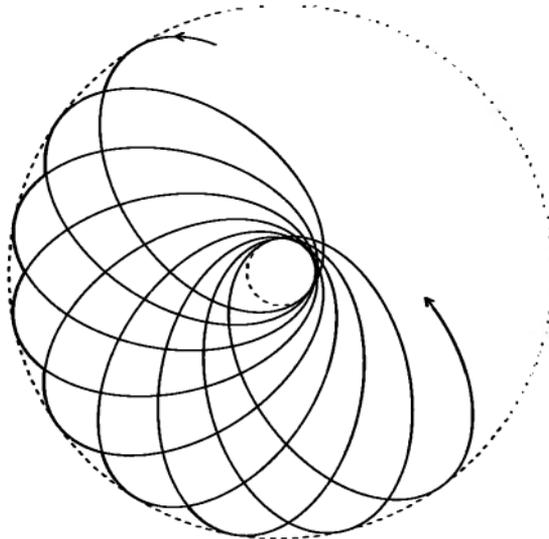


Figure 5. Précession du périhélie d'une orbite planétaire.

2.1. « Enfin Newton vint... »

L'un des grands succès de la théorie de la gravitation universelle de Newton est d'avoir réussi à expliquer les lois de Kepler, à partir de l'hypothèse que le Soleil attire

chaque planète par un champ de forces dérivant d'un potentiel en $-c/r$ (où r est la distance au Soleil, et c une constante positive proportionnelle au produit de la masse du Soleil par celle de la planète ; on néglige les attractions mutuelles des planètes, dont les masses sont infimes par rapport à celle du Soleil).

En fait, (et cette remarque est également due à Newton) la « loi des aires » n'est pas particulière au potentiel en $-c/r$, et est également valable pour tout potentiel « central », c'est-à-dire invariant par rotation autour du Soleil. Par contre, le fait que les orbites soient périodiques (et soient des ellipses !) est typique du potentiel en $-c/r$, les autres potentiels centraux donnant lieu en général à un phénomène de « précession du périhélie » (Fig. 5) : c'est ainsi par exemple que la théorie de la gravitation d'Einstein, en introduisant de petites corrections au potentiel newtonien, a réussi à rendre compte avec précision de la précession observée du périhélie de la planète Mercure.

Nous allons ici étudier les mouvements d'un point matériel soumis au potentiel central $V(r)$, en nous plaçant pour simplifier dans un espace à deux dimensions ; en fait le cas de l'espace à trois dimensions se réduit facilement à celui que nous allons étudier, car on montre que tout mouvement dans un potentiel central à trois dimensions se situe dans un plan contenant l'origine.

2.2. Mouvement dans un potentiel central à deux dimensions

Dans l'espace des phases $\mathbb{R}^2 \times \mathbb{R}^2$, muni des coordonnées canoniques $p = (p_1, p_2)$, $q = (q_1, q_2)$, les équations différentielles du mouvement peuvent s'écrire sous la forme de Hamilton (cf. chap. 3, § 1), avec la fonction de Hamilton :

$$H(p, q) = \frac{p^2}{2m} + V(r)$$

$$\left(p^2 := p_1^2 + p_2^2, r = \|q\| = \sqrt{q_1^2 + q_2^2} \right).$$

En utilisant le fait que V ne dépend que de r , on vérifie que la fonction

$$K(p, q) = q_1 p_2 - q_2 p_1 \quad (\text{« moment angulaire »})$$

est aussi une constante du mouvement.

Remarquons que $K = r p_t$, où p_t est la composante tangentielle de l'impulsion (Fig. 6).

Il en résulte que si l'on repère le point q par ses coordonnées polaires (r, φ) , l'évolution temporelle de φ est liée à celle de r par la loi $m r^2 \dot{\varphi} = K$.

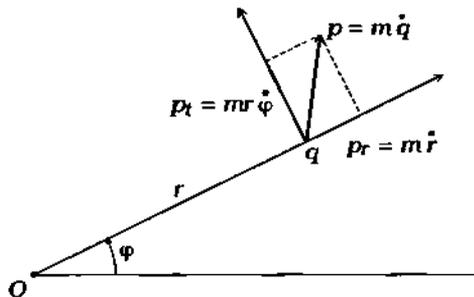


Figure 6. Décomposition de l'impulsion en composantes radiale et tangentielle.

Comme $\frac{1}{2}r^2\dot{\varphi}$ n'est autre que la dérivée par rapport au temps de l'aire balayée par le segment $0q$, on voit ainsi que la propriété pour K d'être une constante du mouvement équivaut à la loi des aires de Kepler.

Pour achever de décrire le mouvement, il reste à décrire l'évolution temporelle de r . En remarquant que $p^2 = p_r^2 + p_l^2 = p_r^2 + \frac{K^2}{r^2}$, on voit que l'évolution du couple (r, p_r) est régie par les équations :

$$\begin{cases} \dot{r} = p_r/m \\ \frac{p_r^2}{2m} + \frac{k^2}{2mr^2} + V(r) = E \end{cases}$$

où k [resp. E] désigne la valeur constante de K [resp. H]. On reconnaît là les équations, dans le plan des phases, d'un oscillateur à une dimension dans un potentiel $V_k(r) = V(r) + \frac{k^2}{2mr^2}$, auquel on peut appliquer l'étude du paragraphe 0.

Dans le cas newtonien $V(r) = -c/r$, le « potentiel effectif » $V_k(r)$ a l'allure représentée sur la figure 7.

Les hypothèses du paragraphe 0 sont vérifiées si l'énergie E est négative (Fig. 7) : dans ce cas r évolue de façon périodique entre un minimum r_- et un maximum r_+ . Ce qui est typique du cas newtonien, c'est que la période de cette oscillation coïncide avec le temps que met l'angle φ à faire un tour : c'est pourquoi les orbites képlériennes sont périodiques.

Par contre, si l'on perturbe si peu que ce soit le potentiel newtonien, le décalage entre la période de r et celle de φ créé un phénomène de « précession », du type représenté sur la figure 5.

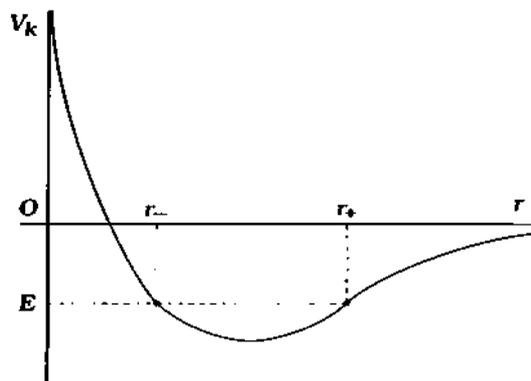


Figure 7. Potentiel effectif dans le cas newtonien.

Lecture recommandée Landau et Lifschitz, *Mécanique* (Physique théorique, tome 1, Éditions Mir, Moscou, 1966), chapitre 3.

PROBLÈME SUR LE THÈME D'ÉTUDE 3^{bis}

- 1) En utilisant les formules du paragraphe 0, écrire sous forme de relation intégrale entre r et φ l'équation des trajectoires pour un potentiel central quelconque. En déduire que, lorsque r parcourt un cycle complet (c'est-à-dire un aller-retour entre r_- et r_+), la variation de l'angle φ est égale à :

$$\Delta\varphi = \oint \omega = 2 \int_{r_-}^{r_+} \omega, \quad \text{où } \omega = \frac{k dr/r^2}{\left[2m(E - V(r)) - \frac{k^2}{r^2}\right]^{\frac{1}{2}}}$$

\oint désigne l'intégration le long d'un cycle complet de r : on remarquera que la détermination de la racine change de signe entre l'aller et le retour.

- 2) (Facultatif : pour les amateurs de calculs explicites.)
Calculer explicitement les relations intégrales de 1) dans le cas $V(r) = -c/r$, et en déduire les lois de Kepler (pour le calcul des intégrales il sera commode de faire le changement de variable $s = 1/r$).

- 3) (Variante pour paresseux.)
La périodicité des trajectoires keplériennes, c'est-à-dire le fait que $\Delta\varphi = 2\pi$ pour le potentiel $V(r) = -c/r$, se démontre sans efforts par la méthode des résidus : l'intégrand ω est méromorphe dans le plan complexe \mathbb{C} fendu le long de $[r_-, r_+]$, avec 0 comme seul pôle ; comme les points r_- et r_+ sont des points de branchement d'ordre 2, au voisinage duquel ω se comporte comme $(r - r_{\pm})^{-\frac{1}{2}}$, l'intégrale $\oint \omega$ peut encore s'écrire $\int_C \omega$, où C est le contour représenté sur la figure 8 ; comme ω est holomorphe à l'infini⁴ on a $\int_C \omega = \int_{\Gamma} \omega$, où Γ désigne le contour représenté sur la figure 8⁵ ; on conclut en remarquant que le résidu de ω en 0 vaut $1/i$.



Figure 8.

⁴ Par le changement de variable $s = 1/r$ l'intégrand devient $\frac{-k ds}{\sqrt{2m(S + cs) - k^2 s^2}}$, qui est holomorphe en $s = 0$.

⁵ En raisonnant sur la sphère de Riemann (cf. Étude 4^{bis}) on voit que les chemins C et Γ se déduisent l'un de l'autre par une déformation évitant les singularités de ω .

TRAVAUX PRATIQUES

À propos de hérissons...

On a vu au paragraphe 1.4 qu'il existe sur toute sphère de dimension impaire un champ de vecteurs ne s'annulant nulle part. Si ce résultat était vrai pour la sphère usuelle S^2 , il serait possible de « peigner » un hérisson roulé en boule, ce qui éviterait de se piquer. Le théorème suivant est donc d'une grande importance pour les hérissons.

Théorème *Il n'existe pas sur la sphère S^2 de champ de vecteurs continu ne s'annulant nulle part.*

Pour les mathématiciens, l'importance de ce théorème ne tient pas tant dans le résultat lui-même que dans les idées qu'il met en jeu. Un peu d'expérimentation sur les champs de vecteurs de S^2 va nous donner un avant-goût de ces idées.

Matériel à se procurer : quelques balles de tennis (ou autres sphères) et de quoi dessiner dessus (crayons feutres) ; on dessinera dessus le « pôle nord », le « pôle sud » et « l'équateur ».

Travaux pratiques

i) Dessinez les trajectoires intégrales du champ de vecteurs v de l'exercice 1. (N.B. : oubliez les calculs de l'exercice 1 ; la réflexion à laquelle vous êtes invité ici est de nature purement géométrique.)

Ce champ de vecteurs est non singulier sur toute la sphère à l'exception d'un seul point (le pôle nord). Le hérisson est ainsi « peigné » partout sauf au pôle nord, et il ne reste plus qu'à « finir le travail » en modifiant notre champ de vecteurs v au voisinage du pôle nord.

ii) Sur une autre balle de tennis (pour ne pas surcharger le dessin), dessinez la restriction à l'équateur du champ de vecteurs v . On se pose le problème de construire sur l'hémisphère nord un champ de vecteurs non singulier coïncidant sur l'équateur avec v .

Le fait que l'hémisphère nord soit difféomorphe à un disque (par exemple par projection stéréographique à partir du pôle sud) permet de se ramener à travailler sur un disque du plan, ce qui est plus commode pour dessiner (la balle de tennis peut être remplacée par une feuille de papier).

iii) Dessinez sur une feuille de papier un champ de vecteurs non singulier sur le disque. Sur une feuille de papier vierge, recopiez sa restriction au cercle (bord du disque) et comparez-la au dessin obtenu dans la question i). Si un champ de vecteurs du plan vous est donné *seulement sur le cercle* (non nul en tout point du cercle), pensez-vous être capable de deviner s'il s'étend ou non en un champ de vecteurs non singulier sur le disque ?

iv) Peut-être notre échec à peigner le hérisson vient-il de ce que notre champ de vecteurs v initial était mal choisi ? Recommencez l'expérience avec d'autres champs de vecteurs non singuliers sur l'hémisphère sud, et essayez de comprendre pourquoi cela ne marche pas non plus.

Indices des champs de vecteurs et caractéristique d'Euler des surfaces

Cet « intermède » apporte des éléments de réponse aux questions soulevées par les « travaux pratiques sur les hérissons ».

Ce sera notre premier contact avec un sujet très riche, la *topologie sur les variétés*, que nous développerons davantage dans la deuxième partie (chap. 7 et fin du chap. 8). L'exposé est fortement inspiré par le livre d'Arnold, *Équations différentielles ordinaires*, dont il développe le paragraphe 36 (chap. V). Il y sera largement fait appel à l'intuition, mais la plupart des arguments avancés seront justifiés rigoureusement dans la deuxième partie.

Les questions qui suivent relèvent davantage de la topologie que du calcul différentiel : sauf à quelques endroits faciles à identifier, les champs de vecteurs n'auront pas besoin d'être différentiables et seront simplement supposés *continus*.

1. Indice d'un champ de vecteurs le long d'un lacet du plan

Soit v un champ de vecteurs non singulier (c'est-à-dire ne s'annulant nulle part) sur un ouvert U de \mathbb{R}^2 , et considérons dans U un point mobile $\lambda(t)$ qui revient à son point de départ au bout d'un temps fini T (que l'on prendra égal à 1 par convention) : on dit alors que l'application (continue) $\lambda : [0, 1] \rightarrow U$ est un *lacet* de U .

Quand le temps t s'écoule de 0 à 1, la direction du vecteur (non nul !) $v(\lambda(t))$ varie continûment ; il en est donc de même de l'angle qu'elle fait avec une direction fixe du plan (angle orienté de demi-droites) ; comme le vecteur retrouve au temps 1 sa position initiale, la variation d'angle de l'instant 0 à l'instant 1 est multiple de 2π , soit $2\pi n$ ($n \in \mathbb{Z}$). L'entier relatif n est appelé **indice du champ de vecteurs v le long de λ** , et nous le noterons $\text{ind}_\lambda(v)$.

1.1. Exemples

La figure 1 donne les indices de quelques champs de vecteurs remarquables, non singuliers dans $U = \mathbb{R}^2 \setminus \{0\}$, le long du lacet λ défini par $\lambda(t) = (\cos 2\pi t, \sin 2\pi t)$ (cercle trigonométrique parcouru une fois dans le sens direct).

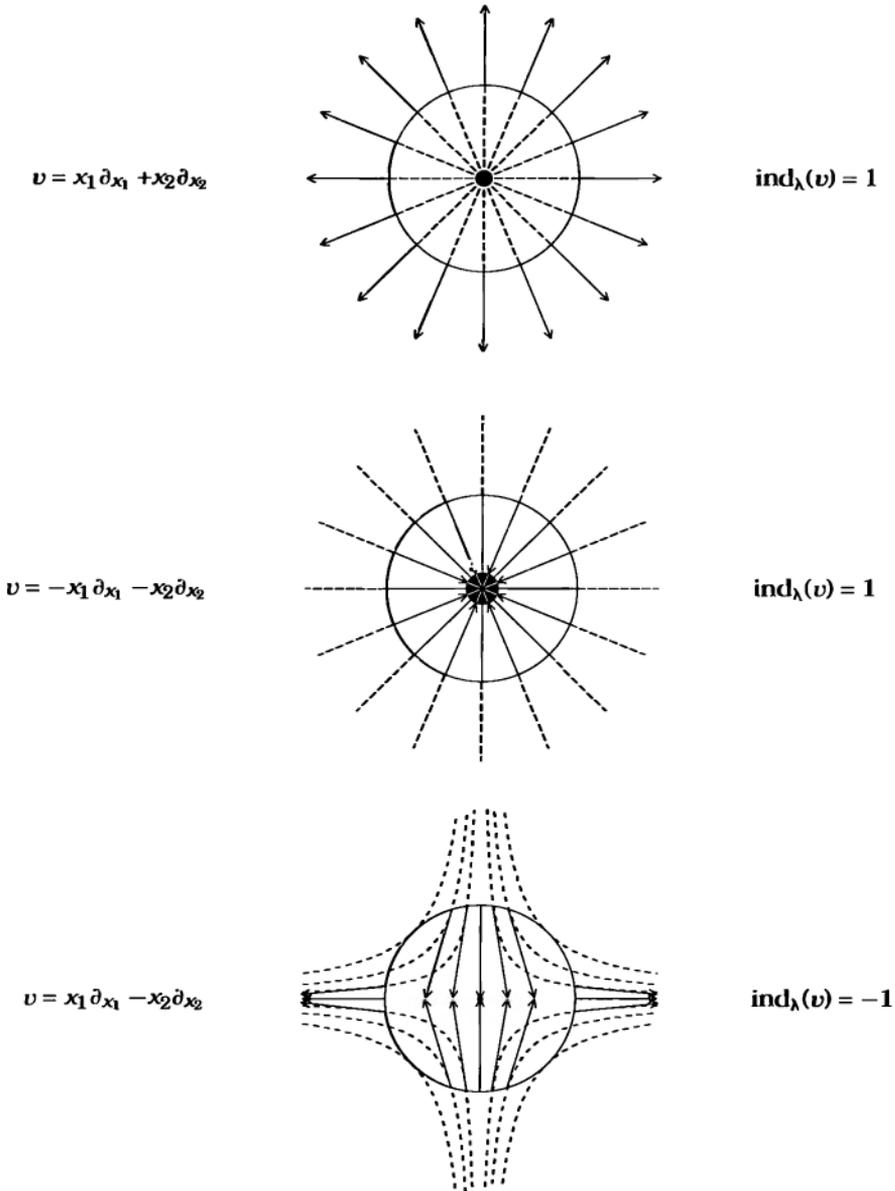


Figure 1. Indices de quelques champs de vecteurs remarquables (les pointillés sont des trajectoires intégrales).

1.2. Invariance de l'indice par déformation

Lemme 1.— $\text{ind}_\lambda(v)$ est invariant par déformation continue de λ dans U .

Lemme 2.— $\text{ind}_\lambda(v)$ est invariant par déformation continue du champ de vecteurs v (pourvu qu'il reste non singulier au cours de la déformation).

Ces deux lemmes sont intuitivement évidents, si l'on pense qu'une déformation continue (de λ , ou de v) ne peut modifier l'angle que de façon continue, et se traduit donc par une modification continue de l'indice ; or celui-ci est entier, et ne peut donc que rester constant.

En voici quelques conséquences remarquables.

Supposons que λ soit le cercle trigonométrique parcouru une fois dans le sens positif.

Alors :

Corollaire du lemme 1 : si v n'a aucun point singulier dans le disque unité, $\text{ind}_\lambda(v) = 0$;

Corollaire du lemme 2 : si v est partout « sortant » (ou partout « entrant ») le long du bord du disque, $\text{ind}_\lambda(v) = 1$.

Preuve des corollaires

1. Déformer λ par homothétie jusqu'à le réduire au lacet constant (point immobile au centre du disque !) et appliquer le lemme 1.
2. Se ramener, par déformation du champ de vecteurs (lemme 2), au premier (ou au deuxième) des exemples 1.1 (cf. Fig. 2). □

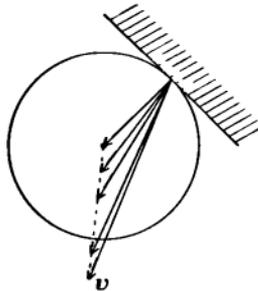


Figure 2. Comment déformer un vecteur « entrant » pour le faire « pointer vers le centre ».

Voici maintenant deux « grands » théorèmes, qui découlent eux aussi des deux lemmes en prenant pour λ un cercle.

Théorème de d'Alembert Tout polynôme de degré ≥ 1 admet au moins une racine complexe.

Preuve

Identifiant \mathbb{C} à \mathbb{R}^2 , considérons le polynôme étudié comme définissant un champ de vecteur v :

$$v(z) = z^n + a_1 z^{n-1} + \dots + a_n$$

(supposer le coefficient de z^n égal à 1 ne restreint évidemment pas la généralité). Soit λ un cercle de centre 0, de rayon r assez grand pour que $r^n > |a_1| r^{n-1} + \dots + |a_n|$. Quel que soit $\tau \in [0, 1]$, le champ de vecteurs $v_\tau(z) = z^n + \tau(a_1 z^{n-1} + \dots + a_n)$ est alors non singulier sur ce cercle, de sorte que (d'après le lemme 2) $\text{ind}_\lambda(v_\tau)$ ne dépend pas de τ . En particulier, $\text{ind}_\lambda(v) = \text{ind}_\lambda(v_0)$, qui vaut n puisque $v_0(z) = z^n$ et $\arg z^n = n \arg z$. Ayant un indice non nul le long de λ , v a donc au moins un point singulier dans le disque de rayon r , d'après le corollaire du lemme 1. □

Théorème du point fixe Toute application continue d'un disque fermé dans lui-même a au moins un point fixe.

Preuve

Soit $f : \bar{D} \rightarrow \bar{D}$ une application continue d'un disque fermé dans lui-même. Notant $v(x)$ le vecteur de \mathbb{R}^2 d'origine x et d'extrémité $f(x)$ (cf. Fig. 3), on obtient un champ de vecteurs v continu sur \bar{D} .

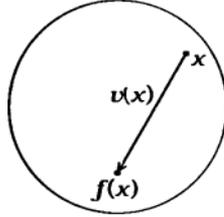


Figure 3.

Dire que f n'a pas de point fixe, c'est-à-dire que v n'a aucun point singulier ni dans D ni sur le bord. D'après le corollaire du lemme 1, cela devrait impliquer que $\text{ind}_{\partial D}(v) = 0$, où ∂D désigne le bord orienté de D . Or comme $v|_{\partial D}$ est purement entrant (cf. Fig. 4), son indice vaut 1 d'après le corollaire du lemme 2, d'où contradiction.

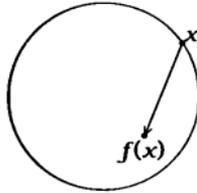


Figure 4.

□

2. Indice d'un point singulier isolé d'un champ de vecteurs du plan

Soit v un champ de vecteurs défini sur un ouvert U de \mathbb{R}^2 , et supposons que $s \in U$ soit un point singulier isolé de v ; autrement dit, s est l'unique point singulier de v dans un disque D , de centre s et de rayon assez petit.

Soit alors λ un lacet de $D_s^* = D_s \setminus \{s\}$, tournant une fois autour de s dans le sens direct. Bien que cette condition laisse beaucoup de liberté dans le choix de λ , il est intuitivement clair qu'à déformation de λ près le choix est unique. D'après 1.2, lemme 1, l'indice de v le long de λ ne dépend donc pas de ce choix; nous l'appellerons **indice du point singulier s de v** , et le noterons $\text{ind}_s(v)$.

2.1. Exemples : indice d'un point singulier simple

On appelle **point singulier simple** d'un champ de vecteurs un point singulier où la partie linéaire du champ de vecteurs est non dégénérée, c'est-à-dire (dans des coordonnées choisies de façon que le point singulier soit à l'origine) tel que :

$$v(x) = Ax + w(x)$$

où A est une matrice *inversible*, et $w \in o(\|x\|)$.

Comme A est inversible, on peut minorer $\|Ax\|$ par $c\|x\|$, où c est une constante indépendante de x (en effet, pour tout y on a $\|A^{-1}y\| \leq \|A^{-1}\| \|y\|$; en posant $y = Ax$ on obtient l'inégalité voulue avec $c = 1/\|A^{-1}\|$). Pour tout x dans un voisinage assez petit de l'origine on a donc $\|w(x)\| < \|Ax\|$, de sorte que pour tout $\tau \in [0, 1]$, le champ de vecteurs :

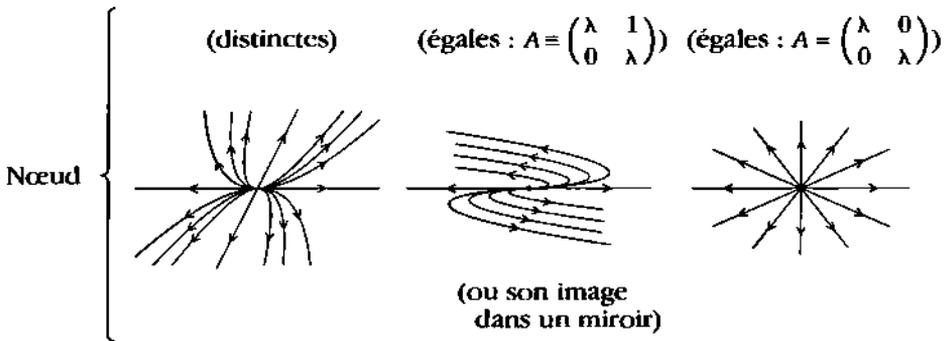
$$v_\tau(z) = Ax + \tau w(z)$$

a l'origine comme unique point singulier dans ce voisinage.

Nous en concluons (§ 1.2, lemme 2) que l'indice en un point singulier simple d'un champ de vecteurs est égal à celui de sa partie linéaire.

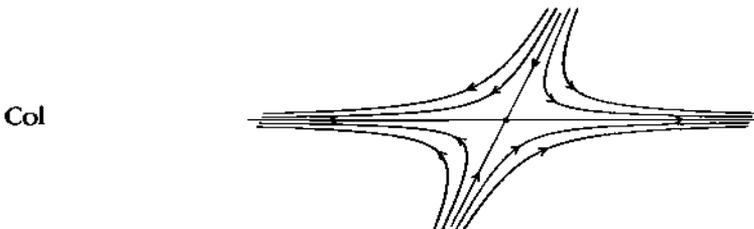
La figure 5 représente les différentes allures possibles des trajectoires intégrales d'un champ de vecteurs linéaire de \mathbb{R}^2 (c'est donc la réponse à l'exercice 2, b) du chapitre 3 !).

- Valeurs propres réelles > 0



- Valeurs propres réelles < 0
(mêmes figures, avec le sens des flèches inversé).

- Valeurs propres réelles de signes opposés



■ Valeurs propres imaginaires conjuguées $\lambda = \alpha + i\beta$, $\bar{\lambda} = \alpha - i\beta$

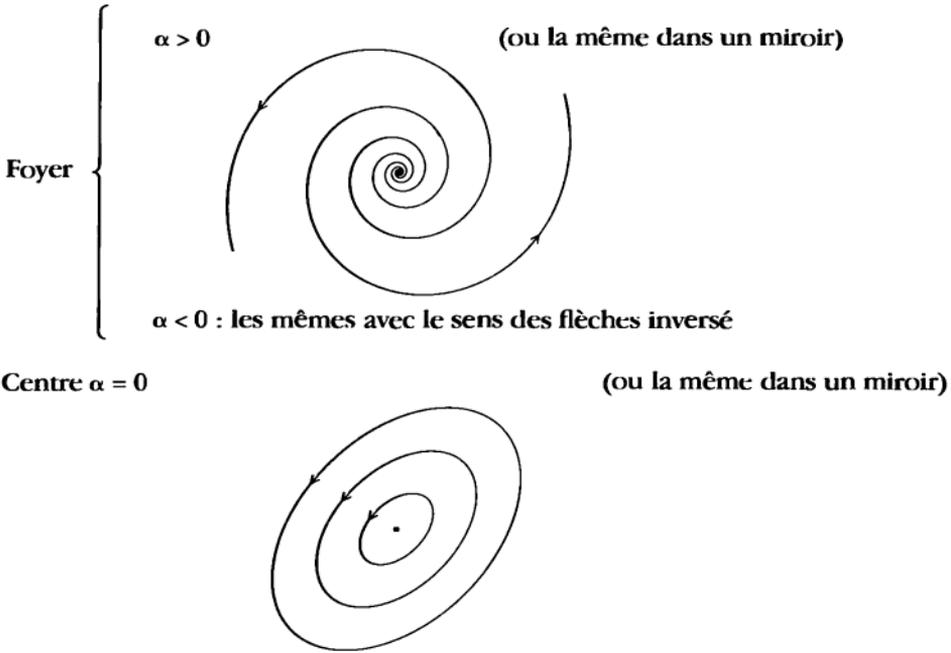


Figure 5. Trajectoires intégrales d'un champ de vecteurs linéaire de \mathbb{R}^2 .

Une simple inspection de ces dessins fournit la proposition suivante.

PROPOSITION

L'indice d'un point singulier simple vaut -1 si la partie linéaire du champ de vecteurs est du type « col », $+1$ dans tous les autres cas.

Corollaire En tout point critique quadratique non dégénéré d'une fonction f , l'indice du gradient de f vaut $+1$ dans le cas d'un sommet ou d'une cuvette, -1 dans le cas d'un col.

2.2. Retour au cas général

Théorème d'invariance de l'indice

$\text{ind}_s(v)$ est invariant par tout difféomorphisme local $\phi : \mathbb{R}^2, s \dots \rightarrow \mathbb{R}^2, s$.

Preuve

Désignons par ϕ_*v l'image du champ v par le difféomorphisme ϕ (cf. Chap. 3, § 3.0).

Il s'agit de démontrer que $\text{ind}_s(\phi_*v) = \text{ind}_s(v)$.

■ Première étape : se ramener au cas où ϕ est un difféomorphisme linéaire.

Supposant $s = 0$ pour simplifier les notations, écrivons ϕ sous la forme $x \mapsto Lx + \psi(x)$, où $L = D\phi_0$ est une application linéaire inversible et $\psi \in \mathcal{O}(\|x\|)$. Pour tout $\tau \in [0, 1]$, définissons ϕ_τ par $\phi_\tau(x) = Lx + \tau\psi(x)$. Il résulte du théorème d'inversibilité locale que ϕ_τ est une famille de difféomorphisme locaux de \mathbb{R}^2 , 0 dépendant

différentiablement de τ . Plus précisément l'application $\phi : (x, \tau) \rightarrow (\phi_\tau(x), \tau)$ est un difféomorphisme entre un ouvert U de $\mathbb{R}^2 \times [0, 1]$ contenant $0 \times [0, 1]$ et un ouvert V du même type (appliquer le théorème d'inversibilité locale au voisinage de chaque point $0 \times \tau_0$, et remarquer que les inverses des difféomorphismes locaux ainsi obtenus se recollent, par unicité, quand τ_0 parcourt $[0, 1]$).

Comme $[0, 1]$ est compact, V contient un ouvert de la forme $V_0 \times [0, 1]$, où V_0 est un voisinage fixe de 0 dans \mathbb{R}^2 . Pour tout $\tau \in [0, 1]$, le champ de vecteurs $\phi_{\tau*}v$ est donc bien défini dans V_0 , où il réalise une déformation continue entre $\phi_{*}v$ et L_*v .

Considérant la restriction de $\phi_{\tau*}v$ à un lacet λ de $V_0^* = V_0 \setminus \{0\}$ tournant une fois autour de 0 dans le sens positif, on conclut grâce à 1.2, lemme2, que :

$$\text{ind}_0(\phi_*v) = \text{ind}_0(L_*v)$$

■ Deuxième étape : montrer que $\text{ind}_0(L_*v) = \text{ind}_0(v)$

... Au lecteur de jouer ! □

Remarquons que le théorème est vrai même si $\det L < 0$ (difféomorphisme « inversant l'orientation ») : en effet, écrivant $(L_*v)(x) = (Lv)(L^{-1}x)$, on voit que L apparaît deux fois dans la définition de L_* , de sorte que les changements d'orientation se compensent (la transformation $x \mapsto L^{-1}x$ change un lacet λ en un lacet tournant dans « l'autre sens », mais la correspondance $v \mapsto Lv$ fait aussi tourner les vecteurs « dans l'autre sens »).

Question Supposons que quelqu'un, connaissant ce théorème mais pas la proposition du paragraphe 2.1, veuille démontrer le résultat final de 2.1 sur l'indice du gradient d'une fonction en un point critique quadratique non dégénéré. Remarquant que le résultat est évident pour $f = \pm x_1^2 \pm x_2^2$ (cf. exemples du § 1.1), il conclut : « D'après le lemme de Morse, le cas général se ramène à ce cas particulier. » Ce raisonnement est-il correct ?

2.3. Additivité des indices sur un domaine simple du plan

Soit D un « domaine simple » de \mathbb{R}^2 , c'est-à-dire un ouvert dont l'adhérence \bar{D} est homéomorphe à un disque fermé, de sorte que D est homéomorphe à l'intérieur de ce disque. On notera ∂D le bord de D « orienté dans le sens direct », c'est-à-dire dans le sens de parcours d'un promeneur qui longe le bord de D en laissant D sur sa gauche.

Théorème d'additivité des indices

Soit v un champ de vecteurs à singularités isolées dans \bar{D} , sans singularités sur le bord. Alors :

$$\text{ind}_{\partial D}(v) = \sum_{s \in D} \text{ind}_s(v)$$

Preuve

Remarquons que \bar{D} étant compact et les points singuliers étant isolés, ceux-ci sont en nombre fini. S'il n'y en a qu'un (ou aucun !) la proposition est une conséquence immédiate de 1.2, lemme 1, et de la définition 2.0 de l'indice. Le cas général se ramène à ce cas par le schéma de la figure 6, en remarquant que :

$$\text{ind}_{\partial D}(v) = \text{ind}_{\partial D_1}(v) + \cdots + \text{ind}_{\partial D_N}(v)$$

car les variations d'angles de v le long de $\partial D_1, \partial D_2, \dots, \partial D_N$ se compensent le long des composantes « mitoyennes » des bords orientés de D_1, \dots, D_N (les composantes mitoyennes sont orientées en sens inverse !).

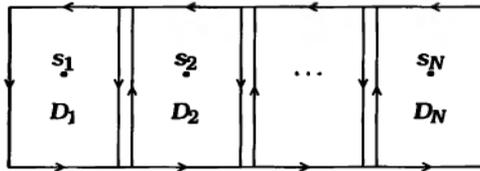


Figure 6.

□

Corollaire Soit f une fonction C^∞ sur un voisinage de \bar{D} , constante sur le bord, et n'ayant que des points critiques quadratiques non dégénérés dont aucun n'est sur le bord. Alors :

$$n_0 - n_1 + n_2 = 1$$

où n_0 désigne le nombre de cuvettes, n_1 le nombre de cols, n_2 le nombre de sommets (appliquer le lemme 2 du paragraphe 1.2.1 et le corollaire de la proposition du paragraphe 2.1).

Travaux pratiques Vérifier, en regardant une carte d'état-major, que le corollaire s'applique bien au relief d'une île montagneuse.

3. Caractéristique d'Euler d'une surface compacte

Les « surfaces » de ce paragraphe seront toujours supposées lisses, c'est-à-dire que ce seront des variétés à deux dimensions (de classe C^∞).

3.1. Définition de la caractéristique d'Euler en termes d'indices

Considérons sur une surface S un champ de vecteurs v ayant $s \in S$ comme point singulier isolé. Son image par une carte x de S , centrée en s , sera un champ de vecteurs x_*v de \mathbb{R}^2 ayant l'origine comme point singulier isolé, et il résulte de la proposition du paragraphe 2.2 (appliquée au changement de carte) que le nombre $\text{ind}_0(x_*v)$ ne dépend pas du choix de la carte x : nous l'appellerons **indice du point singulier** s de v , et le noterons $\text{ind}_s(v)$.

Étant donné sur une surface S un champ de vecteurs n'ayant que des singularités isolées et en nombre fini, nous poserons :

$$\chi(S, v) = \sum_s \text{ind}_s(v)$$

(où la somme porte sur tous les points singuliers).

Théorème Si la surface S est compacte, le nombre $\chi(S, v)$ ne dépend pas du choix de v , champ de vecteurs à singularités isolées sur S . On le note $\chi(S)$, et on l'appelle **caractéristique d'Euler de la surface S** .

L'idée de la démonstration consiste à décomposer S en « cellules » (homéomorphes à des domaines simples de \mathbb{R}^2) sur lesquelles on appliquera le théorème d'additivité des indices (§ 2.3). Mais attention : pour une telle cellule $D \subset S$, quel sens cela a-t-il de parler de $\text{ind}_{\partial D}(v)$? A priori aucun, comme le montre l'exemple suivant.

3.2. L'exemple de la sphère

S = la sphère S^2 ; D = l'un des deux hémisphères nord ou sud ; ∂D = l'équateur (orienté dans un sens à préciser).

Pour pouvoir parler de $\text{ind}_{\partial D}(v)$, on aurait envie de considérer la façon dont varie, quand a parcourt ∂D , l'angle que fait $v(a)$ « avec un vecteur fixe » de $T_a S^2$. Mais comme le plan tangent $T_a S^2$ varie avec a , cela n'a aucun sens. Par contre, en utilisant la structure d'espace vectoriel euclidien (orienté) des $T_a S^2$ (structure induite par le plongement de la sphère dans \mathbb{R}^3), cela a un sens de parler de la variation, quand a parcourt l'équateur, de l'angle orienté de deux vecteurs $v(a), w(a) \in T_a S^2$.

Précisons les questions d'orientation. On peut munir les espaces tangents $T_a S^2$ de l'orientation définie « par la normale extérieure » (cf. Fig. 7) : c'est l'orientation naturelle pour quelqu'un qui se promène sur la sphère, considérée comme bord d'un globe solide !

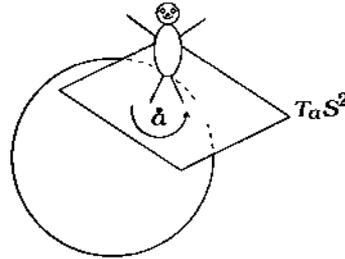


Figure 7. Orientation de S^2 « par la normale extérieure » (le sens direct de rotation dans $T_a S^2$ est symbolisé par un petit cercle orienté).

Avec cette convention, en désignant par D^+ et D^- les hémisphères nord et sud respectivement, on obtient comme bord orienté ∂D^+ l'équateur parcouru dans un sens (cf. Fig. 8), tandis que ∂D^- est l'équateur parcouru dans l'autre sens.

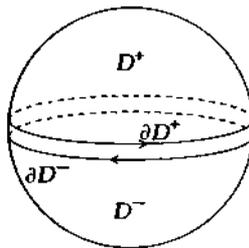


Figure 8. Bords orientés des deux hémisphères.

Considérons donc la façon dont varie, quand a parcourt ∂D^+ , l'angle orienté dans $T_a S^2$ des deux vecteurs $v(a)$ et $w(a)$. Pour calculer cette variation, que nous noterons

$\text{var}_{\partial D^+}(\widehat{v, w})$, donnons-nous une carte x d'un voisinage de $\overline{D^+}$, transformant $\overline{D^+}$ en le disque unité $\overline{D} \subset \mathbb{R}^2$ et respectant l'orientation, c'est-à-dire transformant le lacet ∂D^+ en le cercle trigonométrique ∂D parcouru dans le sens direct (ce sera le cas par exemple en prenant pour x la projection stéréographique sud).

Lemme $\text{var}_{\partial D^+}(\widehat{v, w}) = \text{var}_{\partial D}(x_*v, x_*w)$.

Dans le cas de la projection stéréographique, ce lemme est évident pour qui sait que la projection stéréographique conserve les angles. En fait, nous verrons dans la deuxième partie que les nombres entiers dont ce lemme affirme l'égalité peuvent (de même que la notion d'*indice*) être définis de façon *purement topologique*, l'usage des *angles* dans ces questions n'étant qu'un artifice.

Admettant donc le lemme, récrivons son membre de droite :

$$\frac{1}{2\pi} \text{var}_{\partial D}(x_*v, x_*w) = \text{ind}_{\partial D}(x_*w) - \text{ind}_{\partial D}(x_*v)$$

Or, d'après la proposition du paragraphe 2.3, on a :

$$\text{ind}_{\partial D}(x_*v) = \chi(D, x_*v) = \chi(D^+, v)$$

(par définition de χ en termes d'indices).

Compte tenu de cette égalité et de l'égalité analogue pour w , on a donc :

$$\frac{1}{2\pi} \text{var}_{\partial D^+}(\widehat{v, w}) = \chi(D^+, w) - \chi(D^+, v)$$

Le même travail pour l'hémisphère sud donne :

$$\frac{1}{2\pi} \text{var}_{\partial D^-}(\widehat{v, w}) = \chi(D^-, w) - \chi(D^-, v)$$

égalité qui, ajoutée à la précédente donne bien :

$$0 = \chi(S^2, w) - \chi(S^2, v)$$

puisque ∂D^+ et ∂D^- sont le même lacet *parcouru dans les deux sens opposés*. Nous avons ainsi démontré le théorème 3.1 dans le cas de la sphère.

EXERCICE En utilisant un champ de vecteurs de votre choix, calculer $\chi(S^2)$ et en déduire le « théorème du hérisson ».

3.3. Cas général

Supposons construit un « pavage » de S par des « cellules » (de dimension 2) : chaque cellule est homéomorphe à un domaine simple de \mathbb{R}^2 , et munie d'une orientation ; les cellules sont en nombre fini, ne se chevauchent pas, et leurs adhérences recouvrent S .

Pour chaque cellule D , un raisonnement identique à celui que nous avons fait en 3.2 pour un hémisphère montre que :

$$\frac{1}{2\pi} \text{var}_{\partial D}(\widehat{v}, \widehat{w}) = \chi(D, w) - \chi(D, v)$$

où ∂D désigne le bord orienté de D (on suppose bien entendu que le bord de D ne contient aucun point singulier de nos champs de vecteurs).

En sommant cette égalité sur toutes les cellules, on trouve :

$$\frac{1}{2\pi} \sum_D \text{var}_{\partial D}(\widehat{v}, \widehat{w}) = \chi(S, w) - \chi(S, v)$$

et tout le problème est de se convaincre que le membre de gauche est nul.

Ce membre de gauche est une somme de variations d'angles le long du « squelette de dimension 1 » de notre pavage (union des frontières de toutes nos cellules). Décomposons ce squelette en un nombre fini de « segments » (homéomorphes à des intervalles de \mathbb{R}), assez courts pour que chaque segment σ soit mitoyen à deux cellules D' , D'' et pas davantage (il peut arriver que $D' = D''$ comme le montre l'exemple du segment σ de la figure 9 ; mais les raisonnements qui suivent n'en sont pas affectés, le segment σ partageant toujours *localement* S en deux morceaux).

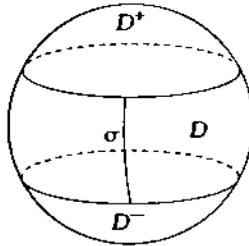


Figure 9. Autre exemple de décomposition de la sphère en cellules.

Au voisinage d'un tel segment σ , deux situations peuvent se produire.

- Premier cas : les orientations choisies pour D' et D'' « concordent » au voisinage de σ (Fig. 10).

Alors le long de σ , l'orientation de $\partial D'$ est opposée à celle de $\partial D''$ de sorte que (comme cela se passait dans notre raisonnement 3.2) les variations d'angles $\text{var}_{\partial D'}$ et $\text{var}_{\partial D''}$ le long de σ se compensent.



Figure 10. Orientations concordantes

- Deuxième cas : au voisinage de σ , les orientations choisies pour D' et D'' sont opposées (Fig. 11).

Alors, le long de σ , l'orientation de $\partial D'$ est la même que celle de $\partial D''$; mais la mesure de l'angle orienté \widehat{v}, \widehat{w} change de signe selon que le plan tangent $T_a S$ est

orienté comme D' ou comme D'' : il en résulte donc, là encore, une compensation des deux variations d'angles $\text{var}_{\partial D'}$ et $\text{var}_{\partial D''}$.



Figure 11. Orientations opposées

Remarque Dans le cas où la surface S est plongée dans \mathbb{R}^3 , on peut s'arranger pour que seul le premier cas se présente (on dit alors que la surface est *orientable*). En effet, comme S est compacte, son complémentaire $\mathbb{R}^3 \setminus S$ contient un « voisinage de l'infini » (complémentaire d'une grosse boule) ; appelons « extérieur de S » la composante connexe de $\mathbb{R}^3 \setminus S$ qui contient ce voisinage de l'infini ; son bord n'est autre que la surface S , que l'on peut ainsi « orienter par la normale extérieure », comme nous l'avons fait pour le cas de la sphère.

DEUXIÈME PARTIE
POINT DE VUE INTRINSÈQUE
SUR LES VARIÉTÉS

Introduction

Le calcul différentiel est né avec l'idée de décrire l'évolution temporelle de « quantités fluentes » (comme les appelait Newton). Dans la première partie de ce livre, les quantités fluentes étaient des n -uplets de nombres (indépendants ou liés par des relations *a priori*) : « points mobiles » dans \mathbb{R}^n ou dans des sous-variétés de \mathbb{R}^n .

Mais la mécanique, ainsi que la géométrie, ne se contentent pas de vouloir décrire le mouvement des *points* : elles veulent aussi décrire le mouvement des *corps* (solides par exemple), ou le mouvement des *figures*.

Considérons par exemple la plus simple des figures qu'on puisse rencontrer en géométrie : une droite du plan. Comment décrire un mouvement de cette figure ? Ayant choisi dans le plan des coordonnées cartésiennes (x, y) , supposons que pendant un certain intervalle de temps la droite D ne soit jamais « verticale », c'est-à-dire parallèle à l'axe des y . Une droite D non verticale peut être représentée par une équation de la forme :

$$y = px + q \tag{1}$$

et la donnée de sa position équivaut à la donnée de deux nombres p et q , que nous pouvons appeler « coordonnées » de la droite D . La donnée d'un mouvement de D équivaut ainsi à la donnée de deux fonctions $p(t)$ et $q(t)$, et l'on dira que le mouvement est différentiable si ces deux fonctions sont différentiables.

Sur un intervalle de temps où la droite D peut être verticale, la représentation (1) ci-dessus n'est plus adaptée. Mais quitte à raccourcir l'intervalle de temps, on peut alors supposer que D n'est jamais « horizontale », et procéder comme ci-dessus en remplaçant la représentation (1) par :

$$x = \tilde{p}y + \tilde{q} \tag{1}$$

ce qui fournit un autre « système de coordonnées » (\tilde{p}, \tilde{q}) , lié au précédent (dans le cas des droites ni verticales ni horizontales) par les relations :

$$\begin{cases} \tilde{p} = 1/p \\ \tilde{q} = -q/p \end{cases} \quad (p \neq 0, \tilde{p} \neq 0) \tag{2}$$

Comme les relations (2) sont différentiables, de même que les relations réciproques qui leur sont équivalentes :

$$\begin{cases} p = 1/\tilde{p} \\ q = -\tilde{q}/\tilde{p} \end{cases}$$

l'usage de l'un ou l'autre des systèmes de coordonnées (p, q) ou (\tilde{p}, \tilde{q}) conduit à la même notion de différentiabilité du mouvement sur les intervalles de temps où D n'est ni verticale ni horizontale.

Les quelques lignes qui précèdent illustrent tout le contenu du chapitre 4 : l'ensemble des droites du plan est un exemple de ce qu'on appelle une **variété différentielle** ; le fait qu'on puisse caractériser localement la position d'une droite D par deux « coordonnées » (par exemple (p, q) si D n'est pas verticale, ou (\tilde{p}, \tilde{q}) si D n'est pas horizontale) se traduit en disant que cette variété M est à deux dimensions, le couple (p, q) ou (\tilde{p}, \tilde{q}) étant un **système de coordonnées locales** (définissant une **carte locale**) de cette variété.

Enfin, notre définition d'un « mouvement différentiable » illustre la notion de **différentiabilité d'une application** (ici l'application de \mathbb{R} dans M qui à t associe $D(t)$).

Passons maintenant au contenu du chapitre 5. Ayant compris ce qu'est un mouvement différentiable, nous voulons comprendre ce qu'est le **vecteur vitesse** d'un tel mouvement.

Étant donnée une droite D_0 , considérons toutes les façons possibles de la « mettre en mouvement » de façon différentiable (applications différentiables $t \mapsto D(t)$ telles que $D(0) = D_0$). Nous dirons que deux tels mouvements sont **tangents** si, dans un système adapté de coordonnées locales (par exemple les coordonnées p, q , si D_0 n'est pas verticale) les dérivées en 0 de chacune des fonctions $p(t), q(t)$ sont les mêmes pour ces deux mouvements. Dans l'ensemble de tous les mouvements possibles de D_0 , la relation de tangence est une relation d'équivalence. La classe d'équivalence d'un mouvement est ce qu'on appelle le **vecteur vitesse** de ce mouvement. Le mot « vecteur » se justifie par le fait que l'ensemble de ces classes d'équivalence forme effectivement un **espace vectoriel**, de dimension égale à celle de M (ici 2 : les deux nombres $p'(0), q'(0)$ sont les « coordonnées » du vecteur vitesse dans la carte (p, q)). Cet espace vectoriel est appelé **espace tangent à M en D_0** , et noté $T_{D_0}M$.

Sur la lancée des explications précédentes, le lecteur est vivement encouragé à attaquer directement l'étude 5 (familles à un paramètre de droites du plan), remettant à plus tard la lecture des chapitres 4 et 5 : la lecture de ces chapitres lui sera plus profitable quand il pourra la considérer comme une « mise en ordre théorique » d'idées avec lesquelles il se sera déjà familiarisé par la pratique.

Variétés différentielles

L'idée des notions que nous présentons dans ce chapitre remonte à Bernhard Riemann (1826-1866) : dans une dissertation intitulée « Sur les hypothèses qui servent de base à la géométrie » (1854), Riemann introduisait sans la définir une notion qu'il appelait *m-dimensionale Mannigfaltigkeit*, terminologie qui s'est conservée de nos jours (ainsi que sa traduction anglaise *m dimensional manifold*) ; beaucoup de vieux ouvrages français utilisent l'expression « multiplicité m fois étendue » (ou « à m dimensions »). On dit aujourd'hui en français *variété à m dimensions*, terminologie qui malheureusement est aussi appliquée en français à une classe d'objets non nécessairement lisses, pour lesquels les Anglais remplacent le mot *manifold* par *variety*.

Bien que la notion introduite par Riemann se soit rapidement imposée comme une notion centrale de la géométrie, il a fallu près d'un siècle pour que les mathématiciens arrivent à la formuler de façon précise. Avant de donner cette formulation, je vais essayer de suggérer par quelques exemples l'idée intuitive de Riemann, et avant tout l'idée capitale de *dimension* d'une variété.

Exemple 1.- L'ensemble des droites affines du plan est une variété à deux dimensions.

Pourquoi deux dimensions ? Pour se donner une droite affine du plan on peut commencer par se donner sa direction : ce premier choix dépend d'un seul paramètre (« un degré de liberté »), par exemple l'angle que fait la direction choisie avec une direction de référence.

Il reste ensuite à préciser la position de la droite choisie au sein de la famille des droites affines de direction donnée : ce deuxième choix dépend lui aussi d'un seul paramètre, par exemple l'abscisse du point d'intersection avec une droite transverse à la direction donnée.

On a donc en tout deux *degrés de liberté*.

Exemple 2.- L'ensemble des droites affines de l'espace \mathbb{R}^3 est une variété à quatre dimensions.

En effet,

- i) la direction dépend maintenant de deux degrés de liberté (pourquoi ?) ;
- ii) une fois choisie la direction, il reste encore deux degrés de liberté (par exemple, le choix de la position du point d'intersection avec un plan transverse à la direction donnée).

EXERCICE Quelle est la dimension de la variété des positions d'un corps solide de l'espace \mathbb{R}^3 n'ayant aucune symétrie particulière ?

Pour mettre en forme l'idée de variété, on définira d'abord la notion de *variété topologique* (§ 0), qui nous suffira pour comprendre précisément ce qu'est la dimension. Puis les *variétés différentielles* seront définies comme des variétés topologiques munies d'une structure supplémentaire qui nous permettra de faire du calcul différentiel, en définissant les applications différentiables, leurs applications tangentes, etc., comme nous l'avons déjà fait pour les ouverts de \mathbb{R}^n et les sous-variétés de \mathbb{R}^n .

0. Variétés topologiques

0.0.

Une **variété topologique à m dimensions** est un espace topologique M dont tout point a admet un voisinage ouvert U homéomorphe à un ouvert de \mathbb{R}^m .

La donnée d'un tel homéomorphisme :

$$x : U \rightarrow \underline{U} \subset \mathbb{R}^m$$

est appelée *carte (locale) de M* (au voisinage de a).

L'ouvert $U \subset M$ est le *domaine de la carte*.

Un atlas \mathcal{A} de M est une famille de cartes x dont les domaines U_x recouvrent M :

$$M = \bigcup_{x \in \mathcal{A}} U_x .$$

Pour tout couple de cartes x, \tilde{x} de domaines respectifs $U_x, U_{\tilde{x}}$ on définit l'homéomorphisme de **changement de carte**, que nous noterons abusivement $\tilde{x} \circ x^{-1}$, par le diagramme de la figure 1.

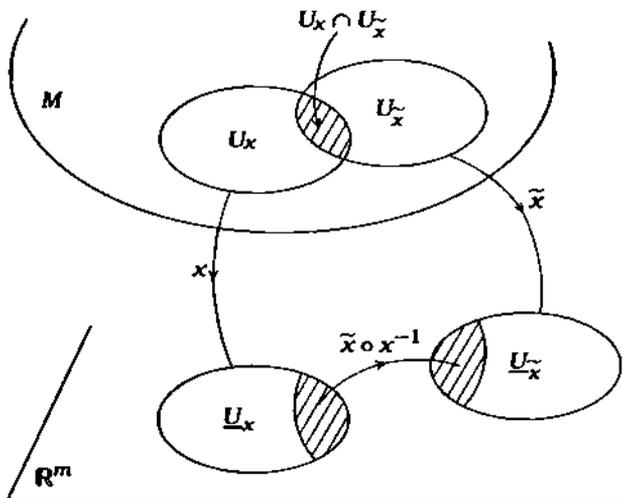


Figure 1.

$\tilde{x} \circ x^{-1}$ est donc une abréviation pour :

$$(\tilde{x}|_{U_x \cap U_{\tilde{x}}}) \circ (x^{-1}|_{x(U_x \cap U_{\tilde{x}})})$$

c'est un homéomorphisme entre les ouverts de \mathbb{R}^m hachurés sur la figure, $x(U_x \cap U_{\tilde{x}})$ et $\tilde{x}(U_x \cap U_{\tilde{x}})$ (homéomorphisme vide si $U_x \cap U_{\tilde{x}} = \emptyset$!).

Remarque Invariance topologique de la dimension Un théorème célèbre de topologie assure qu'un ouvert de \mathbb{R}^m ne peut pas être homéomorphe à un ouvert de $\mathbb{R}^{m'}$ pour $m' \neq m$. (N.B. : rappelons que ce théorème est évident dans le cas des difféomorphismes ; mais il est loin de l'être pour les homéomorphismes !) Il en résulte que la notion de dimension d'une variété topologique est définie sans ambiguïté : une variété topologique à m dimensions ne peut être en même temps une variété topologique à m' dimensions, avec $m' \neq m$.

0.1. Retour à l'exemple 1 de l'introduction : l'ensemble des droites affines du plan

Dire que cet ensemble est une variété topologique (de dimension 2) n'a a priori aucun sens tant qu'on n'a pas défini une topologie sur l'ensemble en question.

L'intuition suggère que cette topologie nous soit en quelque sorte imposée (par décret divin ?) : la notion d'une « suite de droites du plan tendant vers une droite donnée » ne devrait-elle pas avoir un sens évident dépourvu de toute ambiguïté ? Mais essayez donc de la mettre en forme... (le débat est ouvert).

Au cas où vous n'auriez pas réussi, une autre stratégie possible consiste à se demander ce qui peut subsister de la notion d'atlas quand on n'a pas su mettre une topologie sur l'ensemble à étudier : tout ce qu'on peut alors demander à une carte, c'est d'être une *bijection* (la notion d'homéomorphisme n'a plus de sens) ; par contre cela garde un sens d'exiger que les changements de cartes soient des homéomorphismes entre ouverts de \mathbb{R}^m .

0.2. Atlas continu

Un *atlas continu* (de dimension m) sur un ensemble M est une collection \mathcal{A} de bijections :

$$x : U_x \rightarrow \underline{U}_x \subset \mathbb{R}^m$$

entre des parties U_x de M recouvrant M et des ouverts \underline{U}_x de \mathbb{R}^m , telle que pour tous $x, \tilde{x} \in \mathcal{A}$ les bijections :

$$\tilde{x} \circ x^{-1} : x(U_x \cap U_{\tilde{x}}) \rightarrow \tilde{x}(U_x \cap U_{\tilde{x}})$$

soient des homéomorphismes entre ouverts de \mathbb{R}^m .

PROPOSITION

Étant donné un atlas continu \mathcal{A} (de dimension m) sur un ensemble M , il existe une et une seule topologie sur M telle que les bijections $x \in \mathcal{A}$ soient des homéomorphismes entre ouverts de M et ouverts de \mathbb{R}^m . Autrement dit, la donnée d'un atlas continu sur un ensemble M définit sur cet ensemble une structure de variété topologique.

La preuve de cette proposition est un « yoga » facile de topologie générale : on définit la topologie de M en disant qu'une partie $U \subset M$ est ouverte si et seulement si $\forall x \in A \ x(U \cap U_x)$ est un ouvert de \mathbb{R}^m .

0.3. Encore l'exemple 1 : construction d'un atlas continu sur l'ensemble M des droites affines du plan

Notant (x, y) les fonctions coordonnées dans le plan, désignons par U l'ensemble des droites non verticales (c'est-à-dire non parallèles à l'axe des y) et par \tilde{U} l'ensemble des droites non horizontales (c'est-à-dire non parallèles à l'axe des x). Évidemment $M = U \cup \tilde{U}$.

Une droite non verticale peut être considérée comme le graphe d'une application affine $x \mapsto y = px + q$ de \mathbb{R} dans \mathbb{R} . La donnée de $D \in U$ équivaut ainsi à la donnée des deux nombres (p, q) et l'on définit ainsi une bijection :

$$\begin{aligned} \phi : U &\rightarrow \mathbb{R}^2 \\ D &\mapsto (p(D), q(D)). \end{aligned}$$

De même, une droite non horizontale peut être considérée comme le graphe d'une application affine $y \mapsto x = \tilde{p}y + \tilde{q}$, et l'on définit ainsi une bijection.

$$\begin{aligned} \tilde{\phi} : \tilde{U} &\rightarrow \mathbb{R}^2 \\ D &\mapsto (\tilde{p}(D), \tilde{q}(D)). \end{aligned}$$

$U \cap \tilde{U}$ est l'ensemble des droites ni verticales ni horizontales : il est envoyé par ϕ sur $\mathbb{R}^* \times \mathbb{R}$ ($D \in U$ est non horizontale si et seulement si $p(D) \neq 0$), et par $\tilde{\phi}$ sur $\mathbb{R}^* \times \mathbb{R}$ également ($D \in \tilde{U}$ est non verticale si et seulement si $\tilde{p}(D) \neq 0$).

Le changement de carte $\tilde{\phi} \circ \phi^{-1}$ est l'application de $\mathbb{R}^* \times \mathbb{R}$ dans lui-même qui à (p, q) associe $(\tilde{p} = 1/p, \tilde{q} = -q/p)$. Cette application est continue (et même différentiable !), et il en est de même de son inverse, donné par la même formule avec (p, q) , (\tilde{p}, \tilde{q}) échangés.

1. Structures différentielles sur une variété

1.0. Atlas différentiable

Renforçons la définition 0.2 d'un atlas continu en exigeant que les changements de cartes $\tilde{x} \circ x^{-1}$ soient non seulement des homéomorphismes mais des *difféomorphismes* (entre ouverts de \mathbb{R}^m).

On obtient ainsi la notion d'*atlas différentiable* (si l'on demande aux changements de cartes d'être des difféomorphismes *analytiques* on parlera d'*atlas analytique*).

La donnée sur M d'un atlas différentiable \mathcal{A} permet de définir la notion de *fonction différentiable sur M* : une fonction $f : M \rightarrow \mathbb{R}$ (ou \mathbb{R}^p) est dite différentiable (pour l'atlas \mathcal{A}) si pour toute carte $x \in A$ la fonction $(f|_{U_x}) \circ x^{-1}$ est différentiable (sur son domaine de définition $\underline{U}_x \subset \mathbb{R}^m$).

Propriétés de la notion de différentiabilité d'une fonction

- i) C'est une notion locale : une fonction f est différentiable sur M pour l'atlas \mathcal{A} si et seulement si, pour tout $a \in M$, il existe un voisinage U de a tel que $f|_U$ soit différentiable pour l'atlas $\mathcal{A}|_U$ (atlas de U dont les cartes sont les $x|_{U \cap U_x}$, $x \in \mathcal{A}$).
- ii) Cette notion ne dépend du choix de l'atlas différentiable \mathcal{A} qu'à équivalence près, au sens de la définition que voici.

1.1. Équivalence d'atlas. Structure différentielle sur une variété

Deux atlas différentiables \mathcal{A} et $\tilde{\mathcal{A}}$ de M sont dits **différentiablement équivalents** si, pour tous $x \in \mathcal{A}$, $\tilde{x} \in \tilde{\mathcal{A}}$, le changement de carte $\tilde{x} \circ x^{-1}$ est différentiable (ce qui revient à dire que l'atlas $\mathcal{A} \cap \tilde{\mathcal{A}}$ est différentiable).

On appelle **structure différentielle** sur une variété topologique¹ M la donnée, à équivalence différentiable près, d'un atlas différentiable de M .

On appelle **variété différentielle**² la donnée d'une variété topologique et d'une structure différentielle sur cette variété.

On définit de même la notion de **variété analytique** en remplaçant partout dans ce qui précède le mot « différentiable » par « analytique ».

Exemples de base Les sous-variétés de \mathbb{R}^n (chap. 1 et 2) sont des variétés différentielles (cf. chap. 2, § 3 pour la construction d'atlas).

Nous venons par ailleurs de rencontrer un exemple de variété différentielle qui n'est pas donnée comme sous-variété de \mathbb{R}^n : l'ensemble des droites affines du plan (l'atlas construit en 0.3 était différentiable, et même analytique).

Voici d'autres exemples très importants.

1.2. Espaces projectifs

Soit E un espace vectoriel sur \mathbb{R} de dimension finie $m+1 \geq 2$. L'ensemble des droites vectorielles de E est appelé **espace projectif** sur E et noté $\mathbb{P}(E)$ (si $E = \mathbb{R}^{m+1}$ on écrit $\mathbb{P}(\mathbb{R}^{m+1}) = \mathbb{R}\mathbb{P}^m$ ou simplement \mathbb{P}^m).

On peut considérer $\mathbb{P}(E)$ comme l'espace quotient de $E \setminus 0$ par la relation d'équivalence « proportionnalité des vecteurs » :

$$u \sim u' \iff \exists \lambda \in \mathbb{R}^* , u' = \lambda u$$

L'espace projectif est muni canoniquement d'une structure de variété différentielle (et même analytique) de dimension m dont les cartes peuvent être définies comme suit.

¹ En fait la topologie de M n'a pas besoin d'être donnée au départ, puisque la donnée d'un atlas suffit à la définir (cf. Prop. 0.2).

² ou, selon la terminologie traditionnelle, *variété différentiable*. La terminologie « variété différentielle » (qui est celle de Bourbaki) tend à souligner qu'il ne s'agit pas d'énoncer une *propriété* de l'objet appelé variété (comme quand on dit d'une fonction qu'elle est différentiable), mais de *donner une structure* sur cet objet (comme un « espace topologique » est un espace muni d'une structure topologique). Par contre *dans le cas des sous-variétés* la différentiabilité est effectivement une *propriété* de l'objet (puisque la structure est déjà imposée par l'espace ambiant).

Ayant choisi une base de E , notons $X = (X_0, X_1, \dots, X_m)$ la base duale (système de coordonnées linéaires sur E). Se donner une droite $D \in \mathbb{P}(E)$ c'est se donner, à multiplication par un scalaire près, un $(m+1)$ -uplet de nombres non tous nuls X_0, X_1, \dots, X_m (les composantes d'un vecteur quelconque de D). Un tel $(m+1)$ -uplet est appelé **système de coordonnées homogènes** de la droite D , et on note $X_0 : X_1 : \dots : X_m$ sa classe d'équivalence à proportionnalité près.

Si la droite D n'est pas contenue dans un hyperplan $\{X_0 = 0\}$, la classe $X_0 : X_1 : \dots : X_m$ contient un unique représentant de la forme $(1, x_1, \dots, x_m)$ (où $x_i = X_i/X_0$). On a ainsi défini une bijection :

$$x = (x_1, \dots, x_m) : U_X \rightarrow \mathbb{R}^m$$

où U_X désigne l'ensemble des droites non contenues dans l'hyperplan $\{X_0 = 0\}$.

Ayant ainsi construit les cartes, vérifions que les changements de cartes sont différentiables (et même analytiques) sur leurs ouverts de définition.

Deux systèmes X et Y de coordonnées linéaires sur E sont liés par une formule de changement de coordonnées :

$$Y_i = \sum_{j=0}^m a_{ij} X_j \quad i = 0, 1, \dots, m$$

où (a_{ij}) est une matrice inversible.

En notant $y = (y_1, \dots, y_m) : U_Y \rightarrow \mathbb{R}^m$ la carte associée au système Y , on vérifie immédiatement que le changement de carte $y \circ x^{-1}$ a pour ouvert de définition :

$$x(U_X \cap U_Y) = \{x \in \mathbb{R}^m \mid a_{00} + a_{01}x_1 + \dots + a_{0m}x_m \neq 0\}$$

et qu'il est défini sur cet ouvert par la formule :

$$y_i = (a_{i0} + a_{i1}x_1 + \dots + a_{im}x_m) / (a_{00} + a_{01}x_1 + \dots + a_{0m}x_m)$$

($i = 1, 2, \dots, m$), évidemment analytique en les variables x_1, \dots, x_m .

Remarque Au lieu de l'atlas pléthorique que nous venons de construire (autant de cartes que de bases de E !) on aurait pu se contenter d'un atlas à $m+1$ cartes, extrait du précédent en choisissant sur E un système de coordonnées linéaires X et en le permutant circulairement (les domaines de définition respectifs de ces $m+1$ cartes sont $U_i = \{D \in \mathbb{P}(E) \mid X_i(D) \neq 0\}$, $i = 0, 1, \dots, m$). L'intérêt de notre construction pléthorique est de montrer que la structure différentielle de $\mathbb{P}(E)$ ne dépend que de E et pas du choix d'une base.

1.3. Variétés grassmanniennes

E étant un espace vectoriel de dimension strictement supérieure à p (disons, $\dim E = p + q$, $q \geq 1$), on appelle **grassmannienne des p -plans de E** , et on note $\mathbb{G}_p(E)$ l'ensemble des sous-espaces vectoriels de dimension p de E (l'espace projectif correspond donc au cas particulier $p = 1$). La grassmannienne $\mathbb{G}_p(E)$ est munie canoniquement d'une structure de variété analytique de dimension pq , au moyen des cartes que voici.

Soit $X = (X_1, \dots, X_p; X_{p+1}, \dots, X_{p+q})$ un système de coordonnées linéaires sur E ; on utilisera les notations abrégées $X^\circ = (X_1, \dots, X_p)$, $X^* = (X_{p+1}, \dots, X_{p+q})$. Soit U_X l'ensemble des p -plans transverses au q -plan $\{X^0 = 0\}$: dire qu'un p -plan est dans U_X , c'est dire que c'est le graphe d'une application linéaire de l'espace des X° dans l'espace des X^* ; en désignant par $x = (x_{ij})$ la matrice de cette application linéaire, on obtient ainsi une bijection :

$$x = (x_{ij})_{\substack{j=1, \dots, p \\ i=1, \dots, q}} : U_X \rightarrow \mathbb{R}^{pq}.$$

Le calcul des changements de cartes est un simple exercice d'algèbre matricielle :

notant $\left(\begin{array}{c|c} a^{\circ\circ} & a^{\circ*} \\ \hline a^{*\circ} & a^{**} \end{array} \right)$ la matrice qui fait passer des coordonnées $X = (X^\circ, X^*)$ aux coordonnées $Y = (Y^\circ, Y^*)$ de E , on trouve que le changement de carte $y \circ x^{-1}$ est donné par la formule matricielle :

$$y = (a^{*\circ} + a^{**}x)(a^{\circ\circ} + a^{\circ*}x)^{-1},$$

avec comme domaine de définition l'ensemble des matrices pour lesquelles $a^{\circ\circ} + a^{\circ*}x$ est inversible.

2. Morphismes de variétés

Le but de ce paragraphe est de définir la notion d'*application différentiable* [resp. *analytique*] d'une variété différentielle [resp. analytique] dans une autre.

Nous nous placerons, pour fixer les idées, dans le cas différentiable, mais le discours serait exactement le même dans le cas analytique (au changement d'adjectif près). L'adjectif en question sera sous-entendu quand nous parlerons des variétés, cartes, atlas,... (toutes les variétés seront supposées munies d'une structure différentielle [resp. analytique], et nous ne considérerons que des cartes ou atlas compatibles avec cette structure).

2.0. Applications différentiables

Soient M et N deux variétés. Une application $f : M \rightarrow N$ est dite *différentiable* si les coordonnées du point $f(a) \in N$ dépendent différentiablement des coordonnées du point $a \in M$. Plus précisément, f est dite **différentiable au voisinage de a_0** si, pour une carte locale x de M au voisinage de a_0 et pour une carte locale y de N au voisinage de $f(a_0)$, l'application $y \circ f \circ x^{-1}$ est différentiable au voisinage de $x(a_0)$. (N.B. : $y \circ f \circ x^{-1}$ est une notation abrégée : en fait il faut restreindre x^{-1} et f aux ouverts hachurés sur la figure 2.)

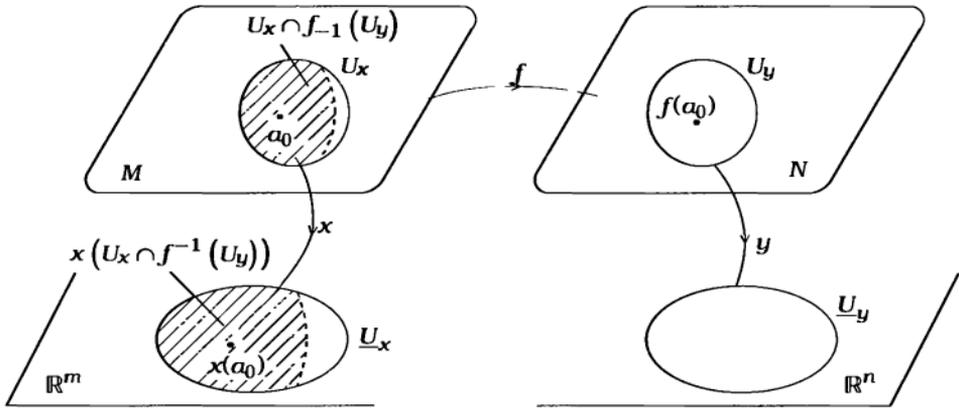


Figure 2.

Notons que la propriété de différentiabilité au voisinage de a_0 ne dépend pas du choix des cartes x, y puisque par définition d'une structure différentielle les changements de cartes sont différentiables.

Une application $f : M \rightarrow N$ est dite **différentiable** si elle l'est au voisinage de tout point $a_0 \in M$, autrement dit s'il existe un atlas \mathcal{A}_M de M et un atlas \mathcal{A}_N de N tels que $\forall x \in \mathcal{A}_M, y \in \mathcal{A}_N$ l'application $y \circ f \circ x^{-1}$ soit différentiable.

Cas particulier des fonctions numériques Dans le cas $N = \mathbb{R}$ (ou \mathbb{R}^n) on retrouve bien la notion de différentiabilité introduite au paragraphe 1.

Stabilité par composition

Il résulte immédiatement de la définition que si $f : M \rightarrow N$ et $g : N \rightarrow P$ sont différentiables, $g \circ f : M \rightarrow P$ est différentiable. Par ailleurs, l'application identique \mathbb{I}_M est évidemment différentiable !

Nous pouvons ainsi parler de la *catégorie des variétés différentielles* (cf. chap. 2, § 3.2), dont les objets sont les variétés différentielles et dont les « morphismes » sont les applications différentiables.

Les isomorphismes de cette catégorie sont appelés **difféomorphismes**. Si l'on travaille dans la catégorie des variétés analytiques (où les morphismes sont les applications analytiques), les isomorphismes sont appelés *difféomorphismes analytiques* (ou *isomorphismes analytiques*).

Exemple Une carte (locale) d'une variété M est un isomorphisme entre un ouvert U de M et un ouvert \underline{U} de \mathbb{R}^m .

Produits

Si M et N sont deux variétés (différentielles, ou analytiques...) il existe sur l'ensemble produit $M \times N$ une structure canonique de variété (différentielle, analytique...), structure dite de **variété produit**, caractérisée par la propriété suivante : une application f d'une variété P (différentielle, analytique...) dans $M \times N$ est un morphisme si et seulement si ses deux composantes $pr_M \circ f$ et $pr_N \circ f$ sont des morphismes (où pr_M

et pr_N désignent les projections canoniques de $M \times N$ sur M et N respectivement) ; en particulier pr_M et pr_N sont elles-mêmes des morphismes (faire $f = \mathbb{I}_{M \times N}$).

Cette structure de variété produit peut être définie à partir d'un atlas \mathcal{A}_M de M et d'un atlas \mathcal{A}_N de N , au moyen de l'« atlas produit » dont les cartes sont les :

$$(x \times y : U_x \times U_y \rightarrow \underline{U}_x \times \underline{U}_y)_{x \in \mathcal{A}_M, y \in \mathcal{A}_N}.$$

Évidemment, $\dim M \times N = \dim M + \dim N$. On définit de même le produit d'un nombre fini quelconque de variétés.

Additif : Variétés analytiques complexes

En remplaçant partout \mathbb{R}^m par \mathbb{C}^m , et en exigeant des changements de carte qu'ils soient *analytiques complexes*, on obtient la catégorie des **variétés analytiques complexes**. Un exemple important est l'**espace projectif complexe** $\mathbb{C}P^m$, défini comme l'ensemble des droites vectorielles complexes de \mathbb{C}^{m+1} (sous- \mathbb{C} -espaces vectoriels à une dimension de \mathbb{C}^{m+1}), muni de l'atlas analytique complexe défini comme en 1.2. La **droite projective complexe** $\mathbb{C}P^1$, encore appelée **sphère de Riemann**, joue un rôle important en théorie des fonctions d'une variable complexe ; c'est une variété analytique complexe à une dimension, qu'on peut identifier à \mathbb{C} complété par l'ajout d'un « point à l'infini » ; on peut aussi la considérer comme une variété analytique réelle à deux dimensions, et en tant que telle elle est difféomorphe à la sphère S^2 (cf. étude 4^{bis}).

De la même façon, le tore $\mathbb{T}^2 = S^1 \times S^1$ peut être considéré soit comme une variété analytique réelle à deux dimensions, soit comme une variété analytique complexe à une dimension (quotient de \mathbb{C} par un réseau Ω) ; mais alors que la structure analytique réelle est définie de façon canonique, la structure analytique complexe ainsi définie sur \mathbb{T}^2 dépend du choix du réseau (à similitude directe près), et l'on obtient ainsi une *infinité continue de structures non isomorphes* : cf. étude 4^{ter}.

2.1. Exemples d'isomorphismes en géométrie projective

i) **Homographies de $\mathbb{P}(E)$** . Tout automorphisme linéaire d'un espace vectoriel E transforme les droites vectorielles en droites vectorielles, et définit donc une bijection de $\mathbb{P}(E)$ dans lui-même, qui est évidemment un isomorphisme analytique, appelé **homographie** de $\mathbb{P}(E)$.

ii) **Isomorphisme canonique entre la grassmannienne $G_p(E)$ et sa « duale » $G_q(\check{E})$** (E est un espace vectoriel de dimension $p+q$, et \check{E} l'espace vectoriel dual).

La donnée dans E d'un sous-espace vectoriel F de dimension p équivaut à la donnée dans \check{E} d'un sous-espace vectoriel F^\perp de dimension q , dit « espace orthogonal » à F , défini comme l'ensemble des formes linéaires sur E qui s'annulent sur F . On obtient ainsi une bijection canonique entre $G_p(E)$ et $G_q(\check{E})$, dont on vérifie immédiatement que c'est un difféomorphisme analytique : si $F \in G_p(E)$ est définie dans un système de coordonnées X de E (base de \check{E}) par la matrice $x(F)$, $F^\perp \in G_q(\check{E})$ sera définie dans le système de coordonnées dual par la matrice transposée.

En particulier, si E est un espace vectoriel de dimension $m+1$, on a un isomorphisme canonique entre $\mathbb{G}_m(E)$ et $\mathbb{P}(E)$: l'ensemble des hyperplans vectoriels de E s'identifie à l'espace projectif sur le dual de E (en d'autres termes, à un hyperplan de E on associe son équation, définie à multiplication par un scalaire près).

2.2. Sous-variétés et plongements

De même que nous avons défini au chapitre 1 la notion de sous-variété de \mathbb{R}^n , nous pouvons définir la notion de **sous-variété d'une variété** M : un ensemble $S \subset M$ est dit *lisse en a de dimension ℓ* (ou de codimension $p = m - \ell$, où $m = \dim M$) s'il existe une carte locale $x = (x_1, \dots, x_m)$ de M au voisinage de a qui « redresse » S , c'est-à-dire telle que S soit définie dans le domaine de cette carte par l'annulation de p coordonnées choisies parmi les x_i (disons, comme au chapitre 2, les p dernières $x_{\ell+1}, \dots, x_m$).

Une *sous-variété* de M est un sous-ensemble S lisse en chacun de ses points. La structure de variété de M induit alors sur S une structure de variété que l'on peut définir ainsi : on considère une famille de redressements locaux de S dont les domaines recouvrent S ; les changements de carte correspondants sont différentiables (ou analytiques, selon la catégorie dans laquelle on travaille), et leurs restrictions à $\mathbb{R}^\ell \times \{0\}$ sont à valeurs dans $\mathbb{R}^\ell \times \{0\}$; ces restrictions définissent donc un atlas différentiable (ou analytique) de S , dont la classe d'équivalence différentiable ou analytique ne dépend évidemment pas de la famille de redressements choisie (pourvu que cette famille soit compatible avec la structure donnée sur M).

Si S est une sous-variété de M , l'injection canonique $i_S : S \rightarrow M$ est évidemment un morphisme, appelé **plongement canonique de la sous-variété** S . Plus généralement, on appelle **plongement** d'une variété N dans une variété M un morphisme :

$$i_S \circ \alpha : N \xrightarrow{\alpha} S \xrightarrow{i_S} M,$$

composé d'un isomorphisme α et d'un plongement canonique i_S de sous-variété.

Exemple de sous-variété Le graphe d'un morphisme $f : M \rightarrow N$ est évidemment une sous-variété de la variété produit $M \times N$.

Morphismes de sous-variétés Tout ce qui a été dit au paragraphe 3 du chapitre 2 sur les morphismes de sous-variétés de \mathbb{R}^n se transpose sans changement au cas où les espaces ambiants sont eux-mêmes des variétés, à l'exception des versions *globales* des énoncés du § 3.3, dont la preuve utilisait des propriétés métriques de \mathbb{R}^n .

2.3. Encore un peu de géométrie projective

(N.B. : Ce qui suit est vrai aussi bien en géométrie projective complexe qu'en géométrie projective réelle.)

i) Sous-espaces projectifs. Soit $F \subset E$ une inclusion d'espaces vectoriels de dimensions finies ($\dim E = m + 1$, $\dim F = \ell + 1$). Elle induit une inclusion évidente $\mathbb{P}(F) \subset \mathbb{P}(E)$, qui fait de $\mathbb{P}(F)$ une sous-variété de $\mathbb{P}(E)$.

En effet, en appliquant la construction 1.2 à un système de coordonnées linéaires (X_0, X_1, \dots, X_m) tel que $X_{i+\ell}, \dots, X_m$ soit un système d'équations de F dans E , on obtient un redressement x de $\mathbb{P}(F)$ sur l'ouvert U_x des droites de E non contenues

dans l'hyperplan $\{X_0 = 0\}$; or toute droite $D \subset F$ appartient à un tel ouvert U_X (il suffit de choisir pour X_0 une fonction linéaire telle que $X_0|_D \neq 0$).

Une telle sous-variété d'un espace projectif est appelée **sous-espace projectif**. Les sous-espaces projectifs de codimension 1 ($\ell = m - 1$) sont appelés **hyperplans projectifs**. Par ailleurs, on appelle **droite projective** [resp. **plan projectif**] tout espace projectif de dimension 1 [resp. 2], qu'il soit ou non plongé dans un espace projectif de dimension supérieure.

ii) **Lien entre géométrie affine et géométrie projective.** Nous nous proposons ici de démontrer que le complémentaire d'un hyperplan projectif est muni canoniquement d'une structure d'espace affine.

Soit F un hyperplan vectoriel (X_0, X_1, \dots, X_m) tel que F soit donné par l'équation $X_0 = 0$. La carte x associée à ce système de coordonnées par la construction 1.2 a pour domaine $\mathbb{P}(E) \setminus \mathbb{P}(F)$, qu'elle envoie isomorphiquement sur \mathbb{R}^m . Bien que cet isomorphisme dépende du choix du système de coordonnées X , il est facile de voir qu'un changement de coordonnées linéaires sur E , laissant F invariant, a pour effet de transformer $x = (x_1, \dots, x_m)$ par une transformation linéaire affine. La structure affine de $\mathbb{P}(E) \setminus \mathbb{P}(F)$ induite par celle de \mathbb{R}^m via l'isomorphisme x ne dépend donc pas du choix de X .

Exemple Dans le cas de \mathbb{P}^1 la carte x est simplement l'application qui, à une droite vectorielle « non verticale » de \mathbb{R}^2 , associe sa pente, et l'« hyperplan projectif » $\mathbb{P}(F)$ se réduit à un seul point (la droite verticale).

En étendant la notion de pente à tout \mathbb{P}^1 par la convention que la droite verticale est « de pente infinie », on obtient ainsi une bijection canonique entre la droite projective \mathbb{P}^1 et $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$, droite réelle « complétée » par l'ajout d'un « point à l'infini ».

Remarque Plus généralement, on montre dans les cours de géométrie projective élémentaire comment associer *canoniquement* à tout espace affine \mathbb{A} à m dimensions une paire $(\overline{\mathbb{A}}, \mathbb{A}_\infty)$ telle que $\mathbb{A} = \overline{\mathbb{A}} \setminus \mathbb{A}_\infty$, où $\overline{\mathbb{A}}$ est un espace projectif à m dimensions de \mathbb{A}_∞ un hyperplan projectif appelé « hyperplan de l'infini ».

Exemple : **Complétion d'une droite affine \mathbb{A} en une droite projective $\overline{\mathbb{A}} = \mathbb{A} \cup \{\infty\}$.** Oubliant la géométrie projective pour ne nous intéresser qu'à la structure de variété, nous pouvons décrire la structure analytique de $\overline{\mathbb{A}}$ au moyen d'un atlas à deux cartes x, \bar{x} défini comme suit.

Soit $x : \mathbb{A} \xrightarrow{\cong} \mathbb{R}$ une coordonnée affine de \mathbb{A} , et soit a_0 « l'origine » de cette coordonnée affine (point de \mathbb{A} tel que $x(a_0) = 0$) ; en posant $\overline{\mathbb{A}} = \mathbb{A} \setminus \{a_0\}$, on définit $\bar{x} : \overline{\mathbb{A}} \xrightarrow{\cong} \mathbb{R}$ par $\bar{x}(a) = 1/x(a)$ si $a \neq \infty$, $\bar{x}(\infty) = 0$. Il est clair que la variété analytique $\tilde{\mathbb{A}}$ ainsi obtenue est isomorphe à \mathbb{P}^1 .

iii) **Variétés projectives.** Le lieu des zéros d'un polynôme homogène de $m+1$ variables X_0, X_1, \dots, X_m est évidemment invariant par homothéties de centre O . On appelle **cône algébrique** de \mathbb{R}^{m+1} un sous-ensemble de \mathbb{R}^{m+1} défini par l'annulation d'un ou plusieurs polynômes homogènes (de degrés éventuellement différents).

Étant invariant par homothéties, un cône est une union de droites vectorielles, les **génératrices** du cône.

L'ensemble de toutes les génératrices d'un cône algébrique est un sous-ensemble fermé de \mathbb{P}^m (la base du cône). Les sous-ensembles de \mathbb{P}^m obtenus de cette façon sont appelés **variétés projectives**.

Attention ! Nous sommes ici dans le contexte de la géométrie algébrique, où le mot « variété » est utilisé pour désigner des sous-ensembles d'un espace projectif (ou affine) définis par des équations polynomiales, sans se préoccuper du caractère lisse ou non de ces sous-ensembles : une « variété projective » n'est donc pas nécessairement une « variété » au sens du paragraphe 1 !

Cette fâcheuse confusion terminologique n'existe pas en anglais, où le mot consacré en géométrie algébrique est *variety*, à distinguer du mot *manifold*.

Exemple Soit f un polynôme homogène vérifiant la condition :

$$\partial_0 f, \dots, \partial_m f \text{ ne s'annulent qu'en } 0. \quad (*)$$

Alors le cône d'équation $f = 0$ est lisse en dehors de 0, et sa base S est une hypersurface lisse de \mathbb{P}^m .

Exercice Démontrer la dernière affirmation ; S est donnée dans la carte $x = (X_1/X_0, \dots, X_m/X_0)$ par l'équation $f(1, x_1, \dots, x_m) = 0$, et par des équations analogues dans les cartes qui s'en déduisent par permutation circulaire de (X_0, X_1, \dots, X_m) .

Dans la situation ci-dessus, le polynôme homogène f est appelé **équation homogène de l'hypersurface** S . Ainsi par exemple, l'hyperplan $\mathbb{P}(F)$ de ii) admet X_0 pour équation homogène.

Cônes quadratiques et hyperquadriques projectives Dans le cas où f est un polynôme homogène de degré 2 (*forme quadratique*), la condition (*) équivaut à dire que la forme quadratique est *non dégénérée*, c'est-à-dire que la matrice de cette forme quadratique est de rang maximum. Le lieu des zéros de f est alors appelé *cône quadratique*.

Un cône quadratique de \mathbb{R}^3 [resp. \mathbb{R}^4] a pour base une courbe lisse de \mathbb{P}^2 [resp. surface lisse de \mathbb{P}^3] appelée *conique projective* [resp. *quadrique projective*] etc. (en dimension supérieure on parlera d'« hyperquadrique »). En enlevant à \mathbb{P}^2 [resp. \mathbb{P}^3] une « droite de l'infini » [resp. un « plan à l'infini »] (cf. ii)) on en déduit les *coniques affines* usuelles, qui sont des courbes lisses du plan [resp. les *quadriques affines*, surfaces lisses de l'espace affine à trois dimensions] : les coniques affines réelles sont de trois types, respectivement appelées *hyperboles*, *paraboles* et *ellipses* selon que la conique projective correspondante coupe la droite de l'infini en deux points, un point (point de tangence) ou zéro point ; si l'on travaille sur \mathbb{C} au lieu de \mathbb{R} deux cas seulement sont à distinguer, selon que la conique projective est tangente ou non à la droite de l'infini (sur \mathbb{C} , un polynôme du second degré a toujours deux racines, distinctes ou confondues).

Dual d'un cône quadratique. Hyperquadrique duale Toute forme quadratique à $m+1$ variables peut s'écrire $Q(X) = \langle X, AX \rangle$, où A est une matrice symétrique et $\langle \cdot \rangle$ désigne le produit scalaire euclidien $\langle X, Y \rangle = \sum_{i=0}^m X_i Y_i$. Dire que la forme quadratique est

non dégénérée équivaut à dire que la matrice A est inversible. Le cône C d'équation $Q(X) = 0$ est alors lisse en tout point $X \neq 0$, et l'hyperplan tangent à C en X s'écrit :

$$\left\{ dX \in \mathbb{R}^{m+1} \mid \langle AX, dX \rangle = 0 \right\}.$$

Autrement dit, une forme linéaire Y sur \mathbb{R}^{m+1} est l'équation d'un hyperplan tangent au cône C si et seulement si elle est de la forme $Y = AX$ avec $X \in C$, c'est-à-dire si elle vérifie l'équation $\langle Y, A^{-1}Y \rangle = 0$.

Conclusion L'ensemble des hyperplans tangents au cône quadratique :

$$C = \left\{ X \in \mathbb{R}^{m+1} \mid \langle X, AX \rangle = 0 \right\}$$

est le cône quadratique :

$$\check{C} = \left\{ Y \in \mathbb{R}^{m+1} \mid \langle Y, A^{-1}Y \rangle = 0 \right\}$$

dit *cône dual* de C .

En remplaçant les cônes par leurs bases on obtient la proposition suivante.

PROPOSITION L'ensemble des hyperplans tangents à une hyperquadrique projective est une hyperquadrique projective, dite duale de la précédente.

La formulation de cette proposition a utilisé la notion d'*hyperplan tangent à une hypersurface projective lisse* $S \subset \mathbb{P}^m$, notion que l'on peut définir comme suit.

Soit C le cône de base S ; ce cône est lisse en tout point $a \neq 0$, et l'hyperplan tangent à C en a est un hyperplan vectoriel de \mathbb{R}^{m+1} qui ne dépend que de la génératrice \hat{a} du cône à laquelle a appartient; à cet hyperplan $F_{\hat{a}}$ est donc associé dans \mathbb{P}^m un hyperplan projectif $\mathbb{P}(F_{\hat{a}})$, qui est par définition l'*hyperplan tangent à S en \hat{a}* .

Le cas où S est une *conique projective* est particulièrement simple : une droite du plan projectif est tangente à S si et seulement si elle coupe S en un seul point.

ÉTUDE 4

Un peu de géométrie plane

La droite projective réelle \mathbb{P}^1 est analytiquement difféomorphe au cercle S^1

La figure 3 suggère une bijection $D \leftrightarrow a$ entre \mathbb{P}^1 et S^1 (D est tangente au cercle si et seulement si le point a est en O). Pour montrer que cette bijection est un difféomorphisme analytique il est instructif de procéder de deux façons.

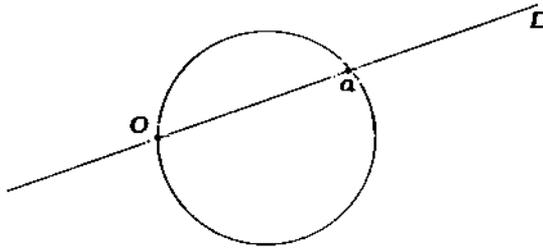


Figure 3.

■ **Première méthode** : on repère les points de S^1 par leurs « coordonnées ambiantes » (X_0, X_1) et ceux de $\mathbb{P}^1 = \overline{\mathbb{R}}$ au moyen de l'une ou l'autre des cartes x, \bar{x} de 2.3 ii) (Fig. 4). Les détails sont laissés au lecteur.

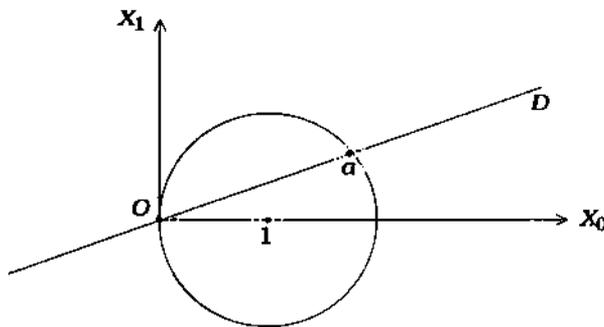


Figure 4.

■ **Deuxième méthode** : on repère les points de S^1 par une « coordonnée angulaire » θ (chap. 2, exercice 3)a), et ceux de \mathbb{P}^1 par la coordonnée $\theta/2$ (Fig. 5).

On fait ainsi le lien avec une propriété bien connue de géométrie plane : les angles orientés de demi-droites sont définis modulo 2π alors que les angles orientés de droites sont définis modulo π .

Le difféomorphisme considéré entre \mathbb{P}^1 et S^1 peut ainsi être vu « de façon intrinsèque » (c'est-à-dire en oubliant le plongement de S^1 dans \mathbb{R}^2) comme le composé des difféomorphismes évidents.

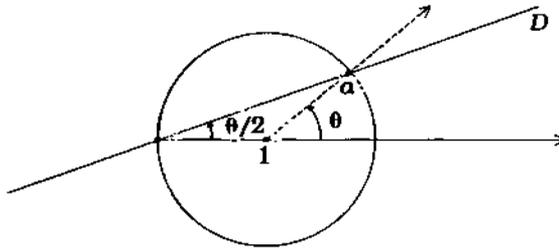


Figure 5.

$$\begin{array}{ccc}
 \mathbb{P}^1 & \longrightarrow & S^1 \\
 \downarrow \mathbb{R} & & \downarrow \mathbb{R} \\
 \mathbb{R}/\pi\mathbb{Z} & \xrightarrow{\text{multiplication par 2}} & \mathbb{R}/2\pi\mathbb{Z}
 \end{array}$$

EXERCICES SUR LE THÈME D'ÉTUDE 4

EXERCICE 1.— Généraliser la construction illustrée par la figure 3 au cas d'une conique de \mathbb{P}^2 ($\mathbb{R}\mathbb{P}^2$ ou $\mathbb{C}\mathbb{P}^2$).

EXERCICE 1*.— Expliciter la construction *duale* de 1) (pour la notion de dualité dans les espaces projectifs, cf. 2.1 ii) et 2.3 iii), proposition finale).

EXERCICE 2.—

- a) Montrer que la variété des droites du plan affine (exemple 1 de l'introduction) s'identifie à un ouvert (que l'on précisera) du plan projectif.
- b) Plus généralement, montrer que l'ensemble des p -plans affines de \mathbb{R}^n s'identifie à un ouvert (que l'on précisera) d'une variété grassmannienne à préciser.

EXERCICE 3.—

- a) Soit P un plan affine ou projectif. On appelle **élément de contact** dans P la donnée d'un couple (a, D) , où a est un point de P et D une droite de P passant par a . Montrer que l'ensemble des éléments de contact de P est une sous-variété de $P \times \check{P}$, produit de P par la variété \check{P} des droites de P .
- b) Généraliser en dimension supérieure : P est un espace affine ou projectif (de dimension n) ; un « élément de contact » dans P est un couple constitué par un point a de P et un hyperplan H (affine ou projectif) passant par a .

ÉTUDE 4bis

La sphère de Riemann $S \approx \mathbb{C}P^1$

En dimension 1, le lien entre géométrie affine et géométrie projective (§ 2.3 ii) s'énonce ainsi : l'espace projectif à une dimension peut être identifié à un espace affine à une dimension complété par l'ajout d'un « point à l'infini ».

Le cas de $\mathbb{R}P^1$ a été explicité en 2.3 ii). Le cas de $\mathbb{C}P^1$ donne une variété analytique complexe à une dimension appelée *sphère de Riemann*, que nous noterons S . La dénomination « sphère » provient de ce qu'en tant que variété différentielle (ou analytique réelle) elle est difféomorphe à la sphère usuelle S^2 .

NOTATIONS $\mathbb{C}P^1$, muni des coordonnées homogènes $(z_0 ; z_1)$, est muni de l'atlas à deux cartes $s_A : A \xrightarrow{\cong} \mathbb{C}$, $s_{\tilde{A}} : \tilde{A} \xrightarrow{\cong} \mathbb{C}$, où :

$$\begin{aligned} A &= \{(z_0 : z_1) \mid z_0 \neq 0\}, & s_A(z_0 : z_1) &= z_1/z_0 \\ \tilde{A} &= \{(z_0 : z_1) \mid z_1 \neq 0\}, & s_{\tilde{A}}(z_0 : z_1) &= z_0/z_1. \end{aligned}$$

Le changement de carte est l'involution σ^* de \mathbb{C}^* qui à z associe $1/z$.

Soit S l'ensemble constitué par \mathbb{C} auquel on adjoit un point noté ∞ , et soit $s : \mathbb{P}^1\mathbb{C} \rightarrow S$ la bijection définie par $s|_A = s_A$, $s(0 : 1) = \infty$. Cette bijection transforme A en \mathbb{C} et \tilde{A} en $\tilde{\mathbb{C}} := \mathbb{C}^* \cup \{\infty\}$. Par composition avec s^{-1} les cartes s_A et $s_{\tilde{A}}$ de $\mathbb{C}P^1$ deviennent deux cartes de S :

$$\begin{array}{ccc} \mathbb{C} & \subset S \supset & \tilde{\mathbb{C}} \\ \downarrow \mathbb{I}_{\mathbb{C}} & & \downarrow \sigma \\ \mathbb{C} & & \mathbb{C} \end{array}$$

où σ coïncide sur \mathbb{C}^* avec l'involution σ^* (c'est-à-dire $z \mapsto 1/z$), et $\sigma(\infty) = 0$.

Nous avons ainsi « transporté sur S » la structure analytique complexe de $\mathbb{C}P^1$.

Remarque sur la topologie de S

- i) Les voisinages ouverts du point à l'infini sont les complémentaires dans S des compacts de \mathbb{C} .
- ii) S est compacte, car c'est l'union de deux compacts \bar{D} et $\sigma^{-1}(\bar{D})$, où \bar{D} désigne le disque unité fermé de \mathbb{C} .

Dans la description qui suit, où S sera identifiée à la sphère usuelle S^2 , ces deux compacts seront les deux hémisphères (sud et nord respectivement).

Construction d'un difféomorphisme (analytique réel) $s : S^2 \rightarrow S$
 S^2 est la sphère usuelle :

$$S^2 = \{(x_0, x_1, x_2) \in \mathbb{R}^3 \mid x_0^2 + x_1^2 + x_2^2 = 1\}.$$

La projection stéréographique nord définit un difféomorphisme :

$$\begin{aligned} s_+ : S^2 \setminus (a_+) &\rightarrow \mathbb{R}^2 = \mathbb{C} \\ (x_0, x_1, x_2) &\mapsto (x_1 + ix_2)/(1 - x_0) \end{aligned}$$

où l'on a noté $a_+ = (0, 0, 1)$ le « pôle nord » (cf. exercice 3)d) du chapitre 2).

En complétant par la convention $s(a_+) = \infty$ on obtient ainsi une bijection $s : \mathbb{S}^2 \rightarrow \mathbb{S}$, dont on sait déjà que c'est un difféomorphisme en restriction à $\mathbb{S}^2 \setminus \{a_+\}$. Reste à montrer que c'est un difféomorphisme au voisinage de a_0 , et par exemple que s envoie $\mathbb{S}^2 \setminus \{a_-\}$ difféomorphiquement sur $\tilde{\mathbb{C}}$ (on a noté $a_- = (0, 0, -1)$ le « pôle sud ») : pour cela on utilise comme carte au but la carte σ , et comme carte à la source la projection stéréographique sud :

$$s_- : (x_0, x_1, x_2) \mapsto (x_1 + ix_2) / (1 + x_0).$$

Dans ces cartes on vérifie que s s'exprime par l'application $z \mapsto \bar{z}$, qui est évidemment un automorphisme analytique réel de \mathbb{R}^2 .

$$\begin{array}{ccc} \mathbb{S}^2 & \xrightarrow{s} & \mathbb{S} \\ \cup & & \cup \\ \mathbb{S}^2 - \{a\} & \longrightarrow & \tilde{\mathbb{C}} \\ \downarrow s_- & & \downarrow \sigma \\ \mathbb{C} & \longrightarrow & \mathbb{C} \\ z & \mapsto & \bar{z} \end{array}$$

PROBLÈME

Morphismes de la sphère de Riemann dans elle-même

Le but de ce problème, en deux parties, est de montrer que les morphismes de \mathbb{S} dans \mathbb{S} s'identifient aux fonctions rationnelles sur \mathbb{C} . On en déduira comme corollaire que les seuls automorphismes analytiques complexes de \mathbb{S} sont les homographies (2.1 i), et que les seuls automorphismes analytiques complexes de \mathbb{C} sont les transformations affines $z \mapsto az + b$.

Exercice préparatoire

- 1) Soit $f : \mathbb{C} \rightarrow \mathbb{C}$ une fonction polynomiale complexe de degré $n \geq 1$.
Montrer qu'en étendant f par la convention $\check{f}(\infty) = \infty$ on obtient un morphisme \check{f} de \mathbb{S} dans \mathbb{S} .
- 2) Par quelle convention peut-on étendre une fraction rationnelle f en un morphisme \check{f} de \mathbb{S} dans \mathbb{S} ?
Sous quelles hypothèses sur f ce morphisme enverra-t-il ∞ sur ∞ ?
- 3) Sous quelles hypothèses sur la fraction rationnelle f le morphisme \check{f} de 2), envoyant ∞ sur ∞ , sera-t-il un difféomorphisme local au voisinage du point à l'infini ?

A. Les morphismes de \mathbb{S} dans \mathbb{S} s'identifient aux fonctions rationnelles sur \mathbb{C} .

- 0) Montrer que les seules fonctions holomorphes sur \mathbb{S} (morphismes de \mathbb{S} dans \mathbb{C}) sont les constantes [cf. « Rappels » ci-après, propriété iii)].
- 1) Soit U un ouvert de \mathbb{C} . Montrer que les morphismes \check{f} de U dans \mathbb{S} s'identifient aux fonctions méromorphes sur U , étendues par la convention $\check{f}(a) = \infty$ si a est un pôle (cf. « Rappels » ci-après, propriété i)).
- 2) Par extension de 1), on appellera fonction méromorphe sur un ouvert U de \mathbb{S} tout morphisme $\check{f} : U \rightarrow \mathbb{S}$. En supposant que U est le domaine d'une carte φ , montrer que \check{f} est méromorphe sur U si et seulement si $\check{f} \circ \varphi^{-1}$ est méromorphe sur $\underline{U} = \varphi(U)$.
- 3) Montrer qu'une fonction rationnelle (quotient de deux polynômes d'une variable complexe) peut être considérée comme la restriction d'une unique fonction méromorphe \check{f} sur \mathbb{S} ; expliciter le nombre $\check{f}(\infty)$, et l'ensemble $\check{f}^{-1}(\infty)$. (N.B : cette question n'est qu'une reformulation de la question 2) de l'exercice préparatoire).
- 4) Soit \check{f} un morphisme de \mathbb{S} dans \mathbb{S} . Montrer que l'ensemble $\check{f}^{-1}(0)$ des zéros de \check{f} , et l'ensemble $\check{f}^{-1}(\infty)$ des pôles de \check{f} , sont des ensembles finis [cf. « Rappels », propriété i)].

En déduire que $f := \check{f}|_{\mathbb{C}}$ est le produit d'une fonction rationnelle par une fonction sans zéros ni pôles sur \mathbb{C} , méromorphe au voisinage de l'infini.

- 5) Montrer qu'une fonction \check{f} méromorphe sur \mathbb{S} , et qui n'a ni zéros ni pôles sur \mathbb{C} , est une constante (si $\check{f}(\infty) \neq \infty$ on appliquera 0 ; le cas $\check{f}(\infty) = \infty$ se ramène à l'autre en raisonnant sur $\sigma \circ \check{f}$ (où σ est une carte d'un voisinage de ∞ dans \mathbb{S}). Conclure.
- 6) De la conclusion ainsi obtenue déduire le corollaire : les seuls automorphismes de \mathbb{S} (difféomorphismes analytiques complexes de \mathbb{S} dans elle-même) sont ceux définis par les fonctions rationnelles $z \mapsto \frac{az+b}{cz+d}$ ($ad - bc \neq 0$), c'est-à-dire (via l'identification $\mathbb{S} = \mathbb{P}^2\mathbb{C}$) les homographies de $\mathbb{P}^1\mathbb{C}$.

B. Application aux « automorphismes » de \mathbb{C} : les seuls difféomorphismes analytiques complexes de \mathbb{C} dans lui-même sont les transformations affines $z \mapsto az + b$ ($a \neq 0$).

- 1) Montrer que tout homéomorphisme h de \mathbb{C} dans lui-même s'étend de façon unique en un homéomorphisme \check{h} de \mathbb{S} dans elle-même, tel que $\check{h}(\infty) = \infty$.
- 2) Montrer que si h est un difféomorphisme analytique complexe il en est de même de \check{h} (le représenter dans une carte au voisinage de l'infini, et appliquer la propriété iv) des Rappels ci-après).
- 3) Conclure à l'aide de A. 6).

APPENDICE

Rappels sur les fonctions holomorphes

Le problème ci-dessus a fait usage des propriétés suivantes des fonctions holomorphes d'une variable complexe :

- o) Si f est holomorphe non nulle dans un ouvert U , son inverse $1/f$ est holomorphe dans U .
- i) Toute fonction holomorphe au voisinage de $a \in \mathbb{C}$ s'écrit localement $f(z) = (z - a)^\nu h(z)$, où $\nu \in \mathbb{N}$ avec h holomorphe non nulle au voisinage de a ; si $\nu \geq 1$ on dit que f admet en a un zéro d'ordre ν ; la fonction $1/f$, holomorphe d'après 0) sur un voisinage épointé de a , est dite méromorphe au voisinage de a , et le point est appelé pôle d'ordre ν de cette fonction.
- ii) Principe du prolongement analytique : une fonction holomorphe sur un ouvert connexe est déterminée sans ambiguïté par sa restriction à n'importe quel ouvert plus petit.
- iii) Une fonction holomorphe et bornée sur tout \mathbb{C} est une constante.
- iv) Théorème des singularités inexistantes de Riemann : une fonction holomorphe bornée sur un disque épointé D^* s'étend en fonction holomorphe sur tout le disque D .

Remarque Les propriétés o), i) et ii) sont vraies pour les fonctions analytiques en général, qu'il s'agisse du cas analytique réel ou analytique complexe (holomorphe). Les propriétés iii) et iv) sont spécifiques du cas complexe, et peuvent se démontrer (facilement) à l'aide d'intégrales de Cauchy.

ÉTUDE 4^{ter}Structures complexes sur le tore \mathbb{T}^2

On appelle **tore standard à n dimensions**, et l'on note \mathbb{T}^n , la variété produit de n exemplaires du cercle trigonométrique S^1 .

Le plongement canonique de S^1 dans \mathbb{R}^2 permet de voir $\mathbb{T}^n = S^1 \times \dots \times S^1$ comme une sous-variété de \mathbb{R}^{2n} . Par ailleurs le tore à deux dimensions \mathbb{T}^2 peut être vu aussi comme une surface de l'espace usuel \mathbb{R}^3 (cf. exercice II 3)b) et schémas de l'étude 2).

Nous allons ici adopter un autre point de vue, basé sur la notion de *coordonnée angulaire* sur le cercle : munissant chaque facteur S^1 de \mathbb{T}^n de son atlas de coordonnées angulaires (cf. chap 2, exercice 3)a)), on obtient sur \mathbb{T}^n un atlas de « systèmes de coordonnées angulaires » dont les cartes sont des n -uplets de fonctions $(\theta_1, \dots, \theta_n)$ définis sur divers produits d'arcs de cercles, n -uplets ne différant les uns des autres (sur les composantes connexes des intersections de leurs domaines de définition) que par des translations $\theta_i \mapsto \theta_i + 2\pi k_i$ ($k_i \in \mathbb{Z}$). Autrement dit, on a un isomorphisme (analytique) canonique :

$$\mathbb{T}^n = \mathbb{R}^n / 2\pi\mathbb{Z}^n \quad (= (\mathbb{R}/2\pi\mathbb{Z})^n),$$

où la structure de variété analytique de l'ensemble quotient écrit à droite est définie par l'atlas ci-dessus.

1 Soit A un espace affine réel à n dimensions, et soit Ω un réseau de translations de A (sous-groupe du groupe des translations de A engendré par n vecteurs linéairement indépendants). Alors il existe sur le tore A/Ω (quotient de A par l'action de Ω) une et une seule structure de variété analytique telle que l'application de passage au quotient $\pi : A \rightarrow A/\Omega$ soit un difféomorphisme (analytique) local.

Preuve

La condition de *difféomorphisme local* signifie que tout $a \in A$ admet un voisinage U tel que $\pi_U = \pi|_U : U \rightarrow \tilde{U}$ soit un difféomorphisme (on a posé $\tilde{U} = \pi(U)$).

Dans notre cas, il est clair que A peut être recouvert par une famille d'ouverts U_α tels que π_{U_α} soit bijectif (par exemple, on pourra prendre pour U_α des « parallélogrammes » de côtés plus courts que les mailles du réseau).

Comme la propriété d'être différentiable est une propriété locale, dire que π est un difféomorphisme local équivaut donc à imposer à ces bijections π_α d'être des difféomorphismes.

Soit alors $x : A \xrightarrow{m} \mathbb{R}^n$ un isomorphisme affine, dont il sera commode de supposer qu'il envoie une base de Ω sur la base canonique de \mathbb{R}^n . Par restriction à U_α , le difféomorphisme x induit un difféomorphisme $x_\alpha : U_\alpha \rightarrow \underline{U}_\alpha$ (ouvert de \mathbb{R}^n), et exiger que π_α soit un difféomorphisme équivaut à exiger la même chose de la bijection \tilde{x}_α définie par le diagramme commutatif suivant :

$$\begin{array}{ccc} U_\alpha & \xrightarrow{\pi_\alpha} & \tilde{U}_\alpha \\ x_\alpha \searrow & & \swarrow \tilde{x}_\alpha \\ & \underline{U}_\alpha & \end{array}$$

On conclut en remarquant que les bijections \tilde{x}_α ne diffèrent les unes des autres (dans les composantes connexes des $\tilde{U}_\alpha \cap \tilde{U}_{\alpha'}$) que par des translations (qui appartiennent à $2\pi\mathbb{Z}^n$) ; elles forment donc un atlas analytique de A/Ω , qui a toutes les propriétés voulues. \square

REMARQUE La démonstration précédente fournit « en prime » un difféomorphisme analytique entre A/Ω et le tore standard $\mathbb{T}^n = \mathbb{R}^n/2\pi\mathbb{Z}^n$.

2 Soit Ω un réseau de $\mathbb{C} = \mathbb{R} \oplus i\mathbb{R}$. Comme les translations de \mathbb{C} sont évidemment des difféomorphismes analytiques complexes, l'atlas de \mathbb{C}/Ω construit par le procédé 1 n'est pas seulement un atlas analytique réel (de variété à deux dimensions) : c'est un atlas analytique complexe, qui définit sur \mathbb{C}/Ω une structure de variété analytique complexe à une dimension, telle que $\pi : \mathbb{C} \rightarrow \mathbb{C}/\Omega$ soit un difféomorphisme local de variétés analytiques complexes.

3 Soient Ω et Ω' deux réseaux de \mathbb{C} se déduisant l'un de l'autre par une similitude directe. Alors les variétés analytiques complexes \mathbb{C}/Ω et \mathbb{C}/Ω' sont isomorphes.

Preuve

On sait que les similitudes directes du plan envoyant l'origine sur l'origine s'identifient aux automorphismes \mathbb{C} -linéaires de \mathbb{C} :

$$\begin{array}{ccc} \mathbb{C} & \xrightarrow{\times a} & \mathbb{C} \\ z & \mapsto & az \end{array} \quad (a \in \mathbb{C}^*).$$

Dire que deux réseaux Ω, Ω' sont directement semblables c'est donc dire qu'il existe un tel automorphisme envoyant Ω sur Ω' , qui induit donc (par passage au quotient) une unique bijection \tilde{a} rendant commutatif le diagramme :

$$\begin{array}{ccc} \mathbb{C} & \xrightarrow{\times a} & \mathbb{C} \\ \pi_\Omega \downarrow & & \downarrow \pi_{\Omega'} \\ \mathbb{C}/\Omega & \xrightarrow{\tilde{a}} & \mathbb{C}/\Omega' . \end{array}$$

Du fait que π_Ω et $\pi_{\Omega'}$ sont des difféomorphismes (analytiques complexes) locaux, on déduit que la bijection \tilde{a} est, comme $\times a$, un difféomorphisme analytique complexe. \square

3* Réciproquement, la condition de similitude directe de Ω et Ω' est nécessaire pour que \mathbb{C}/Ω et \mathbb{C}/Ω' soient isomorphes en tant que variétés analytiques complexes.

La clef de la démonstration est le résultat de l'étude 4^{bis} (problème B) sur les automorphismes analytiques de \mathbb{C} , complété par un argument topologique que nous donnerons à la fin de l'étude 7.

4 Tout réseau Ω de \mathbb{C} est directement semblable à un réseau de la forme $\mathbb{Z} + \tau\mathbb{Z}$, où $\tau \in \mathbb{C}$, $\text{Im}\tau > 0$ (pourquoi ?).

À quelle condition sur τ, τ' les deux réseaux $\mathbb{Z} + \tau\mathbb{Z}, \mathbb{Z} + \tau'\mathbb{Z}$ sont-ils directement semblables ? En déduire l'existence sur le tore \mathbb{T}^2 d'une infinité continue de structures analytiques complexes non isomorphes.

Remarquons que, d'après 1, les structures analytiques réelles sous-jacentes à ces structures analytiques complexes sont, quant à elles, isomorphes.

Tangence

La grande idée du calcul différentiel est d'approcher localement les applications par des applications linéaires. Le but de ce chapitre est d'apprendre à le faire pour les morphismes de variétés différentielles, en définissant l'*application linéaire tangente* à un morphisme de variétés différentielles pointées :

$$f : M, a \cdots \rightarrow N, b$$

En fait, le morphisme n'a besoin d'être donné que dans un voisinage de a (ce que symbolise notre notation $\cdots \rightarrow$), et l'application linéaire tangente ne change pas si l'on remplace f par sa restriction à un (couple de) voisinage(s) plus petit(s)¹.

Après avoir défini l'application linéaire tangente $T_a f$, avec les espaces vectoriels tangents $T_a M$ et $T_a N$ qui en sont la source et le but, nous regarderons comment l'espace tangent $T_a M$ varie quand a parcourt M . Nous obtiendrons ainsi le *fibré tangent* (avec son dual, le *fibré cotangent*), ce qui nous permettra de parler des *champs de vecteurs* et des *formes différentielles de degré 1*.

¹ Pour $f_U^V : U, a \rightarrow V, b$ j'appelle « restriction de f_U^V au couple d'ouverts (\tilde{U}, \tilde{V}) » (avec $a \in \tilde{U} \subset U, b \in \tilde{V} \subset V$) une application $f_{\tilde{U}}^{\tilde{V}}$ rendant commutatif le diagramme :

$$\begin{array}{ccc} U, a & \xrightarrow{f_U^V} & V, b \\ \cup & & \cup \\ \tilde{U}, a & \xrightarrow{f_{\tilde{U}}^{\tilde{V}}} & \tilde{V}, b. \end{array}$$

Comme les propriétés que nous allons étudier sont invariantes par restriction, il est naturel de considérer comme équivalentes des applications $f_1 : U_1, a \rightarrow V_1, b$ et $f_2 : U_2, a \rightarrow V_2, b$ admettant une restriction commune $\tilde{f} : \tilde{U}, a \rightarrow \tilde{V}, b$ ($\tilde{U} \subset U_1 \cap U_2, \tilde{V} \subset V_1 \cap V_2$). Deux telles applications sont dites définir le même *germe* $f : M, a \cdots \rightarrow N, b$. Interpréter notre notation $\cdots \rightarrow$ comme désignant un « germe » (classe d'équivalence pour la relation d'équivalence ci-dessus) est un moyen de rassurer les puristes qui craindraient que le flou de cette notation ne cache un manque de rigueur.

0. Tangence de deux applications, jets d'applications

Soient (M, a) et (N, b) deux variétés différentielles pointées. Le but de ce paragraphe est de définir la notion de *tangence à l'ordre r* de deux (germes d')applications f et $\tilde{f} : M, a \cdots \rightarrow N, b$, où r est un entier ≥ 1 (en fait nous nous limiterons bientôt au cas $r = 1$, mais le cas r quelconque sera l'occasion d'utiles rappels sur les développements limités). Il s'agira d'une relation d'équivalence, que nous noterons $\overset{r}{\underset{a}{\sim}}$. Elle jouira de la propriété fondamentale de *compatibilité avec la composition* : pour g et $\tilde{g} : N, b \cdots \rightarrow P, c$, on a :

$$f \overset{r}{\underset{a}{\sim}} \tilde{f} \text{ et } g \overset{r}{\underset{b}{\sim}} \tilde{g} \Rightarrow g \circ f \overset{r}{\underset{a}{\sim}} \tilde{g} \circ \tilde{f}.$$

0.0. Cas où $M = \mathbb{R}^m$, $N = \mathbb{R}^n$

En munissant \mathbb{R}^m et \mathbb{R}^n de leurs normes d'espaces vectoriels (toutes sont équivalentes), on conviendra que :

$$f \overset{r}{\underset{a}{\sim}} \tilde{f} \text{ si } \|f(x) - \tilde{f}(x)\| = o(\|x - a\|^r).$$

C'est évidemment une relation d'équivalence. Vérifier qu'elle est compatible avec la composition est un exercice élémentaire de calcul infinitésimal, qui est à la base de la théorie usuelle des *développements limités* : on dit qu'une fonction admet un « développement limité à l'ordre r au voisinage du point a » si elle est tangente à l'ordre r en a à un polynôme, de degré total $\leq r$, en les variables $x_1 - a_1, \dots, x_m - a_m$; ce polynôme est alors unique, et le *théorème de Taylor-Young* garantit son existence si f est différentiable en a jusqu'à l'ordre r inclus (N.B. : si f n'est pas une fonction numérique mais une application à valeurs dans \mathbb{R}^n , $n > 1$, on écrit $f = (f_1, \dots, f_n)$ et le « développement limité » de f est le n -uplet constitué par les développements limités des f_i).

La *compatibilité de $\overset{r}{\underset{a}{\sim}}$ avec la composition* est la propriété qui permet de substituer un développement limité dans un autre. Notons bien que cette opération de « substitution d'un développement limité dans un autre » (comme celle de « produit de deux développements limités » dans le cas des fonctions numériques) ne s'identifie pas à l'opération correspondante sur les polynômes mais à cette opération *simplifiée par l'oubli des termes de degré $> r$* : le calcul usuel sur les développements limités n'est donc pas un calcul dans l'anneau des polynômes mais dans l'espace quotient de celui-ci par l'idéal engendré par les monômes de degré $> r$.

0.1. Cas général : tangence à l'ordre r pour deux (germes de) morphismes de M, a vers N, b

Soit x [resp. y] une carte locale de M au voisinage de a [resp. de N au voisinage de b]. Deux morphismes f, \tilde{f} de M, a vers N, b (définis au voisinage de a) sont dits *tangents à l'ordre r en a* ($f \overset{r}{\underset{a}{\sim}} \tilde{f}$) si leurs expressions $y \circ f \circ x^{-1}$ et $y \circ \tilde{f} \circ x^{-1}$ sont tangentes à l'ordre r au point $x(a)$, c'est-à-dire ont même développement limité à l'ordre r en

ce point. La relation $\overset{r}{\sim}_a$ ainsi définie est évidemment une relation d'équivalence. De plus, elle ne dépend pas du choix des cartes (la composition avec des changements de cartes n'affecte pas la relation $\overset{r}{\sim}_{x(a)}$).

Les classes d'équivalence pour cette relation s'appellent des **jets d'ordre r de M, a dans N, b** et l'on note $J_a^r(f)$ la classe de f (« jet d'ordre r de f en a »). Les jets d'ordre r sont donc l'analogie, pour les morphismes de variétés, des développements limités (auxquels on peut les identifier par le choix de cartes locales) ; la différence essentielle avec les développements limités est qu'on ne dispose pas dans chaque classe d'un représentant canonique qui serait l'analogie du polynôme de Taylor (avoir une expression polynomiale dans des cartes est une propriété qui dépend du choix des cartes !). Bien entendu, restreindre f à un voisinage plus petit ne change pas son jet en a .

0.2. L'espace des jets

L'ensemble des jets d'ordre r de M, a dans N, b sera noté $J_a^r(M, N)_b$. L'omission de l'indice a ou b (ou des deux) signifiera qu'on considère l'union des $J_a^r(M, N)_b$ pour toutes les valeurs possibles de a dans M (ou de b dans N , ou des deux).

Le fait que la relation d'équivalence $\overset{r}{\sim}_a$ soit compatible avec la composition permet de définir la **composition des jets** :

$$\begin{aligned} J_a^r(M, N)_b \times J_b^r(N, P)_c &\xrightarrow{\circ} J_a^r(M, P)_c \\ j = J_a^r(f) \ , \ j' = J_b^r(g) &\mapsto j' \circ j := J_a^r(g \circ f). \end{aligned} \tag{o}$$

EXERCICE Montrer que l'espace des jets est une variété différentielle, et que l'application (o) est un morphisme de variétés.

Tout l'exercice se ramène à la remarque évidente que la substitution d'un développement limité dans un autre donne un développement limité dont les coefficients dépendent *polynomialement* de ceux des développements limités initiaux (quelqu'un qui ignore tout des mathématiques sauf l'addition et la multiplication peut en principe comprendre l'algorithme de substitution d'un développement limité dans un autre !).

Les espaces de jets sont ainsi parmi les rares exemples de variétés différentielles munies d'atlas canoniques dont tous les changements de cartes sont polynomiaux.

0.3. L'espace vectoriel des jets d'ordre 1

$J_0^1(\mathbb{R}^m, \mathbb{R}^n)_0$ s'identifie canoniquement à l'espace $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ des applications linéaires de \mathbb{R}^m dans \mathbb{R}^n . De plus, l'application (o) du paragraphe 0.2 n'est autre dans ce cas que la composition des application linéaires :

$$\begin{aligned} \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n) \times \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p) &\xrightarrow{\circ} \mathcal{L}(\mathbb{R}^m, \mathbb{R}^p) \\ \ell \quad \quad \quad \ell' &\mapsto \ell' \circ \ell \end{aligned}$$

Elle est donc *bilinéaire*, et il en résulte que l'atlas canonique de $J_a^1(M, N)_b$ a pour changements de cartes des *automorphismes de l'espace vectoriel* $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$. L'identification (via une carte) de $J_a^1(M, N)_b$ à $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ munit ainsi $J_a^1(M, N)_b$ d'une structure

d'espace vectoriel à $m \times n$ dimensions qui ne dépend pas du choix de la carte. $J_a^1(M, N)_b$ est donc mieux qu'une variété : c'est un espace vectoriel à $m \times n$ dimensions. De plus, la composition des jets d'ordre 1 :

$$J_a^1(M, N)_b \times J_b^1(N, P)_c \xrightarrow{\circ} J_a^1(M, P)_c$$

est une application linéaire.

EXERCICE Montrer que ce qui précède n'est pas vrai pour les jets d'ordre $r > 1$; par exemple si Q est une forme quadratique et ℓ, ℓ' deux formes linéaires il n'est pas vrai que $Q(\ell + \ell') = Q(\ell) + Q(\ell')$ dans l'espace des développements limités d'ordre 2 !

1. Foncteurs tangent et cotangent

1.1. Foncteur tangent

On appelle **espace tangent à M en a** l'espace vectoriel à m dimensions $T_a M := J_0^1(\mathbb{R}, M)_a$.

Interprétation Les éléments de $T_a M$, appelés **vecteurs tangents à M en a** , sont les jets d'ordre 1 de courbes paramétrées $\lambda : \mathbb{R}, 0 \cdots \rightarrow M, a$; le jet $j_0^1(\lambda)$, que nous noterons encore $\lambda'(0)$, est un élément de $T_a M$ qu'il est naturel d'appeler **vecteur vitesse** au temps $t = 0$ du point $\lambda(t)$ mobile sur M .

Pour tout morphisme de variétés $f : M, a \cdots \rightarrow N, b$, la composition des jets permet de définir une application linéaire notée :

$$\begin{aligned} T_a f &: T_a M \rightarrow T_b N \\ v &\mapsto T_a f(v) := j_a^1(f) \circ v \end{aligned}$$

appelée **application tangente à f en a** : considérant v comme le vecteur vitesse d'une courbe paramétrée λ , $T_a f(v)$ est le vecteur vitesse de la courbe $f \circ \lambda$.

Évidemment,

$$T_a \mathbb{I}_M = \mathbb{I}_{T_a M}.$$

De plus, en composant f avec un morphisme :

$$g : N, b \cdots \rightarrow P, c,$$

on obtient un morphisme $g \circ f$ dont l'application linéaire tangente est :

$$T_a(g \circ f) = T_b g \circ T_a f$$

(c'est la propriété de composition des jets $j_a^1(g \circ f) = j_b^1(g) \circ j_a^1(f)$, agissant à gauche sur les vecteurs tangents).

Autrement dit (cf. chap. 2, § 4.2) la correspondance qui aux variétés pointées (et à leurs morphismes) associe leurs espaces tangents (et leurs applications linéaires tangentes) est un *foncteur covariant* de la catégorie des variétés pointées dans la catégorie des espaces vectoriels.

Coordonnées d'un vecteur tangent dans une carte Une conséquence évidente du caractère fonctoriel de T (« foncteur tangent ») est la suivante.

Si f est un isomorphisme (de variétés différentielles pointées), $T_a f$ est aussi un isomorphisme (d'espaces vectoriels).

En particulier, à toute carte locale :

$$x : M, a \cdots \rightarrow \mathbb{R}^m, 0$$

est associé un isomorphisme d'espaces vectoriels :

$$T_a x : T_a M \rightarrow T_0 \mathbb{R}^m = \mathbb{R}^m.$$

L'image par cet isomorphisme d'un vecteur $v \in T_a M$ est un m -uplet de nombres appelés *coordonnées du vecteur v dans la carte x* .

Remarque : translations de l'origine des temps

La « translation des temps » $t \mapsto t - t_0$ établit un isomorphisme canonique entre $J_{t_0}^1(\mathbb{R}, M)_a$ et $J_0^1(\mathbb{R}, M)_a = T_a M$. On peut donc encore considérer comme élément de $T_a M$ le vecteur vitesse au temps t_0 , noté $\lambda'(t_0)$, d'une courbe paramétrée $\lambda : \mathbb{R}, t_0 \cdots \rightarrow M, a$.

1.2*. Foncteur cotangent

On appelle **espace cotangent à M en a** l'espace vectoriel à m dimensions $T_a^* M := J_a^1(M, \mathbb{R})_0$.

Interprétation Les éléments de $T_a^* M$, appelés **vecteurs cotangents à M en a** , sont les jets d'ordre 1 de fonctions $\varphi : M, a \cdots \rightarrow \mathbb{R}, 0$; le jet $j_a^1(\varphi)$, que nous noterons encore $d\varphi_a$, est un élément de $T_a^* M$ qu'il est naturel d'appeler **différentielle de φ en a** .

Pour tout morphisme de variétés $f : M, a \cdots \rightarrow N, b$, la composition des jets permet de définir une application linéaire notée :

$$\begin{aligned} T_a^* f &: T_b^* N \rightarrow T_a^* M \\ \omega &\mapsto T_a^* f(\omega) := \omega \circ j_a^1(f), \end{aligned}$$

appelée **application cotangente à f en a** : considérant ω comme la différentielle en b d'une fonction φ sur N ($\omega = d\varphi_b$), $T_a^* f(\omega)$ est la différentielle en a de la fonction $\varphi \circ f$.

Évidemment,

$$T_a^* \mathbb{I}_M = \mathbb{I}_{T_a^* M}.$$

De plus, en composant f avec un morphisme :

$$g : N, b \cdots \rightarrow P, c$$

on obtient un morphisme $g \circ f$ dont l'application linéaire cotangente est :

$$T_a^*(g \circ f) = T_a^* g \circ T_b^* f$$

(Notez l'inversion des facteurs ! C'est toujours la propriété de composition des jets $J_a^1(g \circ f) = J_b^1(g) \circ J_a^1(f)$, mais agissant à droite sur les vecteurs cotangents.)

Autrement dit, la correspondance qui aux variétés pointées (et à leurs morphismes) associe leurs espaces cotangents (et leurs applications linéaires cotangentes) est un *foncteur contravariant*.

Coordonnées d'un vecteur cotangent dans une carte Comme les foncteurs covariants, les foncteurs contravariants transforment évidemment les isomorphismes en isomorphismes.

En particulier, à toute carte locale :

$$x : M, a \cdots \rightarrow \mathbb{R}^m, 0$$

est associé un isomorphisme d'espaces vectoriels :

$$T_a^* x : \mathbb{R}^m = T_0^* \mathbb{R}^m \rightarrow T_a^* M.$$

L'image réciproque d'un « covecteur » (vecteur cotangent) $\omega \in T_a^* M$ par cet isomorphisme est un m -uplet de nombres appelés *coordonnées du covecteur ω dans la carte x* .

Différentielle d'une fonction prenant en a une valeur quelconque La soustraction d'une constante c aux fonctions établit un isomorphisme canonique entre $J_a^1(M, \mathbb{R})_c$ et $J_a^1(M, \mathbb{R})_0 = T_a^* M$. On peut donc encore considérer comme élément de $T_a^* M$ la *différentielle en a d'une fonction φ prenant en a une valeur quelconque* : $d\varphi_a = J_a^1(\varphi - \varphi(a))$.

1.3. Dualité entre espace tangent et espace cotangent

Il résulte de 0.3 que la composition des jets d'ordre 1 :

$$\begin{array}{ccccc} J_0^1(\mathbb{R}, M)_a & \times & J_a^1(M, \mathbb{R})_0 & \xrightarrow{\circ} & J_0^1(\mathbb{R}, \mathbb{R})_0 \\ \parallel & & \parallel & & \parallel \\ T_a M & \times & T_a^* M & \xrightarrow{\circ} & \mathbb{R} \end{array}$$

est une forme bilinéaire sur le produit des espaces tangent et cotangent.

On va montrer que cette forme bilinéaire est non dégénérée, ce qui permet d'identifier l'espace cotangent au dual de l'espace tangent.

Explicitement, à $v = \lambda'(0)$ et $\omega = d\varphi_a$ (où $\mathbb{R}, 0 \cdots \xrightarrow{\lambda} M, a \cdots \xrightarrow{\varphi} \mathbb{R}, 0$), cette forme bilinéaire associe le nombre $\omega \circ v = (\varphi \circ \lambda)'(0)$, qui mérite d'être appelé, comme au chap. 1, § 1.1, *dérivée directionnelle de φ le long du vecteur v* (et qu'on pourra aussi noter $v\varphi$ comme au chap. 3, § 1.3).

Pour montrer que la forme bilinéaire ainsi définie est non dégénérée, il suffit d'exhiber deux bases, duales l'une de l'autre, des espaces vectoriels $T_a M$ et $T_a^* M$. Pour cela on choisit une carte locale $x = (x_1, \dots, x_m)$ de (M, a) (carte locale de M centrée en a), et l'on définit ∂_{x_i} ($i = 1, \dots, m$) comme le vecteur tangent à M en a dont toutes les coordonnées dans la carte x sont nulles sauf la i -ième qui vaut 1 ; dire qu'un

point mobile λ a ∂_{x_i} pour vecteur vitesse, c'est dire que $(x_j \circ \lambda)'(0) = \delta_{ij}$ (symbole de Kronecker), ce qui s'écrit encore :

$$dx_j \circ \partial_{x_i} = \delta_{ij} \quad (*)$$

(l'indice a est sous-entendu dans dx_j , différentielle en a de la fonction x_j : le contexte indique assez clairement que nous nous plaçons en a).

Nous avons ainsi construit, comme annoncé, deux bases duales l'une de l'autre :

$$\begin{array}{ll} \partial_{x_1}, \dots, \partial_{x_m} & \text{de } T_a M \\ dx_1, \dots, dx_m & \text{de } T_a^* M \end{array}$$

Notations Pour suggérer que les vecteurs cotangents s'identifient à des formes linéaires sur l'espace tangent, on note $\omega(v)$ ce que nous avons précédemment noté $\omega \circ v$. En résumé, pour $\omega = d\varphi_a$, $v = \lambda'(0)$, on écrit :

$$\begin{array}{l} \omega(v) = (\varphi \circ \lambda)'(0) = v\varphi \\ \parallel \\ d\varphi_a(v), \end{array}$$

autant d'écritures différentes pour un même nombre !

Il faut savoir naviguer à l'aise entre ces diverses notations, et savoir les utiliser notamment pour écrire les coordonnées d'un (co)vecteur dans une carte : dans la carte $x = (x_1, \dots, x_m)$ ($x_i \in C_{M,a}^\infty$) les coordonnées d'un vecteur v s'écrivent :

$$dx_1(v), \dots, dx_m(v)$$

tandis que celles d'un covecteur ω s'écrivent :

$$\omega(\partial_{x_1}), \dots, \omega(\partial_{x_m}).$$

Remarquons que, d'après la formule (*), l'identification ainsi obtenue de $T_a M$ et $T_a^* M$ à \mathbb{R}^m (via la carte x) transforme la forme bilinéaire étudiée en le produit scalaire euclidien sur \mathbb{R}^m .

Notons enfin qu'à cause de la dualité entre espaces tangents et cotangents, tout énoncé concernant des espaces tangents admet une traduction en termes d'espaces cotangents et inversement. En pratique, on pourrait dire que l'espace tangent est plus commode pour les raisonnements géométriques, et l'espace cotangent pour calculer (le « calcul différentiel » avec des dx_1, \dots, dx_m tel que nous l'avons pratiqué dans la première partie de ce livre, et tel que l'ont pratiqué avant nous des générations de mathématiciens, est un calcul dans l'espace cotangent).

L'application cotangente, transposée de l'application tangente À tout morphisme de variétés pointées :

$$M, a \cdots \xrightarrow{f} N, b$$

nous avons appris à associer deux applications linéaires :

$$\begin{array}{ccc} T_a M & \xrightarrow{T_a f} & T_b N \\ T_a^* M & \xleftarrow{T_a^* f} & T_b^* N. \end{array}$$

Il résulte immédiatement des définitions 1.1, 1.1* de ces applications qu'elles sont *transposées l'une de l'autre*, c'est-à-dire que :

$$[(T_a^* f)(\omega)](v) = \omega[(T_a f)(v)].$$

Cette « formule de transposition » mérite d'être méditée. Après en avoir compris le « mot à mot » et la justification formelle (facile !), on regardera sa traduction dans des systèmes de coordonnées locales x et y des espaces source et but ; on se convaincra qu'elle ne contient rien de plus que la formule (1.1)' du chapitre 2, dont elle nous propose une double lecture.

■ *Première lecture* (application tangente) : on pense aux $dx_i = dx_i(v)$ comme aux coordonnées d'un vecteur v (vecteur tangent à la source) ; la formule (1.1)' donne alors les coordonnées dy_j du vecteur image par l'application tangente.

■ *Deuxième lecture* (application cotangente) : les dy_i sont des vecteurs cotangents au but, mais l'application cotangente les transforme en vecteurs cotangents à la source, dont la formule (1.1)' donne l'expression dans la base constituée par les dx_i .

2. Le théorème des fonctions implicites entre en scène

À quelles notions de calcul différentiel usuel (dans \mathbb{R}^n) avons-nous fait appel jus'ici dans cette deuxième partie ?

Au chapitre 4, à presque rien : seulement au fait que le qualificatif « différentiable », accolé au substantif « application », se conserve par composition des applications – ce que résume le slogan : « les applications différentiables forment une catégorie ».

Au chapitre 5 est intervenue la notion de *différentielle* d'une application ; mais la seule chose à en retenir était son caractère linéaire, et le fait que la différentielle d'une application composée est la composée des différentielles – ce que résume la notion de « foncteur tangent ».

Les finesses du calcul différentiel commencent lorsqu'entre en scène le théorème des fonctions implicites, ou sa variante géométrique, le *théorème d'inversion locale*.

Théorème *Un morphisme de variétés $f : M, a \cdots \rightarrow N, b$ est un difféomorphisme local si et seulement si $T_a f$ est isomorphisme d'espaces vectoriels.*

Ce théorème, avec toutes ses conséquences (caractérisation des **submersions**, **immersions**, etc.), se démontre exactement comme son analogue pour les sous-variétés de \mathbb{R}^n (relire le § 4.4 du chapitre 2, qui a été rédigé de telle sorte qu'il suffise de gommer le « sous » de « sous-variété » pour obtenir les énoncés qui nous intéressent maintenant).

Comme autres conséquences importantes de ce théorème, il faut mentionner les diverses caractérisations des sous-variété S d'une variété M (cf. chap. 4, § 2.2) : caractérisation comme image réciproque d'un point par une submersion locale (« système d'équations locales » d'une sous-variété, en prenant \mathbb{R}^p comme espace but de la submersion) ; caractérisation comme image locale d'une immersion (« paramétrage local » d'une sous-variété, en prenant \mathbb{R}^l comme espace source de l'immersion) ; dans le langage du chapitre 4, § 2.2, l'analogie de la proposition du chapitre 2, § 2.3, peut s'énoncer ainsi :

toute immersion injective et propre est un plongement.

3. Fibré tangent et champs de vecteurs

À partir de maintenant il sera commode de comprendre le mot « différentiabilité » au sens C^∞ (ou C^ω).

Construire un champ de vecteurs ν sur une variété M , c'est associer à chaque point $a \in M$ un vecteur tangent à M en ce point : $\nu(a) \in T_a M$. Mais nous voudrions que $\nu(a)$ « dépende différentiablement de a » : quel sens donner à cette expression, alors que l'espace $T_a M$ dépend lui-même de a ?

Au chapitre 3, nous avons proposé deux réponses (équivalentes) à cette question : la première (chap. 3, § 1.4) était formulée en termes de composantes du champ de vecteurs dans l'espace ambiant \mathbb{R}^n où notre variété était censée être plongée ; la seconde (chap. 3, § 2.2) utilisait les coordonnées de ce même champ dans une carte locale.

Ne disposant pas a priori d'un plongement de M dans \mathbb{R}^n , nous ne pouvons nous raccrocher qu'à cette deuxième réponse, dont ce qui suit est une reformulation.

3.0. Structure différentielle de l'ensemble des vecteurs tangents

Soit $TM := J_0^1(\mathbb{R}, M)$ l'ensemble de tous les vecteurs tangents en tous les points de M (notation 0.2), et soit $\pi : TM \rightarrow M$ l'application qui, à chaque vecteur tangent $\nu_a \in (T_a M \subset TM)$, associe son « origine » a .

Pour toute carte $x : U \rightarrow \underline{U} \subset \mathbb{R}^m$ d'un ouvert U de M , notons $x : \pi^{-1}(U) \rightarrow \underline{U} \times \mathbb{R}^m$ l'application qui, au vecteur tangent $\nu_a \in T_a M$ ($a \in U$), associe le $2m$ -uplet de nombres :

$$x_1(a), \dots, x_m(a), dx_1(\nu_a), \dots, dx_m(\nu_a).$$

Cette application x est évidemment une bijection.

De plus, pour toute autre carte y , le $2m$ -uplet :

$$y_1(a), \dots, y_m(a), dy_1(\nu_a), \dots, dy_m(\nu_a)$$

est relié au précédent par une formule de « changement de carte » qui s'écrit :

$$\begin{aligned} y &= g(x) \quad (g = \text{difféomorphisme de changement de carte}) \\ dy_i &= \sum_{j=1}^m \partial_j g_i(x) dx_j. \end{aligned}$$

Cette formule définit évidemment un *difféomorphisme entre ouverts de $\mathbb{R}^m \times \mathbb{R}^m$* (différentiable en x_1, \dots, x_m et linéaire en dx_1, \dots, dx_m), de sorte qu'en faisant parcourir à x un atlas de M on obtient un atlas différentiable de TM , dont la classe d'équivalence différentiable ne dépend que de la structure différentielle qu'on s'était donnée sur M .

L'espace TM est ainsi muni d'une structure canonique de variété différentielle à $2m$ dimensions, telle que la projection $\pi : TM \rightarrow M$ soit différentiable. Le couple constitué par cette variété TM munie de sa projection π est ce qu'on appelle le **fibré tangent à M** . Dans le langage général des « espaces fibrés » (qui sera développé au chapitre suivant), l'espace M est la **base** du fibré, TM en est l'**espace total** et π la **projection** ; les espaces $\pi^{-1}(a)$ ($= T_aM$) en sont les **fibres**. Une **section** du fibré est un morphisme $v : M \rightarrow TM$ tel que $\pi \circ v = \mathbb{I}_M$ (ce qui est ici une autre façon de dire que $v(a) \in T_aM$!) : dans le cas présent c'est exactement ce que nous avons décidé d'appeler un **champ de vecteurs** sur M (nous travaillons dans la catégorie des variétés différentielles, où « morphisme » signifie « application différentiable » : dire que v est un morphisme c'est donc dire que les coordonnées $dx_1(v(a)), \dots, dx_m(v(a))$ du vecteur $v(a)$ dans une carte x sont fonctions différentiables de a).

3.1. Courbes intégrales et flot d'un champ de vecteurs

L'usage de cartes locales permet immédiatement de transposer à une variété M ce qui a été dit au chapitre 3 sur l'intégrabilité locale des champs de vecteurs (chap. 3, § 1) : on obtient ainsi les notions de **courbe intégrale** et de **flot local** d'un champ de vecteurs sur M . Dans le cas d'un champ de vecteurs à *support compact*, ces notions se globalisent : les courbes intégrales sont définies pour tout temps, et le flot est défini *globalement sur tout M* comme un groupe à un paramètre d'automorphisme (ϕ^t) de M , c'est-à-dire une application différentiable

$$\begin{aligned} \phi &: \mathbb{R} \times M \rightarrow M \\ t, a &\mapsto \phi^t(a) \end{aligned}$$

telle que

$$\begin{aligned} \phi^0 &= \mathbb{I}_M \\ \phi^{t'} \circ \phi^t &= \phi^{t+t'}. \end{aligned}$$

Un exemple simple de cette situation est fourni par l'étude 5^{bis} .

3.2. Champs de vecteurs vus comme dérivations

Les champs de vecteurs sur M peuvent être identifiés localement aux dérivations qu'ils définissent sur l'algèbre des fonctions différentiables sur M : étant donnée une carte locale x de M au voisinage de a , tout champ de vecteurs sur M au voisinage de a peut s'écrire :

$$v = v_1 \partial_{x_1} + \dots + v_m \partial_{x_m}.$$

où les fonctions $v_i \in C_{M,a}^\infty$ sont les coordonnées du champ de vecteurs v dans la carte x .

3.3. Redressement local des champs de vecteurs non singuliers

Soit v un champ de vecteurs défini au voisinage de a dans M , et tel que $v(a) \neq 0$. Alors il existe une carte locale x de M au voisinage de a , telle que $v = \partial_{x_1}$ (le travail a déjà été fait au chap. 3, § 2.3).

4. Fibré cotangent et formes différentielles de degré 1

De même qu'un champ de vecteurs sur une variété est une section du fibré tangent, un « champ de covecteurs » (ou « forme différentielle de degré 1 ») est une section du « fibré cotangent ». Bien que la définition du fibré cotangent (§ 4.0 ci-après) soit en tous points analogue à celle du fibré tangent, l'étude des formes différentielles de degré 1 pose des problèmes assez différents de l'étude des champs de vecteurs, problèmes que nous comprendrons mieux quand nous les aurons placés dans le cadre plus général des formes différentielles de degré quelconque, en relation avec le *calcul intégral sur les variétés* (chap. 8).

4.0. Structure différentielle de l'ensemble des covecteurs

Soit $T^*M = J^1(M, \mathbb{R})_0$ l'ensemble de tous les vecteurs cotangents en tous les points de M , et soit $\varpi : T^*M \rightarrow M$ la projection évidente.

Pour toute carte $x : U \rightarrow \underline{U} \subset \mathbb{R}^m$ d'un ouvert M , notons $\varpi^{-1}(U) \rightarrow \underline{U} \times \mathbb{R}^m$ l'application qui à un vecteur cotangent $\omega \in T_a^*M$ ($a \in U$) associe le $2m$ -uplet de nombres :

$$x_1(a), \dots, x_m(a), \xi_1 = \omega(\partial_{x_1}), \dots, \xi_m = \omega(\partial_{x_m}).$$

Cette application est évidemment une bijection.

De plus, pour toute autre carte y , le $2m$ -uplet :

$$y_1(a), \dots, y_m(a), \eta_1 = \omega(\partial_{y_1}), \dots, \eta_m = \omega(\partial_{y_m})$$

est relié au précédent par une formule de « changement de carte » qui peut s'écrire :

$$\begin{cases} y = g(x) & (g = \text{difféomorphisme de changement de carte}) \\ \eta_i = \sum_{j=1}^m \frac{\partial x_j}{\partial y_i} \xi_j \end{cases}$$

où $\left(\frac{\partial x_j}{\partial y_i} \right)$ est une abréviation pour la matrice jacobienne (calculée au point $y = g(x)$) de l'application :

$$y \mapsto x = g^{-1}(y).$$

Cette formule définit évidemment un *difféomorphisme entre ouverts de $\mathbb{R}^m \times \mathbb{R}^m$* (différentiable en x_1, \dots, x_m et linéaire en ξ_1, \dots, ξ_m), de sorte qu'en faisant parcourir à x un atlas de M on obtient un atlas différentiable de T^*M , dont la classe d'équivalence différentiable ne dépend que de la structure différentielle qu'on s'était donnée sur M . L'espace T^*M est ainsi muni d'une structure canonique de variété différentielle à $2m$ dimensions telle que la projection $\varpi : T^*M \rightarrow M$ soit différentiable.

Cette variété T^*M , munie de sa projection ϖ sur M , est ce qu'on appelle le **fibré cotangent** à M .

4.1. Formes différentielles de degré 1

Une section du fibré cotangent est ce qu'on appelle une **forme différentielle de degré 1** (ou « champ de covecteurs »). Si $x = (x_1, \dots, x_m)$ est une carte locale de M en a , tout champ de covecteurs sur M au voisinage de a peut s'écrire :

$$\omega = \xi_1 dx_1 + \dots + \xi_m dx_m,$$

où les fonctions $\xi_i \in C_{M,a}^\infty$ sont les coordonnées du champ de covecteurs.

On note Ω^1 [resp. $\Omega_{M,a}^1$] l'ensemble des formes différentielles de degré 1 sur M [resp. l'ensemble des germes de telles formes au voisinage de a]. C'est évidemment un espace vectoriel sur \mathbb{R} , et même un module sur l'anneau $C^\infty(M)$ [resp. $C_{M,a}^\infty$] (le produit d'une forme différentielle par une fonction est encore une forme différentielle !).

De plus, on a une application linéaire :

$$d : C^\infty(M) \text{ [resp. } C_{M,a}^\infty] \rightarrow \Omega^1(M) \text{ [resp. } \Omega_{M,a}^1]$$

qui à une fonction φ associe sa **différentielle** $d\varphi$, dont la valeur en tout point $a \in M$ est $d\varphi(a) := j_a^1(\varphi)$ (cf. 1.2*). Cette application vérifie la propriété (bien connue dans \mathbb{R}^m) :

$$d(\varphi\psi) = \psi d\varphi + \varphi d\psi$$

4.2. Image réciproque d'une forme différentielle

Soit $f : M \rightarrow N$ une application différentiable. À toute forme différentielle $\omega \in \Omega^1(N)$ est associée une forme différentielle $f^*\omega \in \Omega^1(M)$, appelée **image réciproque de ω par f** , définie par :

$$(f^*\omega)(a) = (T_a^*f)(\omega(f(a)))$$

et symbolisée par le diagramme :

$$\begin{array}{ccc} T_a^*M & \xleftarrow{T_a^*f} & T_{f(a)}^*N \\ (f^*)(a) & \longleftarrow & \omega(f(a)). \end{array}$$

En particulier, si $\omega = d\varphi$ ($\varphi \in C^\infty(N)$), il est clair que $f^*\omega = d(\varphi \circ f)$ (revoir au § 1.1* la définition de l'application cotangente). Autrement dit, en notant aussi f^* l'homomorphisme de l'algèbre $C^\infty(M)$ dans l'algèbre $C^\infty(N)$ qui à φ associe $\varphi \circ f$, on a un diagramme commutatif d'applications linéaires :

$$\begin{array}{ccc} C^\infty(M) & \xleftarrow{f^*} & C^\infty(N) \\ \downarrow d & & \downarrow d \\ \Omega^1(M) & \xleftarrow{f^*} & \Omega^1(N). \end{array}$$

Bien entendu, ce qui précède est vrai aussi au niveau local :

$$\begin{array}{ccc} C_{M,a}^\infty & \xleftarrow{f^*} & C_{N,f(a)}^\infty \\ \downarrow d & & \downarrow d \\ \Omega_{M,a}^1 & \xleftarrow{f^*} & \Omega_{N,f(a)}^1. \end{array}$$

De plus, f^* respecte évidemment la structure de module de Ω^1 , en ce sens que pour $\varphi \in C_{N,f(a)}^\infty$, $\omega \in \Omega_{N,f(a)}^1$ on a $f^*(\varphi\omega) = f^*\varphi \cdot f^*\omega$.

Ces diverses remarques fournissent une méthode pratique de calcul de l'image réciproque : si ω est donnée, dans une carte locale $y = (y_1, \dots, y_n)$ de N au voisinage de $f(a)$, par :

$$\omega = \eta_1 dy_1 + \dots + \eta_n dy_n \quad (\eta_i \in C_{N,f(a)}^\infty).$$

$f^*\omega$ sera donnée au voisinage de a par :

$$f^*\omega = f^*\eta_1 d(f^*y_1) + \dots + f^*\eta_n d(f^*y_n).$$

4.3. Différentielles exactes

Une forme différentielle $\omega \in \Omega^1(M)$ est dite **exacte** s'il existe une fonction $\varphi \in C^\infty(M)$ telle que $\omega = d\varphi$, c'est-à-dire $\omega(a) = d\varphi_a$ pour tout $a \in M$.

Le problème de reconnaître si une forme différentielle est exacte peut se décomposer en deux :

- i) *Problème local* : à quelle condition sur $\omega \in \Omega_{M,a}^1$ peut-on construire au voisinage d'un point a donné une fonction $\varphi \in C_{M,a}^\infty$ telle que $\omega = d\varphi$?
- ii) *Problème global* : les conditions locales étant supposées vérifiées partout, peut-on construire φ globalement sur M ?

La réponse au problème local s'exprime simplement dans une carte locale : écrivant $\omega = \xi dx_1 + \dots + \xi_m dx_m$ ($\xi_i \in C_{M,a}^\infty$), on voit que l'exactitude locale de ω signifie l'existence d'une fonction $\varphi \in C_{M,a}^\infty$ telle que $\xi_i = \partial_{x_i}\varphi$; une condition nécessaire évidente est $\partial_{x_i}\xi_j = \partial_{x_j}\xi_i$ pour tous i, j et il résulte de l'étude 3 (exercice 2) que cette condition est suffisante.

La réponse au problème global est plus difficile ; nous en reparlerons au chapitre 8.

Commentaire Nous retrouvons ici, dans un autre langage, des problèmes déjà abordés dans l'étude 3 sur les champs de gradients.

Si l'on peut parler de « champs de gradients » dans \mathbb{R}^n , c'est parce que le produit scalaire euclidien dans \mathbb{R}^n permet d'identifier espace tangent et espace cotangent. Plus généralement, on appelle **variété riemannienne** une variété M dont les espaces tangents T_aM ont été munis d'une *structure euclidienne* (forme bilinéaire définie positive) variant différentiablement avec a : c'est le cas par exemple de \mathbb{R}^n (avec sa structure euclidienne canonique), ou de toute sous-variété de \mathbb{R}^n si on munit ses espaces tangents de la structure euclidienne induite par celle de \mathbb{R}^n . La donnée sur une variété d'une structure riemannienne permet d'identifier espace tangent et espace cotangent, et par conséquent de définir une notion de « champs de gradients » (\Leftrightarrow différentielle exacte). Mais dans les raisonnements mettant en jeu la structure riemannienne on prendra garde à ne pas faire de difféomorphisme ne respectant pas cette structure.

ÉTUDE 5

Familles à un paramètre de droites de plan

Dans toute l'étude, P désignera un plan affine ou projectif, et \check{P} la variété des droites de ce plan.

1 Étude de l'espace tangent à \check{P} en un point D_0

Soit D_0 une droite du plan, qu'il s'agit de « mettre en mouvement ». On se propose de dresser une liste de mouvements simples dont les vecteurs « vitesse initiale » remplissent tout l'espace tangent $T_{D_0}\check{P}$, c'est-à-dire tels que tout mouvement soit tangent à l'instant $t = 0$ à l'un des mouvements de la liste (et à un seul).

Exemple de liste possible (P étant un plan affine euclidien) :

- i) Translations rectilignes uniformes (la droite $D(t)$ reste parallèle à D_0 , son écart à D_0 étant fonction linéaire du temps t) ;
- ii) rotations autour d'un point $a_0 \in D_0$, avec une vitesse angulaire uniforme ω .

Justification Dans des coordonnées cartésiennes (x, y) de P , où D_0 est la droite d'équation $y = 0$, toute droite D voisine de D_0 est donnée par une équation :

$$y = px + q.$$

Le couple $(p = p(D), q = q(D))$ est un système de coordonnées locales de \check{P} au voisinage de D_0 . Dans ces coordonnées, un mouvement de type i) s'écrit :

$$\begin{aligned} p(t) &= 0 \\ q(t) &= ct \end{aligned}$$

et un mouvement de type ii) :

$$\begin{aligned} p(t) &= \operatorname{tg} \omega t \\ q(t) &= -x_0 \operatorname{tg} \omega t \end{aligned}$$

(où x_0 est l'abscisse du point $a_0 \in D_0$).

Les coordonnées du vecteur « vitesse initiale » \check{v} sont pour le mouvement i) :

$$(dp(\check{v}) = 0, \quad dq(\check{v}) = c)$$

et pour le mouvement ii) :

$$(dp(\check{v}) = \omega, \quad dq(\check{v}) = -x_0 \omega)$$

En donnant à c, ω, x_0 toutes les valeurs possibles on remplit bien tout \mathbb{R}^2 , donc tout l'espace tangent à \check{P} .

Corollaire Caractérisation de la direction du vecteur vitesse

Ce qui précède permet de classifier tous les mouvements possibles par la direction de leur vecteur vitesse :

- i) mouvements tangents à des translations ($dp = 0$) : leurs vecteurs vitesses sont tous colinéaires ;
- ii) mouvements non tangents à des translations : la direction de leur vecteur vitesse est caractérisée par la donnée du centre $a_0 \in D_0$ de la rotation à laquelle ils sont tangents. Ce point a_0 est appelé **point caractéristique** du mouvement (ou de son vecteur vitesse) à l'instant $t = 0$.

EXERCICE On fait tourner la droite D_0 autour d'un point a_0 non situé sur D_0 . Montrer que ce mouvement (qui n'est pas tangent à une translation) a pour point caractéristique le pied de la normale abaissée de a_0 sur D_0 .

Remarque Le fait que nos « rotations » ii) aient été choisies avec une vitesse angulaire constante ω n'est évidemment pas essentiel (seule compte la vitesse initiale !). C'est seulement par souci du concret que nous avons fait intervenir dans ce qui précède des notions métriques (« distance », « angle »...). Un travail analogue aurait pu être fait aussi bien dans un plan affine non muni d'une structure euclidienne, et aussi par conséquent (s'agissant de notions locales) dans un plan projectif. L'avantage de travailler dans un plan projectif est qu'on n'a pas à distinguer deux types de mouvements : le point caractéristique est défini pour tout mouvement de vitesse non nulle (les « mouvements de translation » du cas affine sont ceux dont le point caractéristique est « à l'infini »).

Résumé Tout mouvement d'une droite du plan est tangent à un mouvement au cours duquel la droite passe par un point fixe, appelé **point caractéristique** (en convenant que celui-ci est « à l'infini » dans le cas des translations affines). Ce point est défini de façon unique dès lors que le vecteur vitesse \dot{v} n'est pas nul, et il caractérise la direction de \dot{v} .

Dans des coordonnées affines (x, y) où l'équation d'une droite s'écrit :

$$y = px + q.$$

l'abscisse x_0 du point caractéristique est donnée par la formule :

$$x_0 dp + dq = 0.$$

2 Enveloppes

Étant donnée une courbe $\Gamma \subset \check{P}$ (famille à un paramètre de droites du plan), on se pose le problème de trouver une courbe $C \subset P$ dont Γ constitue l'ensemble des tangentes.

Une telle courbe C , si elle existe, est appelée **enveloppe** de la famille de droites.

2.1 Calcul pratique de l'enveloppe

Supposons le problème résolu, c'est-à-dire que Γ soit l'ensemble des tangentes à une courbe lisse $C \subset P$. Dans des coordonnées affines (x, y) où la courbe C est donnée par une équation $y = f(x)$, la pente de la tangente s'écrit $p = f'(x)$, d'où l'on tire $q = f(x) - x f'(x)$.

Différentiant ces formules on obtient $dq = -x dp$, d'où la conclusion : quand le point $a = a(t)$ parcourt C , le mouvement de la tangente $D(t)$ à C en a admet pour point caractéristique à chaque instant le point $a(t)$.

Nous disposons ainsi d'une méthode pratique de calcul de l'enveloppe :

l'enveloppe d'une famille à un paramètre de droites est le lieu des points caractéristiques de cette famille.

2.2 Exemples

Le lecteur est fortement encouragé à essayer d'appliquer la méthode précédente à la famille des normales à un cercle, puis à la famille des normales à une parabole (ce deuxième exemple donne un joli dessin).

2.3 Critères de lissité

Les exemples du paragraphe 2.2 montrent que la notion d'enveloppe n'est pas tout à fait aussi claire que semble le suggérer le paragraphe 2.1. Allons-nous encore parler d'enveloppe si le lieu des points caractéristiques n'est pas une courbe, ou si c'est une courbe non lisse ? Inversement, le fait qu'une courbe C soit lisse entraîne-t-il automatiquement que l'ensemble Γ de ses tangentes soit une courbe lisse de \mathbb{P} , et ne serait-il pas raisonnable de chercher à définir une notion d'enveloppe d'une famille Γ non lisse ?

Si l'on n'envisage que l'aspect local du problème, les deux questions suivantes (symétriques l'une de l'autre) sont les premières à se poser naturellement.

Question C : étant donnée une courbe lisse $C \subset P$, à quelle condition l'application :

$$\gamma_C : C \rightarrow \mathbb{P},$$

qui à a associe la tangente à C en a , est-elle une immersion ?

Question Γ : étant donnée une courbe lisse $\Gamma \subset \mathbb{P}$, à quelle condition l'application :

$$\alpha_\Gamma : \Gamma \rightarrow P,$$

qui à D associe le point caractéristique du mouvement de D le long de Γ , est-elle une immersion ?

Réponse à la question C : avec les notations du paragraphe 2.1, munissons C de la coordonnée locale x , et \mathbb{P} des coordonnées locales (p, q) . Dans ces coordonnées l'application γ_C s'exprime par :

$$x \mapsto (p = f'(x), q = f(x) - x f'(x)),$$

et son application tangente par :

$$(dp = f''(x) dx, dq = -x dp).$$

Cette application tangente est donc injective si et seulement si $d(p|C) \neq 0$, c'est-à-dire $f''(x) \neq 0$, ce qui signifie géométriquement que la courbe C a un contact quadratique avec sa tangente.

Commentaire explicatif : une courbe lisse d'un plan affine (ou projectif) est dite avoir un contact quadratique avec sa tangente [resp. une tangente d'inflexion ordinaire] si dans des coordonnées affines (x, y) centrées au point considéré et telles que la tangente ait pour équation $y = 0$, la courbe peut s'écrire localement $y = \varphi(x)$ avec $\varphi \sim cx^2$ [resp. $\varphi \sim cx^3$] ($c \neq 0$).

Une caractérisation équivalente est la suivante : dans des coordonnées affines quelconques où la courbe est localement le graphe d'une fonction $y = f(x)$, l'abscisse x_0 du point considéré n'est pas un point critique [resp. est un point critique quadratique] de la fonction $x \mapsto p(x) = f'(x)$ (pente de la tangente).

Réponse à la question Γ : toujours avec les notations du paragraphe 2.1 et en supposant Γ non tangente à une translation, on peut prendre p comme coordonnée locale sur la courbe Γ , dont l'équation peut s'écrire $q = g(p)$.

En prenant (x, y) comme coordonnées locales sur P , l'application c_Γ s'exprime par :

$$p \mapsto (x = -g'(p), y = -pg'(p) + g(p)),$$

et son application tangente par :

$$(dx = -g''(p) dp, dy = p dx).$$

Cette application tangente est donc injective si et seulement si $d(x|\Gamma) \neq 0$, c'est-à-dire $g''(p) \neq 0$, ce qui signifie géométriquement que la courbe Γ a un contact quadratique avec sa tangente (rappelons que (p, q) est une carte affine d'un ouvert du plan projectif : cf. Chap. 4, exercice 9).

Conclusion : dans le cas (et seulement dans le cas) où C [resp. Γ] est une courbe lisse ayant un contact quadratique avec la tangente, l'application γ_C [resp. c_Γ] est une immersion – dont l'image locale Γ [resp. C] a aussi un contact quadratique avec sa tangente.

2.4 Singularités génériques

« En général », une courbe lisse choisie « au hasard » n'aura pas *partout* un contact quadratique avec sa tangente : en certains points *isolés* elle aura une tangente d'inflexion ordinaire.

Supposons par exemple que la courbe lisse $\Gamma \subset \check{P}$ ait au point $p = q = 0$ une tangente d'inflexion ordinaire d'équation $q = 0$.

Alors, avec les notations ci-dessus, la fonction g est de la forme :

$$g(p) = cp^3 + o(p^3),$$

où c est une constante $\neq 0$ (disons $c = 1$ pour alléger les notations).

L'application c_Γ s'exprime alors par :

$$p \mapsto (x = -3p^2 + o(p^2), y = -2p^3 + o(p^3)).$$

et son image C a un point de rebroussement à l'origine (Fig. 1).

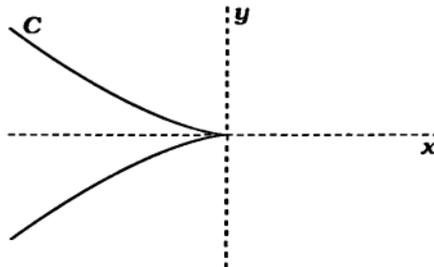


Figure 1.

Ce résultat permet de comprendre pourquoi les enveloppes des familles à un paramètre de droites du plan ont en général des points de rebroussement (cf. par exemple le deuxième exemple du paragraphe 2.2), impossibles à éliminer par petites déformations de la famille (observer, dans un bol de café au lait dont on fera varier l'inclinaison, la « caustique » de réflexion de la lumière le long des parois du bol ; la « caustique » est l'enveloppe des rayons lumineux : cf. Fig. 2).

En effet, il est facile de démontrer que la propriété pour une courbe plane d'avoir une tangente d'inflexion ordinaire est préservée par toutes déformations assez petites de la courbe (utiliser l'exercice 3 de l'étude 1).

Lectures recommandées

Une excellente initiation aux enveloppes (niveau lycée !) est le petit livre de V.C. Boltyanskii (trad. anglaise : *Enveloppes*, Pergamon 1964).

Je recommande aussi, pour la culture, l'introduction de l'article de R. Thom, « Sur la théorie des enveloppes » (*J. Maths Pures Appl. XLI*, 2 (1962)) (reproduite ci-après ; les théorèmes 4 et 5 du paragraphe II se lisent également très bien).

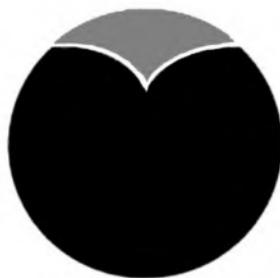


Figure 2. Caustique de réflexion dans un bol de café au lait.

Suite suggérée

Le lecteur peut dès maintenant lire l'introduction de l'étude 6 (enveloppes des familles de courbes), et faire l'« exercice recommandé ».

3 Récréation : centre instantané de rotation

On fait se mouvoir une feuille de papier posée sur une table, sans la soulever de la table.

Problème Montrer qu'à chaque instant son mouvement (supposé différentiable) est tangent soit à un mouvement de translation rectiligne uniforme, soit à un mouvement de rotation (à vitesse angulaire uniforme) autour d'un point du plan.

Commentaire Ce point, qui peut varier avec le temps, est appelé **centre instantané de rotation**. On pourra trouver des exemples concrets dans le livre de Boltyanskii déjà cité.

Indication Montrer que l'ensemble M de toutes les positions possibles de la feuille de papier est une variété, et imiter la démarche décrite au paragraphe 1.

Pour définir la structure de variété de M , on pourra imaginer qu'on a marqué un point de la feuille (par exemple l'un de ses coins) ainsi qu'une demi-droite issue de ce

point (par exemple l'un des bords de la feuille) ; se donner une position de la feuille équivaut donc à se donner, dans le plan de la table, la position du point marqué et la direction de la demi-droite marquée. Un choix de coordonnées cartésiennes dans le plan de la table nous fournit ainsi une bijection entre M et $\mathbb{R}^2 \times \mathbb{S}^1$, bijection qui permet de « transposer » sur M la structure de variété différentielle de $\mathbb{R}^2 \times \mathbb{S}^1$ (exercice : vérifier que deux choix différents de « marquage » de la feuille donnent deux bijections qui se déduisent l'une de l'autre par un difféomorphisme de $\mathbb{R}^2 \times \mathbb{S}^1$, de sorte que la structure différentielle ainsi définie sur M « ne dépend pas du marquage »).

4 Introduction à la géométrie de contact

Le lecteur aura sans doute remarqué que les calculs du paragraphe 2 n'ont fait intervenir qu'un très petit nombre de formules, très simples et n'utilisant que très peu de lettres (x, y, p, q, \dots), mais que les mêmes formules se retrouvent à divers endroits du texte avec des significations différentes, la même lettre pouvant désigner selon le contexte soit une fonction sur P ou \check{P} , soit la restriction de cette fonction à C ou Γ , soit la composée de cette fonction avec l'une des applications γ_C ou γ_Γ .

Une bonne façon de comprendre pourquoi ce sont toujours les mêmes formules qui reviennent est de ne pas travailler dans P ou \check{P} mais dans la variété des **éléments de contact**, que nous noterons \tilde{P} : c'est l'ensemble des couples $(a, D) \in P \times \check{P}$ tels que $a \in D$. Cet ensemble est une sous-variété à trois dimensions de $P \times \check{P}$: dans les coordonnées locales $(x, y; p, q)$ de $P \times \check{P}$, le sous-ensemble \tilde{P} est défini par l'équation $y = px + q$, et peut donc être vu comme le graphe d'une application $(x, p, q) \mapsto y$, ou si l'on préfère comme le graphe d'une application $(x, y, p) \mapsto q$ (de sorte que (x, p, q) aussi bien que (x, y, p) peut être pris comme système de coordonnées locales sur \tilde{P}).

4.1 Le champ de contact sur \tilde{P}

Notons π [resp. $\check{\pi}$] la projection canonique de \tilde{P} sur P [resp. \check{P}]. À tout vecteur tangent $\tilde{v} \in T_{\tilde{a}}\tilde{P}$ ($\tilde{a} = (a, D)$) l'application linéaire tangente à π [resp. $\check{\pi}$] associe un vecteur $v \in T_aP$ [resp. $\check{v} \in T_D\check{P}$].

Lemme Les deux propriétés suivantes sont équivalentes :

- i) v est tangent à D en a ;
- ii) \check{v} admet le point a comme point caractéristique (avec la convention que si $\check{v} = 0$ tout point de D est caractéristique).

L'ensemble des vecteurs $\tilde{v} \in T_{\tilde{a}}\tilde{P}$ vérifiant ces propriétés est un sous-espace vectoriel à deux dimensions de $T_{\tilde{a}}\tilde{P}$.

Preuve

L'une comme l'autre des deux propriétés équivaut à dire que $\omega(\tilde{v}) = 0$, où ω désigne le covecteur :

$$\omega = dy - p dx = x dp + dq.$$

Autrement dit, $\tilde{v} \in \text{Ker } \omega(\tilde{a})$, plan vectoriel défini par l'annulation de la forme linéaire $\omega(\tilde{a})$. □

Conclusion À chaque point \tilde{a} de notre variété à trois dimensions \tilde{P} est ainsi associé, de façon canonique, un sous-espace vectoriel à deux dimensions de l'espace tangent en \tilde{a} à cette variété.

On l'appelle le **plan de contact** en \tilde{a} , et l'on dit que la variété \tilde{P} est munie d'un « champ de plans tangents », le *champ des plans de contact*.

4.2 Courbes legendriennes

Une courbe lisse $A \subset \tilde{P}$ est dite **legendrienne** si elle est partout tangente au champ des plans de contact, c'est-à-dire si sa tangente en tout point est incluse dans le plan de contact en ce point. Avec les notations de la preuve du lemme du paragraphe 4.1, cela équivaut à dire que $\omega|_{\Lambda} = 0$, ce qui donne une façon simple de vérifier qu'une courbe est legendrienne.

On en déduit immédiatement que les courbes suivantes sont legendriennes :

- i) $\pi^{-1}(a)$ ($a \in P$) ;
- ii) $\tilde{\pi}^{-1}(D)$ ($D \in \check{P}$) ;
- iii) l'ensemble C^* des éléments de contact « attachés » à une courbe lisse $C \subset P$: nous entendons par là les couples (a, D) tels que $a \in C$ et tels que D soit tangente à C en a ;
- iv) l'ensemble ${}^*\Gamma$ des éléments de contact « attachés » à une courbe lisse $\Gamma \subset \check{P}$: nous entendons par là les couples (a, D) tels que $D \in \Gamma$ et tels que a soit le point caractéristique des vecteurs tangents à Γ en D .

4.3 Courbes legendriennes « en position générale » par rapport aux projections $\pi, \tilde{\pi}$

Soit $\Lambda \subset \tilde{P}$ une courbe legendrienne, et soit $\tilde{a}_0 = (a_0, D_0)$ un point de Λ . D'après l'exercice 6 du chapitre 2, dire que $\pi|_{\Lambda}$ [resp. $\tilde{\pi}|_{\Lambda}$] est une *immersion* (locale en \tilde{a}_0) équivaut à dire que Λ n'est pas tangente en \tilde{a}_0 à $\pi^{-1}(a_0)$ [resp. à $\tilde{\pi}^{-1}(D_0)$], condition que nous traduirons par l'expression : « Λ est en position générale (en \tilde{a}_0) par rapport à π [resp. $\tilde{\pi}$] ».

En notant C [resp. Γ] le germe de courbe lisse en a_0 [resp. D_0], image locale de cette immersion, on voit immédiatement que Λ coïncide alors, au voisinage de \tilde{a}_0 , avec l'ensemble C^* [resp. ${}^*\Gamma$] des éléments de contact attachés à C [resp. Γ]. En d'autres termes, sous l'hypothèse de position générale il n'y a pas d'autres courbes legendriennes que celles des exemples 4.2 iii) et iv).

Compte tenu de la conclusion de 2.3, nous en déduisons que la propriété pour Λ d'être en position générale à la fois par rapport à π et à $\tilde{\pi}$ équivaut à l'une comme à l'autre des deux conditions suivantes :

- i) $\Lambda = C^*$, où C est un germe de courbe lisse de P en contact quadratique avec sa tangente ;
- ii) $\Lambda = {}^*\Gamma$, où Γ est un germe de courbe lisse de \check{P} en contact quadratique avec sa tangente.

4.4 Courbes legendriennes en position non générale

Remarquons d'abord que $\pi^{-1}(a_0)$ et $\tilde{\pi}^{-1}(D_0)$ ne sont jamais tangentes, de sorte que la perte de position générale ne peut jamais avoir lieu *simultanément* par rapport à π et $\tilde{\pi}$.

Considérons par exemple le cas où Λ est tangente à $\pi^{-1}(a_0)$. N'étant pas tangente à $\tilde{\pi}^{-1}(D_0)$, elle est donc localement de la forme $\Lambda = {}^*\Gamma$, où Γ est une courbe lisse de \check{P} .

Exemple type Soit Γ la courbe de \tilde{P} donnée dans les coordonnées p, q par l'équation $q = p^3$ (cas le plus simple de courbe ayant une tangente d'inflexion). On déduit de l'équation $x dp + dq = 0$ que la coordonnée x d'un élément de contact attaché à Γ vérifie $x = -3p^2$, de sorte que la coordonnée y vérifie $y (= px + q) = -2p^3$.

Dans la carte (x, p, q) [resp. (x, y, p)] la courbe $\Lambda = \ast\Gamma$ apparaît ainsi comme une courbe gauche de \mathbb{R}^3 , tangente à l'axe des p .

La figure 3 i) [resp. 3 ii)] représente cette courbe ainsi que sa projection sur le plan des (p, q) [resp. sur le plan des (x, y)], projection qui présente comme prévu une tangente d'inflexion ordinaire [resp. un point de rebroussement ordinaire] : voir le paragraphe 2.4.

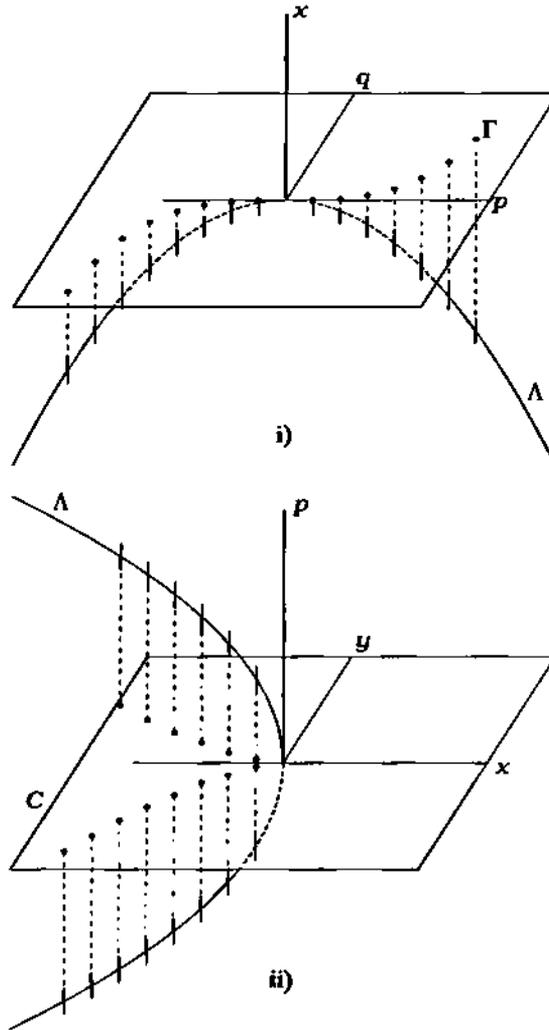


Figure 3. Restrictions des projections $\tilde{\pi}$ et π à la courbe Λ . Sur chacune de ces deux figures la direction de projection a été choisie verticale, de sorte que le plan de contact en chaque point $\tilde{a} \in \Lambda$ est le plan vertical tangent à Λ (ayant comme projection sur le plan horizontal la tangente à la courbe Γ ou C).

Conclusion provisoire L'exemple précédent se généralise immédiatement à $q = p^3 + \alpha(p^3)$, avec essentiellement les mêmes résultats. La propriété conjointe pour Γ [resp. C] d'avoir une tangente d'inflexion ordinaire [resp. un point de rebroussement ordinaire] admet une traduction (que nous donnerons plus tard) en termes de géométrie de la courbe Λ (position de Λ par rapport au champ de directions $\tilde{a} \mapsto \pi^{-1}(\pi(\tilde{a}))$, représenté par les segments verticaux de la figure 3 ii). Cette propriété est stable par petites déformations *legendriennes* de la courbe Λ , mais elle cesse de l'être si l'on oublie de demander à la courbe déformée de rester *legendrienne* (exemple : les courbes $p \mapsto (x = -3p^2, y = -2p^3 + \varepsilon p)$ sont, pour tous $\varepsilon \neq 0$, partout en position générale par rapport à la projection π).

Nous reviendrons plus tard sur ces questions, que nous généraliserons aux familles de courbes sur les surfaces (étude 6). Voici maintenant l'introduction de l'article de René Thom annoncée ci-dessus.

Sur la théorie des enveloppes par M. René THOM (Université de Strasbourg)²

Introduction La récente réforme des études de licence en mathématiques a complètement évincé des programmes la théorie des enveloppes. Sans vouloir discuter ici d'une réforme à beaucoup d'égards bénéfique, je ne puis que trouver cette disparition très regrettable ; rappelons, pour mémoire, le rôle des enveloppes dans la théorie des équations différentielles (intégrales singulières), et des équations aux dérivées partielles ; mais est-il concevable qu'un professeur de lycée ait quelque usage des problèmes de Géométrie élémentaire, sans connaître – ne fût-ce qu'approximativement – les phénomènes généraux de cette théorie ? Même d'un point de vue pratique, la théorie des enveloppes rend compte de phénomènes familiers, sans elle inexpliqués ; pour s'en convaincre, il suffit d'observer, à l'intérieur d'un bol hémisphérique de café au lait convenablement éclairé, la structure cuspidale des caustiques de réflexion, et leur variation lorsque l'éclairage se modifie.

À vrai dire, l'origine de cette condamnation se discerne aisément ; c'est le caractère insuffisant, non rigoureux, de la théorie classique, qu'on doit mettre en cause. Pour tout professeur épris de clarté et de précision, l'exposé des enveloppes, avec sa longue liste de phénomènes exceptionnels : points fixes, points singuliers, courbes stationnaires, etc., tourne rapidement au martyre : car rien n'assure que tous les phénomènes pathologiques ont été effectivement catalogués. Ainsi la théorie, par sa richesse même, refuse de se plier aux canons d'une pédagogie rigoureuse : il faut sans doute voir là une des raisons majeures de son exclusion des programmes.

L'article qui suit, dans sa première partie, a pour but de donner à la théorie des enveloppes les fondements théoriques dont elle a jusqu'à présent manqué. À cet effet, le problème est envisagé avec la plus grande généralité possible : on est ainsi amené à replacer la théorie dans le cadre plus général de la théorie des singularités d'applications différentiables, que divers auteurs ont récemment développée ([3], [6]). Deux idées centrales se dégagent de cette étude :

1° Distinguer parmi les singularités que peuvent présenter les enveloppes, deux grands types : les singularités « stables » (ou « génériques »), qui subsistent lorsqu'on déforme suffisamment peu la famille de sous-variétés dont on prend l'enveloppe, et les singularités instables (« non génériques ») qu'une déformation arbitrairement petite de la famille fait disparaître. Une pédagogie saine devra se borner à la description des singularités stables, qui sont (à ce qu'on peut supposer) en nombre fini, en délaissant provisoirement les singularités instables, qui

² *J. Maths Pures Appl.* XLI, 2 (1962). Reproduit avec l'aimable autorisation des Éditions Gauthier-Villars.

ne pourront de toute façon pas faire l'objet d'une description exhaustive, parce qu'il y en a une infinité.

2° Alors que la pédagogie classique s'efforce (au prix de quelles difficultés !) d'isoler l'enveloppe géométrique au sens classique de composantes telles que variétés stationnaires, etc., usuellement considérées comme parasites, on verra que cette distinction est tout à fait arbitraire, et devrait être rejetée. En effet, l'existence d'une enveloppe au sens large est un phénomène très général ; au contraire, l'enveloppe géométrique au sens classique ne se présente que si des inégalités portant sur le nombre r des paramètres sont vérifiées (en fait, si la dimension r de la famille n'est pas trop différente de la codimension m des sous-variétés qu'on fait varier).

Bien entendu, les considérations qui vont suivre ne pourront pas être transplantées telles quelles dans l'enseignement ; mais du moins, je l'espère, elles indiquent la voie à suivre pour sortir la théorie du chaos où elle se débat encore actuellement. Dans une deuxième partie, j'applique la théorie des singularités des applications différentiables au cas d'enveloppe le plus simple : famille compacte à un paramètre de droites dans le plan projectif. On retrouve ainsi des résultats connus de Géométrie finie [2] ; leur généralisation aux dimensions supérieures ne manquerait pas d'intérêt.

ÉTUDE 5bis

Champs de vecteurs constants sur les tores

1 Nous avons défini dans l'étude 4^{ter} l'atlas des « systèmes de coordonnées angulaires » sur le tore $\mathbb{T}^n = \mathbb{R}^n / 2\pi \mathbb{Z}^n$. Comme deux systèmes de coordonnées angulaires sur \mathbb{T}^n ne diffèrent que par une translation :

$$(\theta_1, \dots, \theta_n) \mapsto (\theta_1 + 2\pi k_1, \dots, \theta_n + 2\pi k_n),$$

il est clair que le champ de vecteurs ∂_{θ_i} (ainsi d'ailleurs que la forme différentielle $d\theta_i$) ne dépend pas du choix du système. Le tore \mathbb{T}^n est ainsi muni canoniquement de n champs de vecteurs $\partial_{\theta_1}, \dots, \partial_{\theta_n}$ (et de n formes différentielles $d\theta_1, \dots, d\theta_n$) linéairement indépendants en tout point.

On appellera **champ de vecteurs constant** sur \mathbb{T}^n toute combinaison linéaire à coefficients constants de ces « champs de vecteurs de base ».

La notion de « champ de vecteurs constant » peut aussi être présentée de façon plus intrinsèque, sans privilégier une base. Soit $\pi : \mathbb{A} \rightarrow \mathbb{A}/\Omega$ l'application canonique de passage au quotient d'un espace affine \mathbb{A} de dimension n par un réseau Ω de translations de \mathbb{A} (cf. étude 4^{ter}, § 1).

Sur \mathbb{A} , les « champs de vecteurs constants » sont les champs de vecteurs invariants par translation. Ils forment un espace vectoriel à n dimensions. Étant invariants en particulier par l'action de Ω , ils définissent donc (via le difféomorphisme local π) des champs de vecteurs sur le tore \mathbb{A}/Ω , que nous appelons *champs de vecteurs constants sur le tore*.

Sur \mathbb{A} , le flot d'un champ de vecteurs constant est un groupe à un paramètre de translations de \mathbb{A} . Toute translation τ de \mathbb{A} induit par passage au quotient un difféomorphisme $\tilde{\tau}$ du tore appelé **translation du tore**, qui rend commutatif le diagramme :

$$\begin{array}{ccc} \mathbb{A} & \xrightarrow{\tau} & \mathbb{A} \\ \downarrow \pi & & \downarrow \pi \\ \mathbb{A}/\Omega & \xrightarrow{\tilde{\tau}} & \mathbb{A}/\Omega \end{array}$$

Le flot d'un champ de vecteurs constant sur le tore est ainsi un groupe à un paramètre de « translations » du tore.

Travaux pratiques

i) En figurant le tore \mathbb{T}^2 par un carré dont les côtés opposés sont identifiés, on dessinera les trajectoires intégrales d'un champ de vecteurs constant $v = a_1 \partial_{\theta_1} + a_2 \partial_{\theta_2}$, pour des valeurs simples de a_1, a_2 . On constatera que l'allure des trajectoires dépend de façon cruciale des propriétés arithmétiques du couple (a_1, a_2) , chaque trajectoire recoupant d'autant moins souvent les bords du carré que a_1 et a_2 sont dans des rapports rationnels simples ; s'ils sont dans un rapport irrationnel, *chaque trajectoire intégrale est dense sur le tore !*

Question : existe-t-il un difféomorphisme du tore dans lui-même qui transforme le champ de vecteurs ∂_{θ_1} en $\partial_{\theta_1} + \partial_{\theta_2}$? Essayer de trouver un énoncé général dont ceci soit un cas particulier.

ii) En considérant le tore comme surface plongée dans \mathbb{R}^3 , on cherchera à redessiner les trajectoires intégrales trouvées en i). Il sera commode d'adopter la convention graphique de la figure 6 ii) de l'étude 2 (« tore vu de face »). Certains de vos dessins ne font-ils pas penser à la figure 5 de l'étude 3^{bis} (précession du périhélie des orbites planétaires) ?

2 Lien avec l'étude 3^{bis}

2.1 Le paragraphe 1 de l'étude 3^{bis} nous a montré que sous certaines hypothèses les courbes intégrales λ d'un oscillateur à une dimension sont des courbes *périodiques*, ayant pour images dans le « plan des phases » (p, q) les *courbes d'énergie constante* C_E .

EXERCICE Sous les hypothèses de l'étude 3^{bis}, paragraphe 1, montrer que la courbe d'énergie constante C_E est *difféomorphe* au cercle $S^1 = \mathbb{R}/2\pi\mathbb{Z}$, par un *difféomorphisme* qui transforme le champ de vecteurs $v_H|_{C_E}$ en un champ de vecteurs constant sur S^1 (idée : paramétrer la courbe C_E par le temps).

2.2 Dans le cas du problème de Kepler à deux dimensions (cf. étude 3^{bis}, § 2), le même raisonnement appliqué au mouvement radial donne dans le « plan des phases radial » (r, p_r) une courbe $C_{E,k}$ qui pour de « bonnes » valeurs de (E, k) (énergie et moment angulaire) sera aussi *difféomorphe* à un cercle. Il en résulte grâce à l'invariance par rotation que dans le plan des phases $\mathbb{R}^2 \times \mathbb{R}^2$ la surface :

$$\mathcal{T}_{E,k} = \{(p, q) \in \mathbb{R}^2 \times \mathbb{R}^2 \mid H(p, q) = E, K(p, q) = k\}$$

est *difféomorphe* à un tore, puisque c'est le produit du « cercle » $C_{E,k}$ par le cercle trigonométrique du plan des q (paramétré par l'angle $\varphi \in \mathbb{R}/2\pi\mathbb{Z}$). La figure 5 de l'étude 3^{bis} peut donc être considérée comme représentant, en projection sur le plan des q , une trajectoire intégrale du champ v_H sur ce tore.

Nous montrerons dans l'étude 8 un résultat analogue à celui de l'exercice ci-dessus et qui se généralise en dimension supérieure :

le *difféomorphisme* entre la variété $\mathcal{T}_{E,k}$ et le tore standard peut être choisi de façon à transformer le champ de vecteurs v_H en un champ de vecteurs constant sur le tore (ce choix est d'ailleurs canonique).

Espaces fibrés

En définissant la notion de champ de vecteurs [resp. champ de covecteurs] sur une variété M , nous avons été amenés à considérer la famille, paramétrée par $a \in M$, de tous les espaces tangents $T_a M$ [resp. espaces cotangents $T_a^* M$], à mettre une structure différentielle sur l'« espace total » de cette famille (cf. chap. 5, § 3 et 4). Nous avons obtenu ainsi la notion de fibré tangent [resp. fibré cotangent].

Nous aurons maintes fois l'occasion de rencontrer des situations analogues, où l'on « recolle » une famille (E_a) d'espaces, paramétrée par un point a d'une variété M , pour obtenir ce qu'on appelle un *fibré sur M* .

Lorsque les espaces que l'on recolle sont des espaces vectoriels, on obtient (sous certaines hypothèses sur la façon de les recoller) ce qu'on appelle un *fibré vectoriel*. S'il s'agit d'espaces affines, ou projectifs, etc., on parle de *fibré affine*, de *fibré projectif*, etc. (par exemple, l'étude des enveloppes de familles de droites du plan conduit à étudier des « fibrés en droites affines »...).

Ce chapitre présente les notions de base de la théorie des espaces fibrés, d'abord dans le cadre très général de la catégorie des espaces topologiques, puis en restreignant à la catégorie des variétés différentielles.

Comme souvent en mathématiques, l'hypergénéralité de la présentation est motivée d'une part par la grande diversité des situations où les notions étudiées ont vocation à intervenir, d'autre part par un souci de clarté conceptuelle (comprendre à quel endroit telle hypothèse intervient...).

En contrepoint à cette présentation hypergénérale, le thème d'étude qui suit (« enveloppes de familles de courbes ») nous offrira l'occasion de manipuler ces notions sous toutes leurs facettes, dans un contexte concret où elles n'apparaîtront pas gratuites (du moins, je l'espère !).

En guise de « mise en train », commençons par un exemple simple, peut-être l'exemple le plus simple de fibré en droites non trivial.

Exemple introductif Soit Ω la variété des droites affines de \mathbb{R}^2 (ou d'un plan affine A). C'est une variété différentielle à deux dimensions, et en associant à toute droite $D \in \Omega$ sa direction on obtient une application différentiable :

$$\pi : \Omega \rightarrow \mathbb{P}^1.$$

où \mathbb{P}^1 est la droite projective, ensemble des directions de droites de \mathbb{R}^2 [resp. A].

Pour tout $\delta \in \mathbb{P}^1$, $\pi^{-1}(\delta)$ est l'ensemble des droites affines de direction δ : c'est donc une droite vectorielle [resp. affine] :

$$\pi^{-1}(\delta) = \mathbb{R}^2 / \delta \quad [\text{resp. } \mathbb{A} / \delta],$$

où $/\delta$ désigne le quotient par la relation d'équivalence qui identifie deux points du plan situés sur une même droite de direction δ .

La variété Ω , munie de sa projection π sur \mathbb{P}^1 , est ce qu'on appelle un *fibré en droites* sur \mathbb{P}^1 . Les « droites » $\pi^{-1}(\delta)$ sont les *fibres* de ce fibré. En choisissant dans le plan un « point marqué » O (par exemple l'origine dans le cas $\mathbb{A} = \mathbb{R}^2$), on peut définir une *section remarquable* $\gamma_0 : \mathbb{P}^1 \rightarrow \Omega$, à savoir l'application qui à δ associe la droite de direction δ passant par O .

Affirmation Toute section (continue) γ du fibré Ω rencontre γ_0 en au moins un point δ .

Supposons en effet qu'il existe une section γ telle que $\gamma(\delta) \neq \gamma_0(\delta)$ pour tout $\delta \in \mathbb{P}^1$. La fibre $\pi^{-1}(\delta)$ est partagée par le point $\gamma_0(\delta)$ en deux demi-droites, dont l'une est distinguée par le fait qu'elle contient le point $\gamma(\delta)$ (cf. Fig. 1, où cette demi-droite distinguée a été symbolisée par une flèche).

Comme la section γ est continue, on doit pouvoir suivre continûment avec δ l'évolution de la demi-droite ainsi distinguée dans $\pi^{-1}(\delta)$. Or, quand δ parcourt tout \mathbb{P}^1 (son argument augmentant de π), la flèche de la figure 1 fait un demi-tour, de sorte que la demi-droite que nous avons distinguée sur la figure est transformée en son opposée : il n'est donc pas possible de « distinguer », continûment avec δ , une demi-droite de $\pi^{-1}(\delta)$.

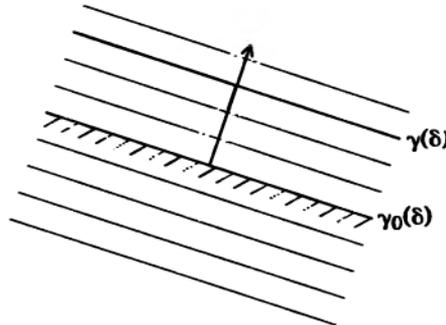


Figure 1. La fibre $\pi^{-1}(\delta)$ (ensemble des droites parallèles à δ) est partagée en deux par le point $\gamma_0(\delta)$.

Remarque S'il avait été possible de distinguer continûment une demi-droite de la fibre, notre fibré en droites aurait été ce qu'on appelle un *fibré trivial*, difféomorphe à un cylindre à base circulaire ($S^1 \times \mathbb{R}$, muni de la projection canonique sur sa base S^1 : rappelons que \mathbb{P}^1 est difféomorphe à S^1). Un modèle en papier de ce cylindre se construit en recollant les deux bords d'une bande infinie $[0, 1] \times \mathbb{R}$ selon le schéma de recollement $(0, y) \sim (1, y)$ (cf. Fig. 2 i)).

En fait, le fibré Ω qui nous intéresse ici est représentable par un modèle analogue, où le schéma de recollement serait $(0, y) \sim (1, -y)$ (cf. Fig. 2 ii)). Pratiquement, un tel recollement n'est pas réalisable sans déchirer le papier, mais il devient réalisable si, au lieu de vouloir représenter tout Ω , on se contente de représenter un voisinage

de la courbe Γ_0 , image de γ_0 (bande hachurée de la Fig. 2 ii) ; le recollement des bords verticaux de la bande selon le schéma 2 ii) est alors réalisable pratiquement par simple torsion de la bande de papier, et l'on obtient la surface célèbre appelée *ruban de Möbius*.

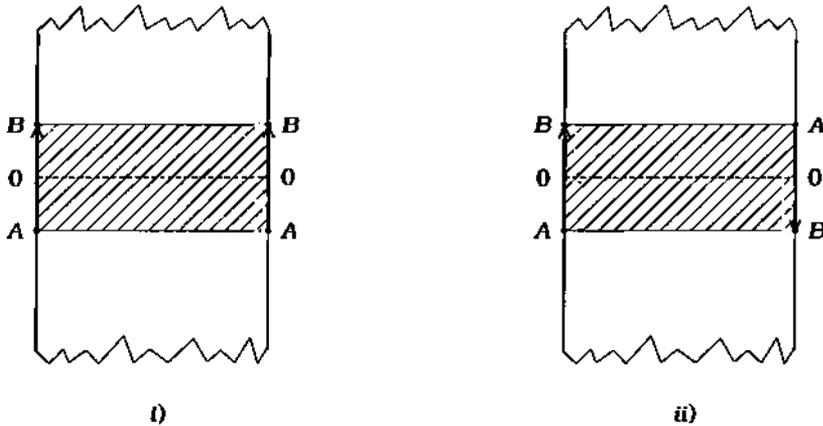


Figure 2.

0. De la notion de famille d'espaces à celle d'espace fibré

Dans ce paragraphe nous oublions provisoirement les structures différentielles pour nous placer dans la catégorie des *espaces topologiques* : le mot « espace » [resp. « application »] signifiera donc « espace topologique » [resp. « application continue »].

0.0. Espaces projetés

Étant donné un espace B , on appelle **espace au-dessus de B** (ou **espace projeté sur B**) la donnée :

$$E = (|E|, \pi_E)$$

d'un espace $|E|$ et d'une application $\pi_E : |E| \rightarrow B$. Cette application π_E est appelée la **projection** de E ; B en est la **base** et $|E|$ l'**espace total**.

Pour $b \in B$, l'espace $E_b = \pi_E^{-1}(b)$ est appelé **fibre de E au-dessus de b** .

On appelle **section** de E une application

$$\sigma : B \rightarrow |E| \quad \text{telle que} \quad \pi_E \circ \sigma = \text{Id}_B.$$

D'un point de vue purement ensembliste (en oubliant la topologie), la donnée de E équivaut à la donnée de la famille $(E_b)_{b \in B}$ (l'espace total $|E|$ étant l'union disjointe des E_b). On pourra donc penser à un espace au-dessus de B comme à une famille d'espaces topologiques $(E_b)_{b \in B}$, dont l'espace total $|E|$ aurait lui-même été muni d'une topologie telle que la projection π_E soit continue, la topologie des E_b étant la topologie induite par l'inclusion $E_b \subset |E|$.

0.1. Morphismes d'espaces projetés

Étant donnée une application $\lambda : B \rightarrow B'$, on appelle **morphisme au-dessus de λ** la donnée, notée

$$f : E \rightarrow E'$$

d'un espace $E = (|E|, \pi_E)$ au-dessus de B , d'un espace $E' = (|E'|, \pi_{E'})$ au-dessus de B' , et d'une application $f : |E| \rightarrow |E'|$ telle que le diagramme

$$\begin{array}{ccc} |E| & \xrightarrow{f} & |E'| \\ \downarrow \pi_E & & \downarrow \pi_{E'} \\ B & \xrightarrow{\lambda} & B' \end{array}$$

soit commutatif.

D'un point de vue ensembliste, cela équivaut à se donner une famille d'applications :

$$(f_b : E_b \rightarrow E'_{\lambda(b)})_{b \in B} ;$$

mais le fait que f soit continue signifie non seulement que les f_b sont continues mais encore qu'elles « dépendent continûment de b ».

Un cas particulier important est celui où tous les espaces considérés ont même base B . On appelle **morphisme d'espaces au-dessus de B** un morphisme au-dessus de \mathbb{I}_B , c'est-à-dire la donnée d'un diagramme commutatif :

$$\begin{array}{ccc} |E| & \xrightarrow{f} & |E'| \\ \searrow \pi_E & & \swarrow \pi_{E'} \\ & B & \end{array}$$

En termes imagés, il s'agit donc d'une famille d'applications $(f_b : E_b \rightarrow E'_b)_{b \in B}$ « se recollant de façon continue ».

Bien entendu, les espaces projetés, munis des morphismes définis ci-dessus, forment une catégorie, et l'on appelle **isomorphismes d'espaces projetés** les morphismes admettant un morphisme inverse. Cette phrase peut se comprendre soit dans la catégorie de tous les espaces projetés soit, si la base B est fixée par le contexte, dans la catégorie des espaces au-dessus de B , en convenant alors de ne considérer que les morphismes au-dessus de \mathbb{I}_B .

0.2. Produits fibrés

Soient $E_i = (|E_i|, \pi_{E_i})$, $i = 1, 2$, deux espaces au-dessus de B . Considérant chacun d'entre eux comme une famille $(E_{i,b})_{b \in B}$, nous allons construire un espace au-dessus de B dont la famille des fibres sera $(E_{1,b} \times E_{2,b})_{b \in B}$.

Cet espace, que l'on note $E_1 \times_B E_2$, sera muni de morphismes canoniques :

$$E_1 \xleftarrow{pr_1} E_1 \times_B E_2 \xrightarrow{pr_2} E_2 .$$

On l'appellera **produit fibré de E_1 par E_2** .

Pour le construire, on définit $\left| E_1 \times_B E_2 \right|$ comme le sous-espace de $|E_1| \times |E_2|$ constitué par les couples (x_1, x_2) tels que $\pi_{E_1}(x_1) = \pi_{E_2}(x_2)$. L'application $\pi_{E_1 \times_B E_2}$ est celle qui, à un tel couple, associe la valeur commune de $\pi_{E_1}(x_1)$ et $\pi_{E_2}(x_2)$. Enfin, l'application pr_i ($i = 1, 2$) est celle qui à un tel couple associe x_i .

Propriété universelle du produit fibré

Pour tout couple de morphismes d'espaces au-dessus de B :

$$f_i : E \rightarrow E_i \quad (i = 1, 2)$$

il existe un unique morphisme :

$$f_1 \times_B f_2 : E \rightarrow E_1 \times_B E_2$$

$$\text{tel que } f_i = pr_i \circ \left(f_1 \times_B f_2 \right) \quad (i = 1, 2).$$

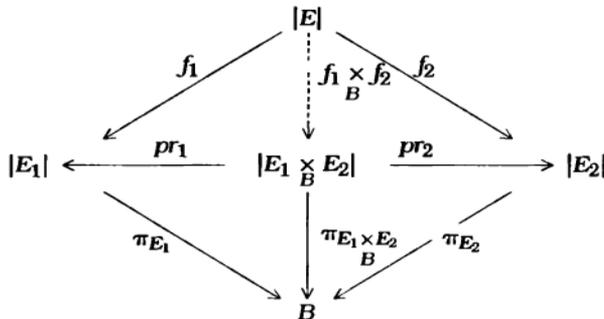
Preuve

Ce morphisme est défini par :

$$\left(f_1 \times_B f_2 \right) (x) = (f_1(x), f_2(x))$$

(remarquer que $\pi_{E_1}(f_1(x)) = \pi_{E_2}(f_2(x))$) puisque f_1 et f_2 sont des morphismes d'espaces au-dessus de B . \square

Illustration diagrammatique



0.3. Changements de base

Soit $E = (|E|, \pi_E)$ un espace au-dessus de B , et soit $\lambda : B' \rightarrow B$ une application.

Considérant E comme une famille $(E_b)_{b \in B}$, nous allons construire un espace au-dessus de B' dont la famille des fibres sera $(E_{\lambda(b')})_{b' \in B'}$. Cet espace, qui sera noté $\lambda^* E$,

sera muni d'un morphisme canonique au-dessus de λ :

$$\begin{array}{ccc} |\lambda^* E| & \xrightarrow{\lambda^E} & |E| \\ \downarrow \pi_{\lambda^* E} & & \downarrow \pi_E \\ B' & \xrightarrow{\lambda} & B \end{array}$$

Nous l'appellerons **espace induit de E par le changement de base λ** (en anglais, *pull back of E by λ* , parfois traduit plaisamment par « rétroirette de E par λ »).

Pour le construire, on définit $|\lambda^* E|$ comme le sous-ensemble de $B' \times |E|$ constitué par les couples (b', x) tels que $\lambda(b') = \pi_E(x)$. L'application $\pi_{\lambda^* E}$ [resp. λ^E] est celle qui à tout couple (b', x) associe b' [resp. x].

Section induite À toute section σ de E est associée une unique section $\lambda^* \sigma$ de $\lambda^* E$ (dite « induite de σ par le changement de base λ ») rendant commutatif le diagramme :

$$\begin{array}{ccc} |\lambda^* E| & \xrightarrow{\lambda^E} & |E| \\ \uparrow \lambda^* \sigma & & \uparrow \sigma \\ B' & \xrightarrow{\lambda} & B \end{array}$$

(exercice : la construire).

Remarque Il saute aux yeux que $|\lambda^* E|$, muni de la projection évidente sur B , n'est autre que le produit fibré de $B' = (B', \lambda)$ par E ; mais ce n'est pas de cette projection-là qu'on le munit.

En relisant, dans cette nouvelle optique, la propriété universelle du produit fibré, on obtient la propriété suivante.

Propriété universelle du changement de base

Tout morphisme $f : E' \rightarrow E$ au-dessus de λ se factorise de façon unique à travers un morphisme $f_{B'}$, au-dessus de B' :

$$f = \lambda^E \circ f_{B'}$$

$$\begin{array}{ccccc} f : |E'| & \xrightarrow{f_{B'}} & |\lambda^* E| & \xrightarrow{\lambda^E} & |E| \\ & \searrow & \swarrow & & \swarrow \\ & & B' & \xrightarrow{\lambda} & B \end{array}$$

Cas particulier : restriction au-dessus d'un sous-espace de la base Si B' est un sous-espace de B , λ étant l'inclusion canonique, l'espace $\lambda^* E$ est noté $E_{B'}$ et appelé **restriction de E au-dessus de B'** ; l'application λ^E est alors, comme λ , une inclusion canonique.

Dans l'interprétation des espaces projetés comme familles d'espaces, l'opération de restriction consiste simplement à remplacer la famille $(E_b)_{b \in B}$ par la sous-famille $(E_{b'})_{b' \in B'}$ ($B' \subset B$).

Remarque La notation $E_{B'}$ est parfois utilisée au lieu de λ^*E lorsque λ , sans être nécessairement une inclusion, est une application canonique de B' dans B (c'est-à-dire lorsque la donnée de B' ne laisse aucun doute sur ce qu'est λ , vu le contexte).

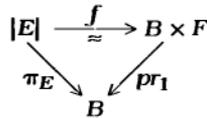
Exemple Étant donnés deux espaces B et F , on note F_B l'espace au-dessus de B défini par $|F_B| = B \times F$, $\pi_{F_B} = p_{r1}$ (première projection).

Cette notation est bien conforme à la remarque précédente, à condition de considérer F comme « espace au-dessus d'un seul point $\{*\}$ » (π_F étant l'application constante de F sur $\{*\}$). F_B est alors l'espace au-dessus de B induit de F par le changement de base constant $B \rightarrow \{*\}$!

0.4. La notion générale d'espace fibré

On appelle **fibré trivial**, de base B et de fibre F , tout espace E au-dessus de B qui est isomorphe à F_B (cf. exemple précédent).

Un tel isomorphisme $f : E \rightarrow F_B$, qui se traduit par un diagramme :



est appelé **trivialisation** du fibré E . Sa restriction à chaque fibre E_b définit un isomorphisme entre celle-ci et la « fibre modèle » F .

Plus généralement, on appelle **espace fibré de base B et de fibre F** un espace E au-dessus de B vérifiant la propriété de **trivialité locale** suivante :

Tout point $b \in B$ admet un voisinage U tel que E_U soit un fibré trivial de base U et de fibre F .

Il est facile de voir que les opérations de *produit fibré* et de *changement de base*, appliqués à des espaces fibrés, donnent encore des espaces fibrés. Plus précisément :

- i) Si E_1 et E_2 sont deux espaces fibrés de même base B et de fibres respectives F_1, F_2 , $E_1 \times_B E_2$ est un espace fibré de fibre $F_1 \times F_2$.
- ii) Si E est un espace fibré de base B et de fibre F , l'espace λ^*E induit de E par tout changement de base $\lambda : B' \rightarrow B$ est un espace fibré de base B' et de même fibre F .

De plus, les propriétés universelles du produit fibré et du changement de base s'énoncent pour les espaces fibrés comme pour les espaces projetés.

1. Fibrés vectoriels

1.0. La catégorie des fibrés vectoriels

Relisons tout le paragraphe 0 avec les restrictions suivantes :

- i) les fibres E_b de tous les espaces projetés sont des espaces vectoriels de dimension finie (sur un corps \mathbb{K} qui pourra être \mathbb{R} ou \mathbb{C}) ;
- ii) les morphismes, en restriction à chaque fibre, sont des applications \mathbb{K} -linéaires entre ces espaces vectoriels.

On obtient ainsi la catégorie des *espaces projetés à fibres vectorielles*, et l'on remarque que les opérations de produit fibré et de changement de base ne font pas sortir de cette catégorie.

En relisant de la même façon le paragraphe 0.4, on obtient la notion de « fibré vectoriel trivial », puis celle de **fibré vectoriel**. La « fibre modèle » d'un fibré vectoriel E est un \mathbb{K} -espace vectoriel F de dimension finie (disons, $F = \mathbb{K}^n$), dont la dimension n est appelée **rang** du fibré vectoriel.

Bien entendu, les opérations de produit fibré et de changement de base opèrent à l'intérieur de la catégorie des fibrés vectoriels.

1.1. Caractérisation des isomorphismes de fibrés vectoriels

PROPOSITION Soit $f : E \rightarrow E'$ un morphisme (au-dessus de \mathbb{I}_B) de fibrés vectoriels de même base B . Alors f est un isomorphisme si et seulement si, pour tout $b \in B$, l'application linéaire f_b est de rang $n = \text{rang } E = \text{rang } E'$.

Preuve

La nécessité est évidente, puisqu'un isomorphisme de fibré vectoriel donne, en restriction aux fibres, des isomorphismes d'espaces vectoriels.

Réciproquement, supposons que pour tout $b \in B$ l'application f_b soit un isomorphisme d'espaces vectoriels. La famille d'applications $(g_b = (f_b)^{-1})_{b \in B}$ est la candidate évidente au titre de « morphisme inverse » de f , le seul problème étant de démontrer que ces applications « dépendent continûment de b ».

Comme la continuité est une propriété locale, on peut se ramener, par trivialisations de E et E' , à travailler sur le fibré trivial :

$$\mathbb{K}_U^n = \left(U \times \mathbb{K}^n \xrightarrow{pr_1} U \right)$$

(U ouvert de B). La famille d'applications (f_b) est ainsi représentée pour $b \in U$ par une famille $(a(b))$ de matrices $n \times n$ dépendant continûment de b , et tout se ramène au lemme ci-après. □

Lemme Une matrice $a(b)$, inversible pour tout b et dépendant continûment de b , a pour inverse une matrice qui dépend continûment de b .

Ce lemme est évident : les coefficients de la matrice inverse sont des fonctions rationnelles des $a_{ij}(b)$ ayant pour dénominateur $\det a(b)$; or l'inverse d'une fonction

continue qui ne s'annule pas est une fonction continue (je souligne cette évidence pour faire observer qu'on peut, dans tout ce qui précède, remplacer l'adjectif « continu » par « différentiable » ou « analytique », par exemple). \square

Corollaire (caractérisation des trivialisations d'un fibré vectoriel)

i) Se donner une trivialisations d'un fibré vectoriel E de rang n , c'est se donner sur $|E|$ n fonctions x_1, \dots, x_n à valeurs dans \mathbb{K} , linéaires et linéairement indépendantes en restriction à chaque fibre.

En effet, un morphisme du fibré vectoriel E dans le fibré vectoriel \mathbb{K}_B^n est une application continue de la forme $|E| \xrightarrow{\pi_E \times x} B \times \mathbb{K}^n$, où $x : |E| \rightarrow \mathbb{K}^n$ est linéaire en restriction à chaque fibre. Dire que ce morphisme est un isomorphisme équivaut, d'après la proposition, à dire que pour tout b l'application linéaire $x|_{E_b}$ est de rang n .

ii) Se donner une trivialisations d'un fibré vectoriel E de rang n équivaut à s'en donner n sections s_1, \dots, s_n partout linéairement indépendantes.

En effet la donnée d'un morphisme s du fibré vectoriel \mathbb{K}_B^n dans E équivaut de façon évidente à la donnée de n sections s_1, \dots, s_n de E , qui définissent s par la formule $s(b, x) = x_1 s_1(b) + \dots + x_n s_n(b)$. Dire que ce morphisme est un isomorphisme équivaut, d'après la proposition, à dire que pour tout b les vecteurs $s_1(b), \dots, s_n(b)$ sont linéairement indépendants.

Commentaire Par définition d'un fibré vectoriel, un tel n -uplet de sections s_1, \dots, s_n [resp. de fonctions x_1, \dots, x_n comme dans i)] existe toujours localement, c'est-à-dire sur les ouverts assez petits de B .

C'est l'analogie pour les fibrés vectoriels de ce qu'est une base [resp. un système de coordonnées linéaires] pour un espace vectoriel.

EXERCICE TEST Les fibrés vectoriels suivants sont-ils triviaux ?

- TS^1 (fibré tangent au cercle) ;
- TS^2 (fibré tangent à la sphère) ;
- $T\mathbb{T}^n$ (fibré tangent au tore $\mathbb{T}^n = S^1 \times \dots \times S^1$).

1.2. Opérations sur les fibrés vectoriels

La base B étant fixée, on peut faire sur les fibrés vectoriels toutes les opérations que l'on fait habituellement sur les espaces vectoriels.

i) **Somme directe** : $E \oplus E' = E \times_B E'$ (la fibre $F \times F' = F \oplus F'$ est la somme directe des espaces vectoriels fibres).

ii) **Quotient par un sous-fibré vectoriel** : on dit que E' est un **sous-fibré vectoriel** de E (notation : $E' \subset E$) si son espace total $|E'|$ est un sous-espace de $|E|$, l'inclusion $|E'| \subset |E|$ définissant un morphisme de fibrés vectoriels (de sorte que chaque fibre E'_b est un sous-espace vectoriel de E_b).

Considérons alors sur $|E|$ la relation d'équivalence :

$$x \sim x' \iff \begin{matrix} \pi(x) = \pi(x') (= b) \\ x - x' \in E'_b \end{matrix}$$

Le quotient de $|E|$ par cette relation d'équivalence, muni de la projection évidente sur B , est un fibré vectoriel de fibre modèle F/F' (espace vectoriel quotient des fibres de E et E').

iii) **Fibré des homomorphismes** : on note $\text{Hom}(E, E')$, ou encore $\mathcal{L}(E, E')$, l'ensemble des morphismes de E dans E' , au sens de la catégorie des fibrés vectoriels sur B ; c'est un fibré vectoriel, de fibre « modèle » $\text{Hom}_{\mathbb{K}}(F, F')$ (ensemble des applications \mathbb{K} -linéaires de l'espace vectoriel F dans l'espace vectoriel F').

iv) **Fibré vectoriel dual** : c'est un cas particulier de la notion précédente. À tout fibré vectoriel E de base B est associé le fibré vectoriel $E^* = \text{Hom}(E, \mathbb{K}_B)$, où \mathbb{K}_B est le fibré vectoriel trivial déduit de \mathbb{K} (corps de base) par le changement de base constant (cf. la notation introduite à la fin du paragraphe 0.3).

Bien entendu, la fibre modèle de E^* est l'espace vectoriel F^* , dual de la fibre de E .

1.3. Morphismes de rang constant entre fibrés vectoriels

PROPOSITION Soit $f : E \rightarrow E'$ un morphisme (au-dessus de \mathbb{I}_B) de fibrés vectoriels de même base B . Supposons que l'application linéaire f_b soit de rang r indépendant de b .

Alors, la famille de sous-espaces vectoriels $(\text{Ker } f_b \subset E_b)_{b \in B}$ [resp. $(\text{Im } f_b \subset E'_b)_{b \in B}$] constitue la famille des fibres d'un sous-fibré vectoriel de E [resp. E'] noté $\text{Ker } f$ [resp. $\text{Im } f$] et est appelé **noyau** [resp. **image**] de f .

On appelle **conoyau** de f le fibré vectoriel quotient :

$$\text{Coker } f = E' / \text{Im } f$$

(dont les fibres sont $\text{Coker } f_b = E'_b / \text{Im } f_b$).

Preuve

Comme on l'a fait pour prouver la proposition du paragraphe 1.1, on se ramène par trivialisations locales au cas d'un morphisme de \mathbb{K}_U^n dans \mathbb{K}_U^p ($n = \text{rang } E$, $p = \text{rang } E'$), qu'on peut écrire sous la forme :

$$y_i = \sum_{j=1}^n a_{ij}(b)x_j \quad i = 1, \dots, p, \tag{*}$$

où $(a_{ij}(b))$ est une matrice dépendant continûment de b , de rang constant r .

Soit $b_0 \in U$. Quitte à permuter les coordonnées x_1, \dots, x_n on peut supposer que les r premières colonnes de la matrice $a(b)$ sont linéairement indépendantes pour $b = b_0$, donc aussi pour b assez proche de b_0 par raison de continuité. Quitte à restreindre U on peut donc supposer cette condition vérifiée sur U tout entier. Le $(n - r)$ -uplet (x_{r+1}, \dots, x_n) est alors pour tout $b \in U$ un système de coordonnées linéaires sur $\text{Ker } f_b$, et définit une trivialisations locale (sur U) de $\text{Ker } f$.

D'autre part, les r premières colonnes de la matrice $a(b)$ forment, pour tout $b \in U$, une base de l'image de l'application linéaire (*), et l'on en déduit une famille de sections de E'_U , linéairement indépendantes et engendrant $\text{Im } f_b$ pour tout $b \in U$. □

2. Espaces fibrés, dans le contexte différentiel

2.1. La différentiabilité rentre en scène

Comme le contexte qui nous intéresse est celui des variétés différentielles, il est naturel de se demander ce qui reste des considérations précédentes quand on remplace la catégorie des espaces topologiques par celle des variétés différentielles. Relisons dans cet esprit le paragraphe 0.

[0.0] : la projection $\pi_E : |E| \rightarrow B$ est maintenant supposée être un morphisme de variétés différentielles. Mais on aimerait que les fibres E_b soient non seulement des sous-espaces topologiques mais des *sous-variétés* de $|E|$.

La façon la plus simple d'en être sûr est d'imposer à π_E d'être une *submersion*, ce que nous ferons désormais.

En résumé, nous appellerons **variété au-dessus de B** une famille de variétés $(E_b)_{b \in B}$ dont l'espace total $|E|$ a été muni d'une structure de variété différentielle telle que la projection π_E soit une submersion, la structure de variété des E_b étant induite par leur statut de sous-variétés de $|E|$.

[0.1] : Rien à changer dans la définition des *morphismes*, si ce n'est que les applications doivent être supposées *différentiables* au lieu d'être seulement continues.

[0.2] Produits fibrés
 [0.3] Changements de base } : pour démontrer que la catégorie des variétés projetées est stable par ces deux types d'opérations, il suffit de démontrer la proposition suivante, ce qui est un excellent exercice d'application du théorème des fonctions implicites.

PROPOSITION Soient $\pi_i : |E_i| \rightarrow B$ ($i = 1, 2$) deux morphismes de variétés différentielles tels que l'un des deux soit une submersion. Alors, en notant $E_i = (|E_i|, \pi_i)$ les deux espaces projetés correspondants, $\left| E_1 \times_B E_2 \right|$ est une sous-variété de $|E_1| \times |E_2|$, et $\pi_1 \times_B \pi_2$ est une submersion.

[0.4] : aucune difficulté d'adaptation. On obtient ainsi la notion de **variété fibrée de base B et de fibre F** , où B et F sont des variétés différentielles.

En adaptant de la même façon le paragraphe 1, on obtient la catégorie des **fibrés vectoriels sur une variété différentielle** : les morphismes de cette catégorie sont des applications différentiables, dont la restriction aux fibres est linéaire.

Comme exemples importants de fibrés vectoriels sur une variété différentielle, regardons ceux qu'on peut construire à partir du fibré tangent.

2.2. Fibrés et morphismes liés au fibré tangent

Au chapitre 5, nous avons appris à construire, sur toute variété différentielle M , un fibré vectoriel TM de rang $n = \dim M$, le *fibré tangent* (chap. 5, § 3). Son dual est le *fibré cotangent* T^*M (chap. 5, § 4).

Remarque sur les notations Dans les exemples qui suivent, nous nous conformons à l'usage courant consistant à désigner par la même notation un espace fibré et son espace total (abandonnant donc la notation $|\cdot|$ pour l'espace total, notation introduite

dans ce cours uniquement dans un but pédagogique). Cette identification du fibré à son espace total est sans danger pour les espaces fibrés dont la projection est « canonique » (c'est-à-dire évidente par le contexte).

Morphismes tangent et cotangent

À tout morphisme $f : M \rightarrow N$ de variétés différentielles nous savons associer (cf. chap. 5) une famille d'applications linéaires :

$$(T_a f : T_a M \rightarrow T_{f(a)} N)_{a \in M}.$$

les applications tangentes à f aux différents points de M . Il est clair, d'après leurs expressions en coordonnées locales, que ces applications « se recollent » différentialement pour former un *morphisme de fibrés vectoriels* :

$$\begin{array}{ccc} TM & \xrightarrow{Tf} & TN \\ \downarrow & & \downarrow \\ M & \xrightarrow{f} & N \end{array}$$

(« morphisme au-dessus de f »), ou – ce qui revient au même d'après la propriété universelle du changement de base – un morphisme de fibrés vectoriels de base M :

$$\begin{array}{ccc} TM & \xrightarrow{Tf_M} & f^*(TN) \\ & \searrow & \swarrow \\ & M & \end{array}$$

De même le recollement des applications cotangentes aux différents points de M fournit un morphisme *transposé* du précédent (en un sens que nous laissons le lecteur expliciter).

$$\begin{array}{ccc} T^*M & \xleftarrow{T^*f_M} & f^*(T^*N) \\ & \searrow & \swarrow \\ & M & \end{array}$$

À titre d'illustration de la « machine » des fibrés vectoriels, faisons l'exercice d'exprimer dans les notations précédentes la notion d'*image réciproque d'une forme différentielle de degré 1* (chap. 5, § 4.2).

La réponse est résumée par le diagramme suivant :

$$\begin{array}{ccccc} T^*M & \xleftarrow{T^*f_M} & f^*(T^*N) & \xrightarrow{f^{T^*N}} & T^*N \\ & \searrow f^*\omega & \downarrow f^*\omega & & \downarrow \omega \\ & & M & \xrightarrow{f} & N \end{array}$$

Commentaire explicatif À toute section ω du fibré T^*N est associée une section $f^*\omega$ du fibré $f^*(T^*N)$, définie par $(f^*\omega)(a) = (a, \omega(f(a)))$ (cf. § 0.3, « section induite »). L'image réciproque de ω par f est la section de T^*M définie par $f^*\omega = T^*f_M(f^*\omega)$. (L'usage de la notation $f^*\omega$ pour l'image réciproque de ω – notée $f^*\omega$ au chapitre 5 – est une « ruse typographique » pour ne pas la confondre avec la section induite $f^*\omega$; aimant trop la notation f^* les mathématiciens l'utilisent à tout bout de champ, et seul le contexte permet de s'y retrouver !)

2.3. Fibrés normal et conormal à une sous-variété $S \subset M$

Soit S une sous-variété de M . TS est alors un sous-fibré vectoriel de TM_S (restriction à S de TM). On note $T_S M$, et l'on appelle **fibré normal à S dans M** , le fibré vectoriel quotient :

$$T_S M := TM_S / TS.$$

Le dual du fibré normal est appelé **fibré conormal à S dans M** , et noté $T_S^* M$.

Par transposition de l'égalité :

$$T_S M = \text{coker } Tj,$$

où $j : S \rightarrow M$ désigne l'inclusion canonique, on obtient l'égalité :

$$T_S^* M = \text{Ker } Tj_S^*,$$

selon laquelle le fibré conormal est l'ensemble des covecteurs (cotangents à M aux points de S) qui s'annulent en restriction à TS .

Interprétation en termes d'équations locales Pour tout $a \in S$, $T_a S$ est un sous-espace vectoriel de $T_a M$, défini par l'annulation de p covecteurs ds_{1a}, \dots, ds_{pa} (où (s_1, \dots, s_p) est un système d'équations locales de S dans M) ; s'annulant sur $T_a S$, ces p covecteurs peuvent être considérés comme un p -uplet de formes linéaires sur l'espace vectoriel $T_a M / T_a S$ (donc un système de coordonnées linéaires sur cet espace puisque celui-ci est de dimension p et que les p formes sont linéairement indépendantes). Autrement dit $(ds_{1a}, \dots, ds_{pa})$ est une base de l'espace vectoriel conormal, dual de l'espace normal.

De cette discussion on pourra retenir la conclusion pratique suivante : la variété $T_S M$ (espace total du fibré normal) admet un système de coordonnées locales de la forme :

$(x_1, \dots, x_\ell$;	$ds_1, \dots, ds_p)$
coordonnées		coordonnées linéaires
dans la base		dans les fibres

où (x_1, \dots, x_ℓ) est un système de coordonnées locales de S et (s_1, \dots, s_p) un système d'équations locales de S dans M ; dans ces coordonnées la projection de $T_S M$ sur sa base S s'exprime tout simplement par la première projection sur \mathbb{R}^ℓ (l'espace des x_1, \dots, x_ℓ).

Pourquoi les notations $T_S M, T_S^* M$? Dans le cas particulier d'une sous-variété réduite à un point ($S = \{a\}$), TM_S se réduit à $T_a M$, et TS à $\{0\}$ (espace vectoriel nul) ; le

« fibré » normal à un point $\{a\}$ s'identifie donc à $T_a M$ (muni de la projection constante sur a !), ce qui explique pourquoi on a choisi la notation $T_S M$ pour désigner, plus généralement, le fibré normal à une sous-variété S .

La même remarque s'applique au fibré conormal, celui-ci s'identifiant à $T_a^* M$ dans le cas $S = \{a\}$.

EXERCICE TEST Quelle relation logique y a-t-il entre les deux propriétés suivantes d'une sous-variété $S \subset M$?

- i) S est définie dans M par un système d'équations globales ;
- ii) le fibré normal $T_S M$ est trivial.

Exemple (cf. introduction du chapitre) Soit Ω la variété des droites affines de \mathbb{R}^n , et soit $\Gamma_0 \subset \Omega$ la sous-variété constituée par les droites passant par l'origine. Cette sous-variété peut-elle être définie dans Ω par une équation globale ?

2.4. Fibrés projectifs et fibrés en grassmanniennes

Soit E un fibré vectoriel de rang r sur une variété différentielle M . Pour tout $p = 1, 2, \dots, r-1$, la construction du chapitre 4, § 1.3, appliquée aux fibres de E , fournit une famille paramétrée par $a \in M$ de variétés grassmanniennes $G_p(E_a)$ (variété des p -plans vectoriels de l'espace vectoriel E_a).

Par une adaptation évidente de la construction faite au chapitre 4, on munit l'espace total de cette famille d'une structure de variété différentielle qui en fait une variété fibrée de base M , le fibré en grassmanniennes $BG_p(E)$.

Cas particuliers importants

- $p = 1$: $G_1(E)$ est par définition le fibré projectif $\mathbb{P}(E)$
- $p = r - 1$: $G_{r-1}(E) = \mathbb{P}(E^*)$

ÉTUDE 6 Enveloppes des familles de courbes

La notion d'enveloppe est un magnifique exemple de notion géométrique très concrète, directement accessible à l'imagination mais extrêmement délicate à cerner avec précision. Depuis les débuts du calcul infinitésimal, les géomètres n'ont cessé de s'y intéresser ; et pourtant il a fallu attendre 1962 (trois siècles après Huygens !) pour voir se dessiner une théorie générale rigoureuse (R. Thom, article déjà cité dans l'étude 5).

Les exemples très simples de la figure 3 suffisent à situer le problème. Tous ces exemples représentent des familles à un paramètre de courbes planes, en l'occurrence des cercles ou des droites (en pointillé sur la figure). Les courbes en trait plein sont les enveloppes de ces familles (courbes fixes auxquelles toutes les courbes en pointillé sont tangentes).

La famille de cercles de la figure 3 i) a pour enveloppe deux cercles concentriques. Dans le cas de la figure 3 ii), le plus petit de ces deux cercles « dégénère » en un point. La figure 3 iii) est également un exemple dégénéré d'enveloppe, cette fois pour une famille de droites. On a représenté sur la figure 3 iv) une façon possible de faire disparaître la dégénérescence, par une déformation aussi petite qu'on veut de la

famille de droites ; l'enveloppe est maintenant une courbe, mais cette courbe n'est pas lisse : elle a trois points de rebroussement.

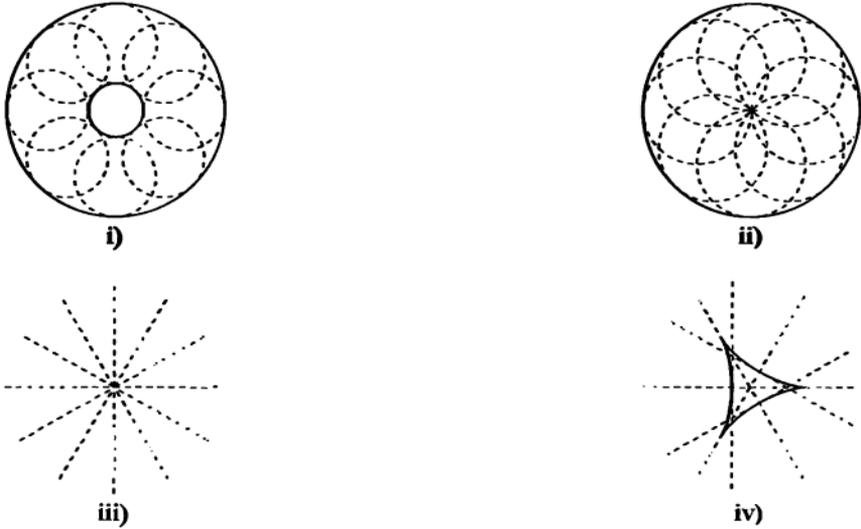


Figure 3.

Un mathématicien aime à ne se poser de problèmes que sur des objets préalablement définis avec précision. La première question est donc : qu'est-ce qu'une enveloppe ? Nous avons eu l'air d'en formuler une définition quelques lignes plus haut quand nous avons écrit : « une courbe fixe à laquelle toutes nos courbes variables sont tangentes ». Cette définition nous interdit donc de parler d'enveloppe dans les cas dégénérés 3 ii) et 3 iii), puisqu'un point n'est pas une courbe ! Quant au cas 3 iv), si l'on veut appeler enveloppe la courbe en traits pleins il faudra expliquer ce que signifie « être tangente à une courbe » lorsque cette courbe n'est pas lisse ! Supposant que nous ayons trouvé une telle définition, se posera alors le problème de savoir si les cas « dégénérés » (comme ceux des figures 3 ii) et 3 iii)) sont bien exceptionnels, en ce sens qu'on peut toujours les faire disparaître par une déformation arbitrairement petite de la famille. Et les singularités de la courbe enveloppe, comme les trois points de rebroussement de la figure 3 iv), peut-on eux aussi les faire disparaître ? L'étude 5 suggère que non, du moins dans le cas des familles de droites. Mais qu'en est-il des familles de courbes plus générales ?

C'est à toute cette série de questions que nous allons essayer de répondre. Mais avant de commencer, voici une « devinette ».

Devinette Parmi les cinq courbes suivantes (Fig. 4), trois ne sont pas des enveloppes de familles de droites proches de celle de la figure 2 iii). Lesquelles ?



Figure 4.

Après cette devinette qui ne fait appel qu'à l'intuition, voici un exercice qui peut être résolu rigoureusement, à condition d'avoir bien compris l'étude 5 sur les familles de droites du plan.

EXERCICE recommandé Donner des équations explicites de familles de droites du plan arbitrairement proches de celle de la figure 3 iii), et dont l'enveloppe est une courbe comme celle de la figure 3 iv). Pouvez-vous construire d'autres exemples analogues, ayant un nombre différent de points de rebroussement ? Ce nombre de points de rebroussement peut-il être prescrit arbitrairement à l'avance ?

(N.B. : Par « point de rebroussement » nous entendons évidemment « point de rebroussement ordinaire », au sens de l'étude 5.)

0 Familles de courbes

Notre objectif est d'approfondir, en la généralisant aux familles de courbes, l'étude 5 sur les enveloppes de familles de droites du plan. N'ayant plus affaire à des droites nous n'avons plus de raisons de supposer que l'espace ambiant est un plan (affine ou projectif) : ce sera, plus généralement, une « surface », c'est-à-dire une variété à deux dimensions, que nous noterons M . Notre objet d'étude sera donc une famille de courbes lisses de M , famille à un paramètre, c'est-à-dire ayant comme espace des paramètres une variété à une dimension. Mais pour pouvoir faire disparaître les « dégénérescences » éventuelles de cette famille, nous voulons nous réserver la possibilité de la déformer à l'intérieur d'une « grande » famille que nous appellerons « famille universelle », paramétrée par une variété Ω de dimension plus grande ; par exemple, dans un problème d'enveloppe d'une famille de droites, on pourra prendre pour Ω la variété de toutes les droites du plan ; dans un problème d'enveloppe d'une famille de cercles la variété de tous les cercles du plan, etc.

0.1 Famille de courbes lisses Prenant donc comme espace des paramètres une variété Ω de dimension quelconque, définissons ce qu'est pour nous une **famille de courbes lisses de M , paramétrée par Ω** : c'est une famille $(S(u))_{u \in \Omega}$, où les $S(u)$ sont des courbes lisses de M vérifiant l'hypothèse suivante : soit

$$S = \{(a, u) \in M \times \Omega \mid a \in S(u)\}$$

l'« espace total » de la famille ; alors S est une hypersurface lisse de $M \times \Omega$, et la projection :

$$\pi_S := pr_2 \mid S : S \rightarrow \Omega$$

est une submersion (dans beaucoup d'exemples intéressants ce sera même une fibration).

Les fibres $\pi_S^{-1}(u)$ de cette submersion (évidemment isomorphes à $S(u)$ par la projection canonique $g := pr_1 \mid S$) seront notées S_u .

$$\begin{array}{ccccc}
 & & M \times \Omega & & \\
 & \swarrow pr_1 & & \searrow pr_2 & \\
 M & & S & & \Omega \\
 \leftarrow g & & \xrightarrow{\pi_S} & & \\
 \cup & & \cup & & \cup \\
 S(u) & \xleftarrow{\cong} & S_u & \xrightarrow{\quad} & \{u\}
 \end{array}$$

0.2 Famille induite Muni de la projection π_S , l'espace S est un « espace au-dessus de Ω », dont l'espace induit par tout changement de base différentiable :

$$\gamma : \Gamma \rightarrow \Omega$$

définit une famille paramétrée par Γ de courbes lisses de M (famille à un paramètre si Γ est une variété à une dimension).

Exemples

i) Prenons pour S la famille de tous les cercles de \mathbb{R}^2 de rayon donné ; la position d'un cercle de la famille peut être repérée par la position de son centre $u \in \Omega = \mathbb{R}^2$; l'espace projeté S est alors un fibré trivial de base \mathbb{R}^2 et de fibre S^1 . En astreignant le centre des cercles à rester sur une courbe $\Gamma \subset \Omega$, on obtient une famille à un paramètre de cercles, qui est un fibré (trivial) de base Γ (le « fibré induit » S_Γ) : cf. par exemple les figures 3 i) et ii), où la courbe Γ est un cercle.

ii) Prenons maintenant pour S la famille de toutes les droites affines de \mathbb{R}^2 . Ω est maintenant la variété (à deux dimensions) de toutes les droites affines de \mathbb{R}^2 , et S est un fibré en droites sur Ω (dit « fibré tautologique », car la fibre $S_u \approx S(u)$ au-dessus d'un « point » $u \in \Omega$ est u lui-même, considéré comme sous-ensemble de \mathbb{R}^2 : un « point » de Ω est une « droite » de \mathbb{R}^2 !).

La figure 3 iii) [resp. iv)] représente une sous-famille à un paramètre de S , que l'on peut paramétrer par \mathbb{P}^1 , et qui est induite par le plongement $\gamma_0 : \mathbb{P}^1 \rightarrow \Omega$ [resp. un plongement voisin $\gamma : \mathbb{P}^1 \rightarrow \Omega$] : cf. les notations de l'introduction du chapitre 6. Remarquons que la famille ainsi induite γ_0^*S [resp. γ^*S] est un fibré en droites non trivial sur \mathbb{P}^1 (pourquoi ?) – ce qui, au passage, démontre que S était un fibré en droites non trivial sur Ω .

On peut remarquer aussi que l'espace total du fibré S n'est autre que la variété des éléments de contact de $M = \mathbb{R}^2$, notée M à la fin de l'étude 5.

1 Points caractéristiques

1.1 Dérive le long d'un vecteur $\delta_u \in T_u \Omega$

Donnons-nous un vecteur $\delta_u \in T_u \Omega$ (vitesse du « point mobile » u), et essayons de « mettre en mouvement » le couple (a, u) dans la variété S avec une vitesse initiale (v_a, δ_u) dont la seconde composante soit précisément le vecteur δ_u donné.

PROPOSITION

La première composante v_a est alors définie de façon unique modulo $T_a S(u)$. Sa classe modulo $T_a S(u)$ définit donc un vecteur « normal » à $S(u)$ en a , que nous noterons

$$\delta_u^S(a) \in T_a M / T_a S(u)$$

et que nous appellerons **dérive de $S(u)$ en a le long de δ_u** (Fig. 5).

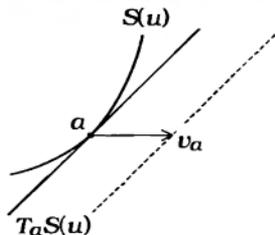


Figure 5. La dérive de $S(u)$ en a est la classe de v_a modulo $T_a S(u)$.

(N.B. : Le mot « dérive » est choisi pour suggérer l'idée d'un bateau qui dérive, poussé par le courant : la dérive de $S(u)$ en a mesure la façon dont la courbe $S(u)$, considérée au voisinage de a , s'écarte de sa position initiale quand u évolue avec la vitesse δ_u .)

Expression de la dérive en coordonnée normale On trouvera au début du paragraphe 1.3 une preuve géométrique de la proposition ci-dessus. Voici, en attendant, une preuve par le calcul, qui a l'avantage de fournir une expression de la dérive « en coordonnée normale ».

Soit F une équation locale de l'hypersurface $S \subset M \times \Omega$, et soit $F_a^u : M, a \cdots \rightarrow \mathbb{R}$ le germe de fonction sur M défini par restriction de F à $M \times \{u\}$: c'est une équation locale de $S(u)$; sa différentielle dF_a^u peut donc être considérée comme une coordonnée normale de S en a .

Dire que (v_a, δ_u) est tangent à S , c'est-à-dire que $dF(v_a, \delta_u) = 0$, ou encore que $dF(v_a, 0) = -dF(0, \delta_u)$.

Cette équation peut encore s'écrire :

$$dF_a^u(v_a) = -\delta_u F, \quad (*)$$

où nous avons noté $\delta_u F$ la dérivée directionnelle de F le long du vecteur δ_u (ou plutôt $(0, \delta_u)$). Il s'agit d'une équation linéaire, dont la solution v_a est unique modulo $\text{Ker } dF_a^u = T_a S(u)$, comme annoncé par la proposition. Cette solution définit donc un unique vecteur normal en a à $S(u)$ (la « dérive »), dont l'équation (*) donne l'expression « en coordonnée locale ».

Dans des coordonnées locales (x, y) de M où $dx|_{TS(u)} \neq 0$ (c'est-à-dire où l'axe des y est transverse à $S(u)$), on peut supposer que l'équation locale F de S est « résolue par rapport à y », autrement dit de la forme :

$$F = y - f(x, u).$$

L'expression (*) de la dérive s'écrit alors :

$$dF_a^u(v_a) = -\delta_u F = \delta_u f(x(a), u).$$

1.2 Points caractéristiques dans une direction δ_u

Supposons choisi un vecteur non nul $\delta_u \in T_u \Omega$. Quand a parcourt $S(u)$, la dérive $\delta_u^S(a)$ dépend différemment de a , en ce sens qu'elle définit une section (différentiable) du fibré normal à $S(u)$:

$$\delta_u^S : S(u) \rightarrow T_{S(u)} M$$

(en effet la fonction $x \mapsto \delta_u f(x, u)$, qui d'après ce qu'on vient de voir donne l'expression locale de δ_u^S en coordonnée normale, est fonction différentiable de x ; cf. aussi paragraphe 1.3 pour une preuve géométrique).

DÉFINITION On dit que le point a de la courbe $S(u)$ est un **point caractéristique pour la famille S , dans la direction δ_u** , si $\delta_u^S(a) = 0$.

C'est un **point caractéristique simple** si c'est un zéro simple de la section δ_S^u , c'est-à-dire si la courbe image de cette section est transverse en a à la section nulle du fibré $T_{S(u)}M$.

Traduction en coordonnées locales Dans des coordonnées locales (x, y) de M où S est donnée par l'équation résolue en y

$$y = f(x, u),$$

le point de $S(u)$ de coordonnées (x, y) est caractéristique dans la direction δ_u si et seulement si

$$\delta_u f(x, y) = 0 ;$$

il est caractéristique simple si et seulement si, de plus,

$$\partial_x \delta_u f(x, u) \neq 0.$$

Remarque Comme la section δ_S^u dépend linéairement de δ_u , la propriété pour un point a d'être caractéristique [resp. caractéristique simple] ne dépend que de la direction dans $T_u\Omega$ du vecteur δ_u .

Exemple i) : familles de droites Reprenons l'exemple 0.2 ii) (famille de toutes les droites du plan). Dans les coordonnées locales $((x, y) ; p, q)$ de $M \times \Omega$ où S est donnée par l'équation $y = px + q$ ($=: f(x, p, q)$), la dérive le long d'un vecteur $\delta \in T\Omega$ s'exprime en coordonnée normale par :

$$\delta f = x dp(\delta) + dq(\delta).$$

Un point caractéristique (dans la direction δ) est un point où cette expression s'annule, ce qui n'est possible que si $dp(\delta) \neq 0$ (à moins que $dq(\delta)$ ne s'annule aussi, ce qui voudrait dire que δ est le vecteur nul).

Conclusion Si $\delta \neq 0$, la droite $S(u)$ contient un unique point caractéristique ; celui-ci est simple ($\partial_x \delta f = dp(\delta) \neq 0$), et son abscisse est donnée par $x = -dq(\delta)/dp(\delta)$.

Exemple ii) : familles de cercles de rayon constant Prenons pour S la famille de tous les cercles de \mathbb{R}^2 de rayon donné r , chaque cercle $S(u)$ étant repéré par la position de son centre $u \in \mathbb{R}^2$ ($=: \Omega$). On peut alors prendre pour équation de S (dans $\mathbb{R}^2 \times \mathbb{R}^2$) la fonction :

$$F(a, u) := \overrightarrow{u\vec{a}}^2 - r^2,$$

où $\overrightarrow{u\vec{a}}^2 = \overrightarrow{u\vec{a}} \cdot \overrightarrow{u\vec{a}}$ (« carré scalaire » du vecteur $\overrightarrow{u\vec{a}}$). Appliquant à cette fonction l'opérateur de dérivation directionnelle le long de δ_u (vecteur vitesse de u) on trouve pour expression de la dérive en a :

$$-\delta_u F = 2\overrightarrow{u\vec{a}} \cdot \delta_u$$

(on a utilisé le fait que $\delta_u(\overrightarrow{u\vec{a}}) = -\delta_u(\overrightarrow{a\vec{u}}) = -\delta_u$).

Le point a du cercle $S(u)$ est donc caractéristique dans la direction δ_u si et seulement si $\overrightarrow{u\vec{a}} \perp \delta_u$. Il est toujours caractéristique simple (pourvu que $\delta_u \neq 0$) : en effet il est clair que la fonction $a \mapsto x(a) = \overrightarrow{u\vec{a}} \cdot \delta_u$ est une coordonnée locale de $S(u)$ au voisinage d'un point caractéristique a_0 (cf. Fig. 6).

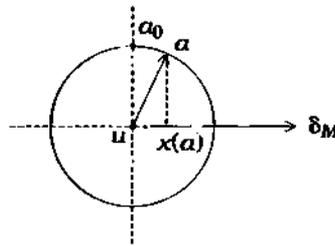


Figure 6.

Exemple iii) : translation rectiligne uniforme d'une courbe plane Dans le plan $M = \mathbb{R}^2$ muni des coordonnées (x, y) , on considère une famille de courbes lisses $(S(t))_{t \in \mathbb{R}}$ telles que $S(t)$ soit déduite de $S(0)$ par la translation de vecteur $(t, 0)$. Alors un point (x_t, y_t) de $S(t)$ est caractéristique (dans la direction ∂_t) si et seulement si la tangente à $S(t)$ en ce point est parallèle à l'axe des x ; il est caractéristique simple si et seulement si $S(t)$ est en contact quadratique avec sa tangente en ce point (exercice : le démontrer). La figure 7 illustre géométriquement ces affirmations pour une courbe $S(0)$ de la forme $y = f(x)$ (cas auquel on peut toujours se ramener au voisinage d'un point caractéristique) ; on y a représenté de deux manières le champ des dérivés :
 1° par le champ (horizontal) des vitesses de translation (schémas de gauche) ;
 2° par le champ orthogonal à $S(0)$ qui est équivalent au précédent modulo $TS(0)$ (schémas de droite) : cette seconde représentation permet de visualiser la transversalité ou non-transversalité de la dérive.

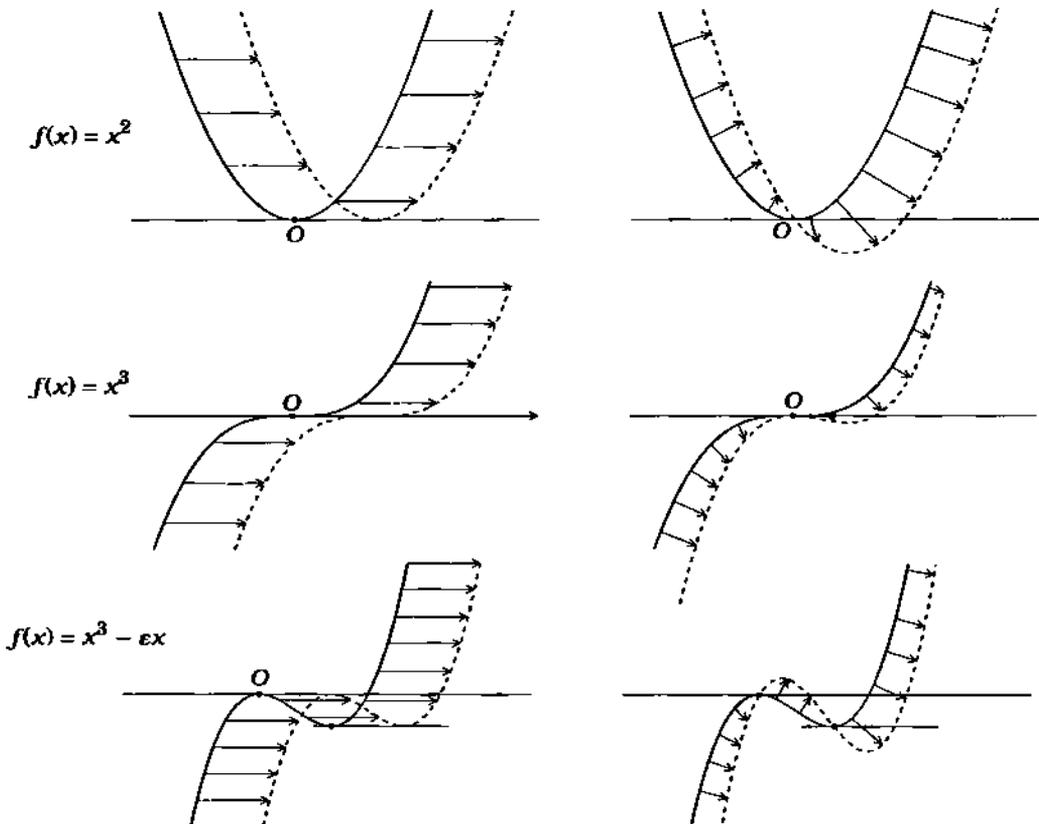


Figure 7. Reconnaître les points caractéristiques simples.

1.3 Compléments sur la dérive

Nous nous proposons ici de retrouver et de préciser les propriétés de la dérive par des méthodes purement géométriques. Aucune hypothèse sur la dimension de la variété M , ou sur la codimension de S , n'est utilisée, de sorte qu'on pourra appliquer ce qui suit à l'étude des enveloppes de variétés de dimension (et codimension) arbitraire.

Preuve de la proposition du paragraphe 1.1

(cf. diagramme final du paragraphe 0.1). Pour tout $s = (a, u) \in S$, l'application tangente $T_s g$ induit un isomorphisme entre $T_s S_u$ et $T_a S(u)$, et passe donc au quotient par ces sous-espaces. Or $T_s S/T_s S_u$ est isomorphe à $T_u \Omega$ (par $T_s \pi_S$).

On obtient ainsi une application linéaire :

$$\nabla_s^S : T_u \Omega \rightarrow T_a M/T_a S(u)$$

que nous appellerons *dérive en s* de la famille (S, π_S) . Le vecteur normal noté $\delta_u^S(a)$ dans la proposition 1.1 n'est autre que $\nabla_s^S(\delta_u)$. □

EXERCICE Vérifier que la surjectivité [resp. l'injectivité] de ∇_s^S équivaut à celle de $T_s g$.

Dépendance différentiable en s Nous nous proposons de montrer que la dérive ∇_s^S dépend différentiablement de $s = (u, a)$. Comme il s'agit d'une application linéaire dont l'espace source et l'espace but dépendent de s , dire qu'elle dépend différentiablement de s n'a de sens que si nous pouvons considérer la source, ainsi que le but, comme la fibre d'un fibré vectoriel de base S .

Pour la source, c'est facile : la famille des $(T_u \Omega)_{s=(a,u) \in S}$ est la famille des fibres du fibré $T\Omega_S$, fibré induit de $T\Omega$ par le changement de base $\pi_S : S \rightarrow \Omega$.

Pour le but, le recollement des espaces normaux $(T_a M/T_a S(u))_{(a,u) \in S}$ va se faire grâce au lemme suivant.

Lemme L'inclusion $T_a M \rightarrow T_s(M \times \Omega)$ ($s = (a, u)$) induit par passage au quotient un isomorphisme $T_a M/T_a S(u) \xrightarrow{\cong} T_s(M \times \Omega)/T_s S$.

Preuve

Le conoyau de cette inclusion n'est autre que $T_u \Omega$; or, il en est de même du conoyau de l'inclusion $T_a S(u) \rightarrow T_s S$. □

Corollaire Les espaces normaux $T_a M/T_a S(u)$ s'identifient aux fibres de $T_S(M \times \Omega)$. Avec cette identification, la dérive en a le long de δ_u s'identifie à la classe de $0 \oplus (-\delta_u)$ modulo $T_s S$.

Preuve

Dire que $(v_a, \delta_u) \in T_s S$, c'est dire que $v_a \oplus 0 = 0 \oplus (-\delta_u)$ modulo $T_s S$. □

Conclusion Définissons le morphisme de fibrés vectoriels

$$\nabla^S : T\Omega_S \rightarrow T_S(M \times \Omega)$$

par

$$\nabla^S = -\pi \circ i,$$

où i et π sont respectivement l'inclusion et la projection canonique :

$$\begin{aligned} i &: T\Omega_S \rightarrow T(M \times \Omega)_S = TM_S \oplus T\Omega_S \\ \pi &: T(M \times \Omega)_S \rightarrow T_S(M \times \Omega) = T(M \times \Omega)_S / TS. \end{aligned}$$

Alors la dérive ∇^S s'identifie à la restriction, au-dessus de $s \in S$, de ce morphisme ∇^S . C'est donc bien ce qu'on peut appeler une famille différentiable (par rapport à s) d'applications linéaires.

2 Enveloppes

Dans tout ce paragraphe, S sera une famille à un paramètre de courbes de la surface M . Nous en noterons Γ (au lieu de Ω) l'espace de base.

Comme $T_u\Gamma$ sera de dimension 1, nous omettrons désormais la mention « dans la direction δ_u » quand nous parlerons de points caractéristiques.

2.1 Étude « à la source » du lieu caractéristique

Le point de vue du paragraphe 1.3 conduit naturellement à étudier les points caractéristiques à la source, c'est-à-dire dans S , plutôt qu'au but, c'est-à-dire dans M .

Au lieu de dire que le point a de $S(u)$ est caractéristique pour S , nous dirons donc désormais que $s = (a, u)$ est un point caractéristique de S . Nous appellerons **lieu caractéristique** le sous-ensemble $C_S \subset S$ formé par les points caractéristiques.

PROPOSITION

Le lieu caractéristique C_S coïncide avec l'ensemble critique de $g : S \rightarrow M$.

Preuve

N'oublions pas que $g|_{S_u}$ est un plongement (ayant pour image $S(u)$), de sorte que pour tout $s \in S$ l'application tangente $T_s g$ est de rang au moins égal à 1. Le lieu caractéristique, ensemble des points s où l'image de $T_s g$ coïncide avec $T_a S(u)$ ($a = g(s)$), est donc précisément l'ensemble des points où $T_s g$ est de rang 1. \square

Conséquences de l'hypothèse de caractéristique simple Il résulte immédiatement de l'hypothèse de caractéristique simple (Déf. 1.2) que :

- i) le lieu caractéristique C_S est une courbe lisse ;
- ii) l'application $\pi_S|_{C_S} : C_S \rightarrow \Gamma$ est un difféomorphisme local.

Autrement dit, les points caractéristiques sont isolés dans chaque fibre S_u , et dépendent différentiablement de $u \in \Gamma$.

Sous cette hypothèse, dont nous verrons plus loin qu'elle est « générique » (§ 3), la courbe lisse C_S est ce que Thom appelle **enveloppe à la source** de la famille de courbes $(S(u))_{u \in \Gamma}$.

Son image par g ,

$$C = g(C_S),$$

qui n'est autre que le « contour apparent » de la surface S par la projection g (cf. fin du chap. 2), est ce que Thom appelle **enveloppe au but**. Mais les exemples de la figure 3 (cas ii) et iii)) montrent que l'enveloppe au but n'est pas toujours une enveloppe au

sens géométrique que nous avons donné à ce mot dans l'introduction (« enveloppe géométrique » dans la terminologie de Thom).

EXERCICE TEST Décrire l'enveloppe à la source et l'enveloppe au but de chacun des exemples de la figure 3 (tous à caractéristique simple d'après le paragraphe 1.2, exemples i) et ii)).

Remarque En fait, Thom travaille avec une hypothèse plus faible que celle de caractéristique simple, qu'on peut formuler de l'une des deux façons suivantes.

i) L'application :

$$s \mapsto T_s g \in \mathcal{L}(T_s S, T_{g(s)} M)$$

est transverse dans $\mathcal{L}(TS, TM_S)$ au sous-fibré constitué par les applications linéaires de rang exactement égal à 1 ;

ii) (formulation équivalente) : l'application $\delta^S : S \rightarrow T_S(M \times \Gamma)$, dérive de S le long de δ , champ de vecteurs non nul de Γ , est transverse à $[0]$ = image de la section nulle du fibré $T_S(M \times \Gamma)$.

EXERCICE Vérifier l'équivalence de ces deux hypothèses (on pourra les traduire en coordonnées locales). En quoi sont-elles plus faibles que l'hypothèse de caractéristique simple (donner un exemple), et quelle information nous donnent-elles sur le lieu caractéristique C_S ?

2.2 Hypothèse générique, garantissant l'existence de l'enveloppe géométrique

Quand s parcourt C_S , l'application tangente $T_s g$ est de rang constant égal à 1, de sorte que son image $\text{Im } T_s g$ est une droite vectorielle de $T_a M$ ($a = g(s)$) que l'on peut considérer comme la fibre d'un fibré vectoriel de rang 1, de base C_S : nous noterons LC_S ce fibré, qui n'est autre que le fibré vectoriel image par Tg_{C_S} du fibré vectoriel TS_{C_S} (cf. chap. 6, Prop. 1.3).

Supposons que S soit à caractéristique simple, de sorte que C_S est une courbe lisse. En composant Tg_{C_S} avec l'inclusion $TC_S \subset TS_{C_S}$, on obtient un morphisme de fibrés en droites de base C_S :

$$\sigma_S : TC_S \rightarrow LC_S.$$

Intuitivement, σ_S est l'application qui au vecteur vitesse d'un point $s = (a, u)$ mobile le long de C_S associe le vecteur vitesse du point image a , vecteur dont on sait qu'il est porté par la droite $LC_{S,s} = \text{Im } T_s g$.

Dans des coordonnées locales $(x, y; t)$ de $M \times \Gamma$ où la surface S est définie par l'équation $y = f(x, t)$ (et admet donc (x, t) comme système de coordonnées locales), la courbe caractéristique C_S s'obtient en résolvant l'équation implicite $\partial_t f(x, t) = 0$, dont nous noterons $x = x^{car}(t)$ la solution. La fibre en s de LC_S est alors la droite vectorielle de $T_a M$ d'équation $dy = p dx$, où $p = p(t) := \partial_x f(x^{car}(t), t)$. Le fibré LC_S peut donc être muni des coordonnées :

$(t$,	$dx)$
coordonnée locale		coordonnée
de la base		de la fibre

et l'application σ_S est celle qui, au vecteur $\partial_t \in TC_S$, associe le vecteur de $LC_{S,s}$ de coordonnées :

$$dx(\sigma_S(\partial_t)) = dx^{car}(t)/dt.$$

Hypothèse (EG) (« enveloppe géométrique ») Le morphisme σ_S n'a que des zéros simples. c'est-à-dire que l'image par σ_S d'un champ de vecteurs non nul de C_S définit une courbe de LC_S transverse à la section nulle.

Dans les coordonnées locales ci-dessus cela signifie que la fonction dx^{car}/dt n'a que des zéros simples, autrement dit que la fonction x^{car} n'a que des points critiques quadratiques (cf. Fig. 8).

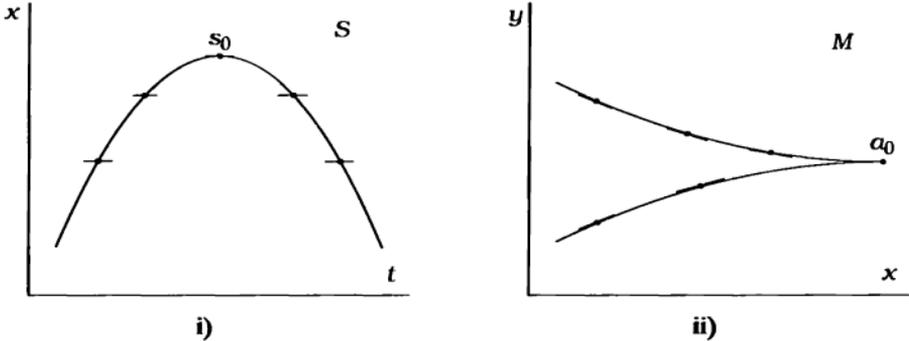


Figure 8. i) La courbe caractéristique, avec le champ de droites $s \mapsto \text{Ker } T_{s,g}$. ii) L'enveloppe au but, avec ses tangentes (les droites $LC_{S,s} = \text{Im } T_{s,g}$).

EXERCICE : points de rebroussement des enveloppes Soit $s_0 = (a_0, u_0)$ un point de S où σ_S s'annule simplement. Montrer qu'avec un choix de coordonnées locales (x, y) de M , centrées en a_0 , tels que $T_{a_0}S(u_0)$ ait pour équation $dy = 0$, la courbe C_S est définie dans $M \times \Gamma$ par des équations de la forme :

$$\begin{aligned} x(t) &= at^2 + o(t^2) & (a \neq 0) \\ y(t) &= bt^3 + o(t^3) & (b \neq 0) \end{aligned}$$

de sorte que son image dans M a un point de rebroussement en a_0 .

Idée : posant $y = f(x, t), p(x, t) = \partial_x f(x, t)$, on commencera par démontrer que la fonction $t \mapsto p(x^{car}(t), t)$ a un zéro simple en $t = 0$.

Conclusion : singularités génériques des enveloppes Sous les hypothèses (CS) (« caractéristique simple ») et (EG) ci-dessus, l'application $g|_{C_S}$ est partout une immersion sauf en des points isolés où elle est du type de l'exercice précédent. Son image C est donc une courbe de M dont les seules singularités sont des points de rebroussement ou des points de « self-intersection » (points où se rencontrent les images de segments disjoints de la courbe C_S).

Cette courbe C est appelée **enveloppe** de la famille de courbes considérée.

2.3 Illustration « à la manière de Huygens »

La figure 9 représente l'enveloppe d'une famille de cercles de rayon constant r dont les centres u parcourent la parabole Γ d'équation $y = x^2/2$. Pour $r < 1$, l'enveloppe se compose de deux courbes lisses disjointes (Fig. 9 i)). Pour $r > 1$ l'une des deux composantes de l'enveloppe acquiert deux points de rebroussement (Fig. 9 ii)). Le « cas limite » $r = 1$ est dégénéré : bien que rien ne le distingue visuellement du cas $r < 1$, ce cas ne vérifie pas l'hypothèse (EG) au point $(a_0 = (0, 1); u_0 = (0, 0))$, et l'enveloppe au point a_0 est seulement de classe C^1 .

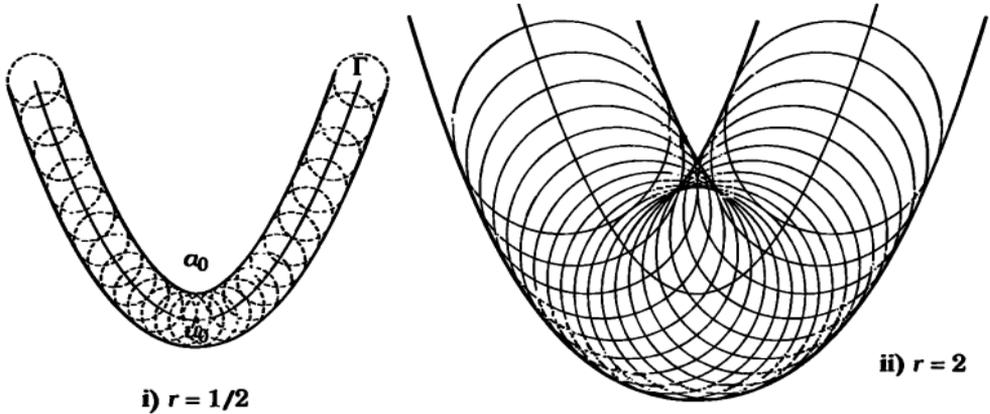


Figure 9.

Quand on dessine le schéma pour diverses valeurs de $r > 1$, on voit les deux points de rebroussement parcourir une courbe qui n'est autre que l'enveloppe des normales à la parabole (Fig. 10). Il est intéressant de comprendre géométriquement pourquoi il en est ainsi : d'après 1.2 (exemple ii)), les deux points caractéristiques du cercle $S(u)$ se trouvent sur la droite $D(u)$ perpendiculaire à Γ en u ; quand u parcourt Γ chacun de ces points $a(u)$ parcourt l'enveloppe avec un vecteur vitesse qui, comme celui de u , est perpendiculaire à $D(u)$ (car la distance r de u à $a(u)$ reste constante) ; comme le point mobile $a(u)$ reste sur la droite mobile $D(u)$, son vecteur vitesse est donc le représentant orthogonal à $D(u)$ de la dérive de la famille de droites $(D(u))_{u \in \Gamma}$; dire que ce vecteur vitesse s'annule, c'est donc dire que $a(u)$ est un point caractéristique de cette famille de droites.

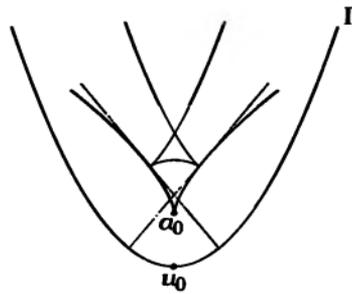


Figure 10.

Affirmation Pour $r \neq 1$ la famille S vérifie l'hypothèse (EG).

La vérification par le calcul ne présente aucune difficulté mais il est plus instructif de raisonner géométriquement.

Désignons par (D, π_D) le fibré en droites de base Γ défini par la famille $(D(u))_{u \in \Gamma}$. Chaque $D(u)$ porte un unique point caractéristique que nous noterons $c(u)$. Quand u parcourt Γ de gauche à droite avec une vitesse non nulle, le vecteur vitesse du point $c(u)$ (toujours porté par la droite $D(u)$) commence par pointer vers le bas (autrement dit vers u), puis pointe vers le haut après franchissement de la valeur u_0 (où il s'annule « simplement », car la famille (D, π) vérifie l'hypothèse (EG)). La fonction distance

de u à $c(u)$ a donc un minimum quadratique en $u = u_0$, de sorte que la courbe caractéristique C_D (enveloppe à la source) de la famille (D, π_D) a la forme représentée sur la figure 11.

Par ailleurs, la courbe caractéristique C_S de S :

$$C_S = D \cap S$$

est la courbe de D définie dans les coordonnées (z, t) de la figure 11 par $z = \pm r$.

Pour $r < 1$, cette courbe ne rencontre pas C_D .

Pour $r > 1$, elle la rencontre transversalement (dans D) en deux points.

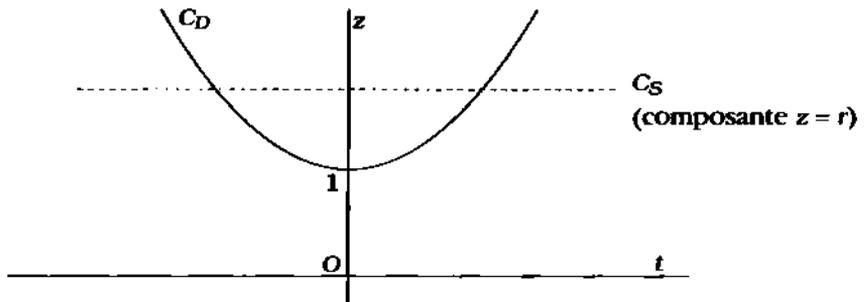


Figure 11. Allure des courbes caractéristiques C_D et C_S , dans le système de coordonnées (z, t) de D ainsi défini :

- t est une coordonnée de la base Γ (par exemple l'abscisse $x(u)$) ;
- z est une coordonnée le long des fibres $D(u)$, définie par la distance du point u au point considéré de $D(u)$, comptée positivement vers le haut.

Cette propriété de transversalité justifie notre affirmation : plus précisément, elle montre que l'application $\sigma_S : TC_S \rightarrow LC_S$ définie en 2.2 ne s'annule jamais pour $r < 1$, et a deux zéros simples pour $r > 1$.

En effet, on a vu que la vitesse du point $a(u)$ est donnée par la dérivée de la famille (D, π_D) , exprimée en coordonnée normale ; or l'expression de cette dérivée est une fonction sur D qui s'annule simplement le long de la courbe C_D , car (D, π_D) est à caractéristique simple (de façon précise, si $\nu : D \rightarrow \mathbb{R}$ est la fonction définie par la coordonnée normale de ∂_t^D , dérivée de D le long du champ de vecteurs ∂_t de Γ , on a $C_D = \nu^{-1}(0)$ avec $d\nu_s \neq 0$ pour tout $s \in C_D$).

3 Généricité des hypothèses du paragraphe 2

Le problème que nous nous posons maintenant est de démontrer le caractère « générique », pour les familles à un paramètre de courbes, des deux propriétés prises comme hypothèses au paragraphe 2 :

« CS » : propriété de *caractéristique simple* ;

« EG » : propriété garantissant l'existence de l'*enveloppe géométrique*.

Que ces propriétés soient « génériques » signifie deux choses :

- (I) elles sont « robustes », c'est-à-dire que si une famille de courbes les vérifie, toute famille suffisamment proche les vérifie aussi ;
- (II) elles sont vérifiées par « presque toute famille », en ce sens que pour toute famille on peut trouver une famille qui en est aussi proche que l'on veut et qui vérifie ces propriétés.

La « robustesse » de (CS), (EG) est à peu près évidente, car la propriété de transversalité de deux sous-variétés (CS), (EG) se conserve par petites déformations de ces sous-variétés.

L'affirmation (II) est plus problématique, ne serait-ce que dans sa signification : comme il y est question de « trouver une famille qui... », le sens (et la véracité !) de l'affirmation dépend de façon cruciale des restrictions éventuelles qu'on a décidé d'imposer à la famille cherchée (cf. introduction du § 0).

Ayant fixé une fois pour toutes une famille $S = (S(u))_{u \in \Omega}$ appelée *famille universelle*, nous conviendrons désormais de ne nous intéresser qu'aux familles à un paramètre γ^*S induites de S par des changements de base $\gamma : \Gamma \rightarrow \Omega$ (où Γ est une variété à une dimension). Notre problème est de choisir la famille universelle S « assez grande » pour que les propriétés (CS) (EG) soient effectivement génériques dans l'ensemble des familles ainsi induites.

En fait nous nous contenterons de résoudre ce problème *localement dans S* , c'est-à-dire qu'au lieu de démontrer l'existence, arbitrairement près de toute famille induite γ_0^*S , d'une famille γ^*S vérifiant *partout* les propriétés (CS) (EG) (comme dans l'exercice recommandé en introduction), nous montrerons seulement comment déformer un *germe de changement de base* $\gamma_0 : \Gamma, t_0 \cdots \rightarrow \Omega, u_0$ de façon que la famille induite γ^*S vérifie les propriétés (CS) (EG) au voisinage d'un point donné de $S(u_0)$.

3.1 L'hypothèse locale d'universalité

Imaginons que nous observions à la loupe un voisinage d'un point $a \in M$. La courbe $S(u)$, dont nous supposons qu'elle passe par a , nous apparaîtra pratiquement rectiligne (Fig. 12). Si le point u se met en mouvement dans Ω , nous verrons la « droite » $S(u)$ se mettre en mouvement dans le champ de notre loupe (supposée immobile), et la question est de décrire ce mouvement.

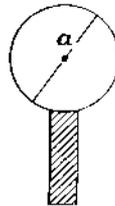


Figure 12. La courbe $S(u)$ « vue à la loupe ».

Il est clair que même d'un point de vue infinitésimal (c'est-à-dire pour un observateur idéal dont la loupe serait infiniment grossissante), la notion de dérive ne décrit qu'incomplètement le mouvement, car elle ne nous renseigne pas sur la façon dont la droite *tourne* dans le champ de la loupe : nous retrouvons ici le fait que les mouvements infinitésimaux d'une droite dans un plan sont décrits par deux paramètres, alors que la dérive de S dans M n'en fournit qu'un seul.

Pour obtenir une description plus complète, nous allons remplacer la surface M par la variété de ses éléments de contact : un **élément de contact** dans une surface M consiste en la donnée, en un point $a \in M$, d'une droite vectorielle $\tilde{a} \subset T_a M$ (on trouvera en appendice la définition générale d'un élément de contact dans une variété de dimension quelconque). L'ensemble \tilde{M} des éléments de contact de M est une variété

à trois dimensions, munie d'une projection $\tilde{\pi} : \tilde{M} \rightarrow M$ (définie par $\tilde{\pi}(\tilde{a}) = a$) qui en fait un fibré en droites projectives : la fibre $\tilde{\pi}^{-1}(a)$ n'est autre que $\mathbb{P}(T_aM)$, ensemble des droites vectorielles de T_aM . Pour tout système de coordonnées locales (x, y) sur M au voisinage de a , les coordonnées (dx, dy) de T_aM fournissent un système de coordonnées homogènes de $\mathbb{P}(T_aM)$; en posant $p = dy/dx$ on obtient un système de coordonnées locales (x, y, p) de \tilde{M} , défini sur l'ouvert des éléments de contact \tilde{a} pour lesquels $dx|_{\tilde{a}} \neq 0$.

Revenons à notre courbe lisse $S(u) \subset M$. En associant à chaque point $a \in S(u)$ la tangente $T_aS(u)$ en ce point (idéalisation mathématique de « ce que voit l'observateur dans sa loupe »), on obtient une section du fibré $\tilde{M}_{S(u)}$. L'image de cette section est une courbe lisse de \tilde{M} que nous noterons $\Sigma(u)$: si $S(u)$ est donnée dans les coordonnées x, y de \tilde{M} par $y = f(x, u)$, la courbe $\Sigma(u)$ est donnée dans les coordonnées x, y, p de \tilde{M} par

$$\Sigma(u) : \begin{cases} y = f(x, u) \\ p = \partial_x f(x, u) \end{cases}$$

En faisant varier u dans Ω on obtient ainsi une famille $(\Sigma(u))_{u \in \Omega}$ de courbes lisses de \tilde{M} dépendant différemment de u , donc un « espace au-dessus de Ω » dont nous noterons Σ l'espace total et Σ_u les fibres (difféomorphes à $\Sigma(u)$) ; le schéma suivant résume nos notations :

$$\begin{array}{ccccc} \tilde{M} & \xleftarrow{\tilde{g}} & \Sigma & \xrightarrow{\pi_\Sigma} & \Omega \\ \cup & & \cup & & \cup \\ \Sigma(u) & \xleftarrow{\cong} & \Sigma_u & \longrightarrow & \{u\} . \end{array}$$

En appliquant à cette famille de courbes la construction du paragraphe 1.3, on obtient une « dérive » :

$$\nabla_{a,u}^\Sigma : T_u\Omega \rightarrow T_a\tilde{M}/T_a\Sigma(u)$$

(notez que cette fois la dérive est à valeurs dans un espace vectoriel à deux dimensions).

Nous pouvons maintenant formuler l'hypothèse fondamentale que notre famille (S, π_S) devra vérifier pour mériter son qualificatif d'« universelle » .

Hypothèse d'universalité Pour tout $(\tilde{a}, u) \in \Sigma$, l'application linéaire $\nabla_{a,u}^\Sigma$ est surjective.

D'après l'exercice du paragraphe 1.3, on peut donner de cette hypothèse la formulation équivalente suivante :

l'application \tilde{g} est une submersion.

Traduction en coordonnées locales : si les coordonnées locales (x, y) ont été choisies de façon que $x(a) = y(a) = 0$ et $dy|_{\tilde{a}} = 0$, on vérifiera que l'hypothèse d'universalité au point (\tilde{a}, u) équivaut à demander à la matrice :

$$\underline{\ell} = \begin{pmatrix} \partial_{u_1} f(0, u) \cdots \partial_{u_n} f(0, u) \\ \partial_x \partial_{u_1} f(0, u) \cdots \partial_x \partial_{u_n} f(0, u) \end{pmatrix}$$

d'être de rang 2 (on a noté $y = f(x, u)$ l'équation de S dans $M \times \Omega$, et $\partial_{u_1}, \dots, \partial_{u_n}$ désignent les dérivées partielles par rapport à un système de coordonnées locales de Ω).

EXERCICE Tester l'hypothèse d'universalité sur quelques familles explicites, comme :

- la famille de toutes les droites du plan ;
- celle de tous les cercles du plan ;
- celle de tous les cercles du plan de rayon donné ; etc.

La fin du paragraphe va être consacré à montrer que, sous l'hypothèse d'universalité, les propriétés (CS), (EG) sont génériques localement, pour les familles induites par les changements de base $\gamma : \mathbb{R} \rightarrow \Omega$ (comme il ne s'agira que d'une étude locale, paramétrer les familles induites par \mathbb{R} ne restreint pas la généralité).

3.2 Généricité locale de la propriété (CS) (« caractéristique simple »)

Il résulte de sa définition même que cette propriété peut s'exprimer en termes de vecteurs tangents $\delta = \gamma'(t)$.

i) Dire que $a \in S(\gamma(t))$ est *caractéristique* pour S^γ , c'est dire que $\nabla_{a,u}^S(\delta) = 0$, où $u = \gamma(t)$.

Dans les coordonnées locales utilisées à la fin du paragraphe 3.2, où $S(u)$ est donnée par l'équation $y = f(x, u)$, cette condition peut s'écrire :

$$\ell_x(\delta) = 0$$

où $x = x(a)$, et où $\ell_x = d_u f(x, u)$ désigne la forme linéaire sur $T_u \Omega$ définie par la différentielle de f considérée comme fonction de u , à x fixé.

ii) Dire que a est *caractéristique simple*, c'est adjoindre à la condition précédente la condition :

$$\ell'_x(\delta) \neq 0.$$

où ℓ'_x désigne la dérivée de ℓ_x par rapport à x .

Soit donc $\gamma_0 : \mathbb{R}, 0 \cdots \rightarrow \Omega$, u_0 un changement de base tel que le point a_0 de coordonnée $x = 0$ dans $S(u_0)$ soit *caractéristique non simple* pour S^{γ_0} , ce qui se traduit par la condition suivante sur $\delta_0 = \gamma'_0(0)$:

$$\begin{cases} \ell_0(\delta_0) = 0 \\ \ell'_0(\delta_0) = 0. \end{cases}$$

Notre problème de généricité locale se ramène à démontrer l'affirmation suivante.

Affirmation Dans tout voisinage de δ_0 on peut trouver des vecteurs δ tels que le système

$$\begin{cases} \ell_x(\delta) = 0 \\ \ell'_x(\delta) = 0 \end{cases}$$

ne soit vérifié en aucun x dans un voisinage de 0.

En effet, en remplaçant alors l'application γ_0 par une application « voisine » $\gamma : \mathbb{R}, 0 \cdots \rightarrow \Omega$, u_0 dont le vecteur vitesse $\gamma'(0) = \delta$ soit choisi comme dans l'affirmation précédente, on obtient une famille induite S^γ aussi voisine qu'on veut de S^{γ_0} (au voisinage de $S(u_0)$), et qui n'a dans $S(u_0)$, au voisinage de a_0 , que des points *caractéristiques simples* (zéros simples de $x \mapsto \ell_x(\delta)$). Comme le fait pour une fonction

d'avoir des zéros simples est préservé par petites déformations, cela restera vrai dans $S(\gamma(t))$ pour tout t assez proche de 0.

Preuve de l'affirmation

Pour tout x voisin de 0, considérons l'application linéaire :

$$\underline{\ell}_x = (t_x, t'_x) : T_{u_0}\Omega \rightarrow \mathbb{R}^2.$$

D'après l'hypothèse d'universalité, cette application est de rang 2. Son noyau $\text{Ker } \underline{\ell}_x$ définit donc un sous-espace vectoriel de codimension 2 de $T_{u_0}\Omega$, dépendant différenciablement de x et contenant le vecteur δ_0 pour $x = 0$. Choisissons dans $T_{u_0}\Omega$ un 2-plan affine A passant par δ_0 et transverse à $\text{Ker } \underline{\ell}_0$. Pour tout x proche de 0, $\text{Ker } \underline{\ell}_x$ rencontre A en un seul point $\delta(x)$, qui dépend différenciablement de x (avec $\delta(0) = \delta_0$). Autrement dit, la correspondance $x \mapsto \delta(x)$ est une courbe paramétrée de A passant par δ_0 au temps $x = 0$. Tout se ramène donc à l'affirmation intuitivement évidente que l'image locale d'une courbe paramétrée dans un plan ne peut pas remplir tout un voisinage d'un point de ce plan (exercice : démontrer cette affirmation). \square

3.3 Généricité locale de la propriété (EG) (« enveloppe géométrique »)

Supposons que la famille induite par le changement de base $\gamma_0 : \mathbb{R} \cdot 0 \cdots \rightarrow \Omega$, u_0 vérifie (CS) mais pas (EG). Plus précisément, supposons que le point $a_0 \in S(u_0)$ de coordonnée $x(a_0) = 0$ soit caractéristique simple pour la famille S^{γ_0} , mais que la solution $x(t)$ de l'équation implicite en x :

$$\partial_t f^\gamma(x, t) = 0 \quad (\text{où } f^\gamma(x, t) = f(x, \gamma(t)))$$

admette en $t = 0$ un point critique non quadratique.

En perturbant aussi peu qu'on veut la fonction $t \mapsto x(t)$, on peut construire une fonction n'ayant au voisinage de $t = 0$ que des points critiques quadratiques (exercice : le démontrer, en raisonnant sur la fonction x' ; dans la mesure où l'on ne veut la construire que localement, il suffira par exemple de rajouter à $x(t)$ un terme linéaire εt , avec ε aussi petit qu'on veut).

Soit $x_\varepsilon(t)$ la fonction ainsi obtenue. Notre problème de généricité locale sera résolu si nous savons résoudre le problème suivant.

Problème Construire une application γ_ε , proche de γ_0 , telle que $x_\varepsilon(t)$ soit la solution de l'équation implicite en x :

$$\partial_t f^{\gamma_\varepsilon}(x, t) = 0.$$

Remarques préparatoires Remarquons que tout changement de base γ assez proche de γ_0 vérifie, comme γ_0 , l'hypothèse de caractéristique simple, on notera $x^\gamma(t)$ la solution de l'équation implicite en x :

$$\partial_t f^\gamma(x, t) = 0.$$

Notant $p(x, u)$ la fonction $p(x, u) = \partial_x f(x, u)$, on posera :

$$\begin{aligned} y^\gamma(t) &= f^\gamma(x^\gamma(t), t) \\ \text{et } p^\gamma(t) &= p(x^\gamma(t), \gamma(t)). \end{aligned}$$

On notera

$$c^\gamma : \mathbb{R} \rightarrow M$$

l'application définie dans les coordonnées (x, y) de M par

$$x = x^\gamma(t) \quad . \quad y = y^\gamma(t)$$

(c'est un paramétrage de l'enveloppe à la source C_{S^γ}) ; et

$$\tilde{c}^\gamma : \mathbb{R} \rightarrow \tilde{M}$$

l'application définie dans les coordonnées (x, y, p) de \tilde{M} par

$$x = x^\gamma(t) \quad . \quad y = y^\gamma(t) \quad . \quad p = p^\gamma(t)$$

(c'est un paramétrage de la « courbe legendrienne » $C_{S^\gamma}^*$, ensemble des « éléments de contact attachés à C_{S^γ} » : cf. Appendice).

Enfin $k^\gamma : \mathbb{R} \rightarrow S$ [resp. $\tilde{k}^\gamma : \mathbb{R} \rightarrow \Sigma$] désignera l'application qui à t associe :

$$\begin{aligned} k^\gamma(t) &= (c^\gamma(t), \gamma(t)) \in S \subset M \times \Omega \\ \text{[resp. } \tilde{k}^\gamma(t) &= (\tilde{c}^\gamma(t), \gamma(t)) \in \Sigma \subset \tilde{M} \times \Omega. \end{aligned}$$

Remarquons que d'après l'hypothèse de caractéristique simple, dp^γ ne s'annule pas au voisinage du point étudié. Rien n'empêche donc de reparamétriser γ de façon que $p^\gamma(t) = t$.

Solution du problème : la donnée de notre problème est une fonction $x_e(t)$, proche de $x^{\gamma_0}(t)$. Notant $y_e(t)$ une primitive de $t x_e'(t)$, proche de $y^{\gamma_0}(t)$, on obtient un paramétrage de courbe legendrienne de \tilde{M} , proche de \tilde{c}^{γ_0} :

$$\begin{aligned} \tilde{c}_e &: \mathbb{R} \rightarrow \tilde{M} \\ t &\mapsto (x = x_e(t) \quad , \quad y = y_e(t) \quad , \quad p = t) \end{aligned}$$

(cette courbe est legendrienne par construction, parce que $dy_e = p(t) dx_e$).

Comme $\tilde{g} : \Sigma \rightarrow \tilde{M}$ est une submersion (hypothèse d'universalité 3.2), cette courbe de \tilde{M} est localement l'image d'une courbe de Σ : il existe une application $\tilde{k}_e : \mathbb{R} \rightarrow \Sigma$, proche de \tilde{k}^{γ_0} , telle que $\tilde{c}_e = \tilde{g} \circ \tilde{k}_e$.

Posant $\gamma_e = \pi_\Sigma \circ \tilde{k}_e$, nous obtenons une solution de notre problème (la vérification est laissée au lecteur).

APPENDICE

Éléments de contact sur une variété et condition d'intégrabilité de Legendre

Bien que motivé par le thème d'étude précédent (fin du § 3), cet appendice peut aussi être étudié indépendamment. Pour des raisons de clarté conceptuelle nous ne nous sommes pas limités au cas de la dimension 2 : M désignera ici une variété de dimension n quelconque.

0 Définition

Un **élément de contact** sur une variété M à n dimensions consiste en la donnée d'un hyperplan vectoriel $\tilde{\alpha}$ de l'espace tangent à M en un point a :

$$\tilde{\alpha} \subset T_a M \quad (\text{codim} \tilde{\alpha} = 1).$$

Dans le cas $n = 2$, un élément de contact est donc une droite vectorielle d'un plan tangent à M (Fig. 13).

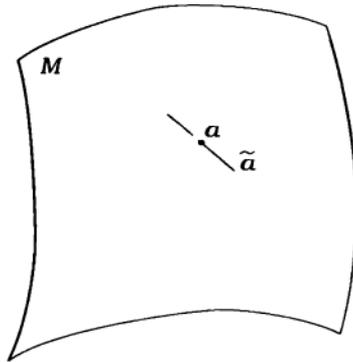


Figure 13. Éléments de contact sur une surface.

Proposition L'ensemble \tilde{M} des éléments de contacts sur M est muni naturellement d'une structure de variété différentielle de dimension $2n - 1$. Cette variété, munie de la projection évidente $\tilde{\pi} : \tilde{M} \rightarrow M$ (qui à un hyperplan vectoriel $\tilde{\alpha} \subset T_a M$ associe a) est un fibré de base M .

Preuve

L'équation d'un hyperplan vectoriel $\tilde{\alpha} \subset T_a M$ est un covecteur non nul $\alpha \in T_a^* M$, unique à multiplication par un scalaire près. L'espace $\tilde{M}_a := \tilde{\pi}^{-1}(a)$ s'identifie donc à l'espace projectif $\mathbb{P}(T_a^* M)$, et le couple $(\tilde{M}, \tilde{\pi})$ au fibré projectif $\mathbb{P}^* M := \mathbb{P}(T^* M)$. \square

1 Éléments de contact attachés à une sous-variété de M

Soit $V \subset M$ une sous-variété de codimension ≥ 1 . Un élément de contact $\tilde{\alpha} \in \tilde{M}_a$ est dit **attaché** à V si $a \in V$ et $\tilde{\alpha} \supset T_a V$. La figure 14 représente, dans le cas $n = 2$, l'ensemble des éléments de contact attachés à une sous-variété V de codimension 1 ou 2.



Figure 14. Ensemble des éléments de contact attachés à une sous-variété V , dans le cas $\dim M = 2$.

PROPOSITION L'ensemble V^* des éléments de contact attachés à V est une sous-variété de \tilde{M} , de dimension $n - 1$.

EXERCICE Le démontrer ; on vérifiera que $V^* = \mathbb{P}(T_V^*M) = \mathbb{P}_V^*M$, fibré projectif associé au fibré conormal à V .

2 Condition d'intégrabilité et champ de contact sur \tilde{M}

Étant donné dans \tilde{M} une sous-variété Λ de dimension $n - 1$, il est naturel de se demander à quelle condition il existe une sous-variété V de M telle que $\Lambda = V^*$. Nous allons donner une condition nécessaire portant sur les espaces tangents à Λ , la « condition d'intégrabilité de Legendre ».

Supposons donc que $\Lambda = V^*$, et soit $\tilde{\alpha}(t)$ un point de Λ mobile avec le temps. Son image $\alpha(t)$ par $\tilde{\pi}$ est un point mobile de V , dont le vecteur vitesse $\nu(t)$ à chaque instant est tangent à V , et appartient donc à l'hyperplan $\tilde{\alpha}(t)$:

$$\nu \in \tilde{\alpha}.$$

Pour tout $\tilde{\alpha} \in \tilde{M}$, notons $\kappa(\tilde{\alpha})$ l'ensemble des vecteurs $\tilde{\nu}$ de $T_{\tilde{\alpha}}\tilde{M}$ dont l'image ν par $T_{\tilde{\alpha}}\tilde{\pi}$ vérifie la condition $\nu \in \tilde{\alpha}$ ci-dessus.

Étant l'image réciproque d'un hyperplan $\tilde{\alpha}$ par une application linéaire surjective, $\kappa(\tilde{\alpha})$ est un hyperplan de $T_{\tilde{\alpha}}\tilde{M}$, autrement dit un élément de contact sur \tilde{M} ! La variété \tilde{M} est ainsi munie canoniquement d'un champ d'éléments de contact κ (section κ du fibré $\mathbb{P}^*\tilde{M}$ des éléments de contact sur \tilde{M}). Ce champ κ est appelé *champ de contact* sur \tilde{M} , et les $\kappa(\tilde{\alpha})$ sont les *hyperplans de contact* (plans de contact dans le cas où M est de dimension 2, de sorte que \tilde{M} est de dimension 3).

La façon dont nous avons introduit ce champ motive la définition suivante.

DÉFINITION Une sous-variété $\Lambda \subset \tilde{M}$ de dimension $n - 1$ est dite vérifier la **condition d'intégrabilité de Legendre** si elle est partout tangente au champ de contact κ (en ce sens que pour tout $\tilde{\alpha} \in \Lambda$, $T_{\tilde{\alpha}}\Lambda \subset \kappa(\tilde{\alpha})$). On appelle **variété legendrienne** une sous-variété de dimension $n - 1$ vérifiant la condition d'intégrabilité de Legendre.

Exemple $\Lambda = V^*$ (V sous-variété de M).

Remarque Soit Λ une variété legendrienne telle que $\tilde{\pi}|_{\Lambda}$ soit une immersion. Pour étudier Λ au voisinage d'un point donné, on peut toujours se restreindre à un voisinage \tilde{U} assez petit pour que $\tilde{\pi}|_{\Lambda \cap \tilde{U}}$ soit un plongement. Alors, en désignant par $V_{\tilde{U}}$ l'hypersurface image de ce plongement, il est facile de voir que $\Lambda \cap U = V_{\tilde{U}}^*$. Cette situation est illustrée par la figure 15.

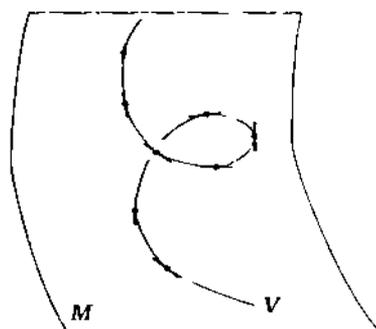


Figure 15. Exemple de courbe legendrienne Λ ($\dim M = 2$) telle que $\tilde{\pi}|_{\Lambda}$ soit une immersion : on a noté V la courbe image de cette immersion.

EXERCICE TEST Donner, dans le cas $n = 2$, un exemple de courbe legendrienne qui n'est pas du type de la figure 14, même localement (l'exemple de la figure 15 est localement dans \tilde{M} du type de la figure 14).

Pour en savoir plus

V. Arnold, A. Varchenko, S. Goussein-Zadé, *Singularités des Applications différentiables*, première partie (trad. française éd. Mir, 1986).

V. Arnold, *Singularities of Caustics and Wave Fronts* (Mathematics and its applications, vol. 62, Kluwer, 1990).

Homotopie et revêtements

Nous abordons ici les premiers rudiments de la *topologie algébrique*, branche des mathématiques qui a pris son plein essor à l'orée du siècle (avec les travaux de H. Poincaré, notamment). En abordant ce chapitre on fera bien de garder présentes à l'esprit les considérations sur les « indices » présentées à la fin de la première partie (Intermède) : réussir à rendre rigoureuses ces considérations est un excellent objectif à se fixer en lisant ce qui suit.

Sauf quelques rares références à la structure différentielle des variétés (§ 2 et 3), toutes les notions de ce chapitre sont purement topologiques.

0. La catégorie d'homotopie

0.0. Familles continues d'applications

D'un point de vue purement ensembliste (en oubliant la topologie), se donner une famille d'applications $(F_t : X \rightarrow Y)_{t \in T}$ équivaut évidemment à se donner une application $F : X \times T \rightarrow Y$ (écrire $F(x, t) = F_t(x)$). Si X, Y, T sont des espaces topologiques et si F est une application continue, il en est évidemment de même des F_t , et l'on dira que $F = (F_t)_{t \in T}$ est une **famille continue d'applications de X dans Y , paramétrée sur T** .

Il est souvent utile d'imposer aux F_t d'envoyer un sous-espace donné de X dans un sous-espace donné de Y . Au lieu de travailler dans la catégorie des espaces topologiques, on travaille alors dans la catégorie des **paires d'espaces topologiques** : un objet de cette catégorie (« paire d'espaces topologiques ») est un couple constitué d'un espace topologique X et d'un sous-espace $A \subset X$; un morphisme $f : (X, A) \rightarrow (Y, B)$ (« application de paires ») consiste en la donnée des paires $(X, A), (Y, B)$ et d'une application continue $f : X \rightarrow Y$ telle que $f(A) \subset B$. Une **famille continue d'applications de (X, A) dans (Y, B) , paramétrée sur T** , est une application de paires :

$$F : (X \times T, A \times T) \rightarrow (Y, B) ;$$

on peut y penser comme à une famille d'applications continues $(F_t : (X, A) \rightarrow (Y, B))_{t \in T}$ « dépendant continûment de $t \in T$ ».

Le cas particulier $A = B = \emptyset$ redonne la notion précédente de famille continue d'applications de X dans Y .

0.1. La relation d'homotopie

Soient f_0 et $f_1 : (X, A) \rightarrow (Y, B)$ deux applications de paires.

Une **homotopie** entre f_0 et f_1 consiste en la donnée d'une famille continue F d'applications de (X, A) dans (Y, B) , paramétrée sur $[0, 1]$, telles que $F_0 = f_0$, $F_1 = f_1$.

S'il existe une telle homotopie, on dit que les applications f_0 et f_1 sont **homotopes**, ce qu'on note ainsi :

$$f_0 \sim f_1.$$

PROPOSITION La relation d'homotopie (entre applications d'une paire donnée (X, A) dans une paire donnée (Y, B)) est une relation d'équivalence.

Preuve

Elle est facile ; par exemple pour démontrer la transitivité : $f \sim g, g \sim h \Rightarrow f \sim h$ il suffit de « mettre bout à bout » les deux homotopies $F : f \sim g, G : g \sim h$ en « reparamétrant » la première sur $[0, 1/2]$ et la seconde sur $[1/2, 1]$. \square

Les classes d'équivalence pour cette relation sont appelées **classes d'homotopie de (X, A) dans (Y, B)** . L'ensemble de ces classes est noté :

$$[(X, A), (Y, B)].$$

On note $[f] \in [(X, A), (Y, B)]$ la classe d'homotopie de l'application $f : (X, A) \rightarrow (Y, B)$.

Exemple : classe d'homotopie de lacets Un **lacet** d'un espace topologique X est une application continue $\lambda : [0, 1] \rightarrow X$ telle que $\lambda(0) = \lambda(1)$. On peut donc considérer qu'il s'agit d'une application de paires :

$$\lambda : (I, \partial I) \rightarrow (X, \{a\})$$

où $I = [0, 1]$, $\partial I = \{0\} \cup \{1\}$, le point a étant le **point de base** du lacet ($a = \lambda(0) = \lambda(1)$). Le point de base a étant fixé, une **homotopie entre lacets de (X, a)** est une famille continue, paramétrée par $t \in [0, 1]$, de lacets λ_t ayant tous le même point de base a .

L'ensemble $[(I, \partial I), (X, \{a\})]$ des classes d'homotopie de lacets de (X, a) est le « groupe fondamental » $\pi_1(X, a)$, dont l'étude fera l'objet du paragraphe 1.

0.2. Composition des classes d'homotopie

PROPOSITION La relation d'homotopie est préservée par composition des applications :

$$\begin{array}{lll} \text{si} & f_0 \sim f_1 & : (X, A) \rightarrow (Y, B) \\ \text{et} & g_0 \sim g_1 & : (Y, B) \rightarrow (Z, C) \\ \text{alors} & g_0 \circ f_0 \sim g_1 \circ f_1 & : (X, A) \rightarrow (Z, C) \end{array}$$

Preuve

Soit $(F_t)_{t \in [0,1]}$ [resp. $(G_t)_{t \in [0,1]}$] une homotopie entre f_0 et f_1 [resp. entre g_0 et g_1].
Alors $(G_t \circ F_t)_{t \in [0,1]}$ est évidemment une homotopie entre $g_0 \circ f_0$ et $g_1 \circ f_1$. \square

La catégorie d'homotopie

La proposition précédente permet de définir la *composée de deux classes d'homotopie* : à toute $[f] \in [(X, A), (Y, B)]$ et $[g] \in [(Y, B), (Z, C)]$ est associé :

$$[g] \circ [f] := [g \circ f] \in [(X, A), (Z, C)].$$

On appelle **catégorie d'homotopie** la catégorie dont les objets sont les paires d'espaces topologiques et dont les morphismes sont les classes d'homotopie d'applications de paires, munies de la loi de composition ci-dessus.

0.3. Équivalence d'homotopie

Les isomorphismes de la catégorie d'homotopie sont appelées **équivalences d'homotopie** : explicitement, se donner une équivalence d'homotopie entre (X, A) et (Y, B) signifie se donner deux applications de paires $f : (X, A) \rightarrow (Y, B)$ et $g : (Y, B) \rightarrow (X, A)$ qui soient « inverses d'homotopie l'une de l'autre », en ce sens que :

$$g \circ f \sim \mathbb{I}_{X,A} \text{ , } f \circ g \sim \mathbb{I}_{Y,B}$$

où $\mathbb{I}_{X,A}$ désigne l'application identique de la paire (X, A) dans elle-même (un exemple important d'équivalence d'homotopie est la « rétraction par déformation » d'un espace X sur un sous-espace Y : cf. le paragraphe 0.4).

Deux objets isomorphes dans la catégorie d'homotopie sont dits avoir **même type d'homotopie**. Il est clair que pour calculer l'ensemble $[(X, A), (Y, B)]$ des classes d'homotopie de morphismes d'un objet dans un autre on peut toujours remplacer l'objet source (X, A) , ou l'objet but (Y, B) , par un objet ayant même type d'homotopie.

0.4. Rétractions – déformations

Soit Y un sous-espace d'un espace topologique X , et soit $i : Y \rightarrow X$ l'inclusion canonique. On appelle **rétraction** de X sur Y une application $r : X \rightarrow Y$ telle que $r \circ i = \mathbb{I}_Y$. On appelle **rétraction par déformation** une rétraction r telle que $i \circ r \sim \mathbb{I}_X$. S'il existe une telle rétraction [resp. rétraction par déformation] on dit que l'espace X est **rétractile** sur Y [resp. **rétractile par déformation** sur Y], ou encore que Y est un **rétracté** [resp. un **rétracté par déformation**] de X .

On définirait de même la notion de rétraction [resp. rétraction par déformation] d'une paire (X, A) sur une « sous-paire » (Y, B) (paire telle que $Y \subset X, B \subset A$).

Espaces contractiles Tout espace est évidemment rétractile sur n'importe lequel de ses points $\{a\}$ (l'application constante $X \rightarrow \{a\}$ est une rétraction !). On dit que X est **contractile** s'il est rétractile *par déformation* sur l'un de ses points.

Exemple i) Toute partie « étoilée » de \mathbb{R}^n est contractile. Un sous-espace $X \subset \mathbb{R}^n$ est dit **étoilé** de centre 0 si $x \in X \Rightarrow tx \in X$ pour tout $t \in [0, 1]$ (Fig. 1).

Notons $r : X \rightarrow \{0\}$ la rétraction évidente, de sorte que $i \circ r$ est l'application nulle de X dans lui-même. On obtient une homotopie F entre $i \circ r$ et \mathbb{I}_X en posant $F_t(x) = tx$.

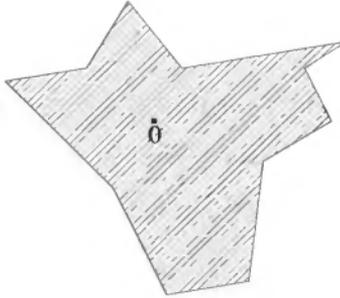


Figure 1. Partie étoilée de \mathbb{R}^2 .

Bien entendu, ce qui précède s'applique aussi aux parties étoilées d'un espace affine, le centre de l'étoile étant un point quelconque de cet espace.

Exemple ii) $X = \mathbb{R}^{n+1} \setminus \{0\}$ est rétractile par déformation sur la sphère S^n .

EXERCICE r désignant la rétraction évidente définie par $r(x) = x/\|x\|$, on construira une homotopie entre $i \circ r$ et \mathbb{I}_X .

EXERCICE TEST 1 Soit X un espace contractile. Montrer à l'aide de 0.3 que toute application $f : X \rightarrow Y$ est homotope à une application constante (qui envoie tout X sur un même point de Y). Si de plus Y est connexe par arcs, montrer que toutes les applications de X dans Y sont homotopes entre elles.

EXERCICE TEST 2 Montrer les égalités suivantes :

$$\begin{aligned} [\mathbb{R}^2 \setminus \{0\}, \mathbb{R}^2 \setminus \{0\}] &= [S^1, \mathbb{R}^2 \setminus \{0\}] \\ &= [\mathbb{R}^2 \setminus \{0\}, S^1] = [S^1, S^1] \end{aligned}$$

Remarque On verra au paragraphe 2.2 que

$$[S^1, S^1] = \mathbb{Z} :$$

les classes d'homotopie d'applications du cercle dans lui-même sont caractérisées par un entier relatif, leur « indice ».

1. Groupe fondamental

Soit X un espace topologique muni d'un point marqué a . Le but de ce paragraphe est de définir une loi de groupe sur l'ensemble des classes d'homotopie de lacets de (X, a) . Ce groupe (définition 1.1) est appelé *groupe fondamental de X* (de point-base a), et noté $\pi_1(X, a)$.

Pour comprendre dans quelle mesure ce groupe dépend du choix de a , il sera commode de faire un travail préparatoire sur les classes d'homotopie de chemins d'origine et d'extrémité quelconques.

1.0. Composition de chemins sur un espace topologique

On appelle **chemin** sur un espace topologique X une application continue :

$$\lambda : I = [0, 1] \rightarrow X.$$

Le point $\lambda(0)$ [resp. $\lambda(1)$] est appelé *origine* [resp. *extrémité*] de λ .

En se fixant l'origine a et l'extrémité b des chemins on obtient la notion d'**homotopie de chemins** d'origine a et d'extrémité b (famille continue $(\lambda_s : I \rightarrow X)_{s \in [0,1]}$ telle que $\lambda_s(0) = a, \lambda_s(1) = b$ pour tout $s \in [0, 1]$).

C'est évidemment une relation d'équivalence, et nous noterons $[I, X]_{a,b}$ l'ensemble des classes d'homotopie de chemins d'origine a et d'extrémité b .

Considérons maintenant deux chemins λ et μ tels que l'extrémité b de λ coïncide avec l'origine de μ (Fig. 2).

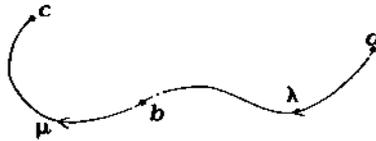


Figure 2.

On définit alors le chemin $\mu\lambda$ par :

$$(\mu\lambda)(t) = \begin{cases} \lambda(2t) & \text{si } t \in [0, 1/2] \\ \mu(2t - 1) & \text{si } t \in [1/2, 1] \end{cases}$$

(penser à un point mobile effectuant d'abord le parcours λ puis de parcours μ , chacun à une vitesse double de façon que le temps mis à effectuer le parcours total soit égal à 1).

(N.B. : Beaucoup d'auteurs notent $\lambda\mu$ ce que nous notons $\mu\lambda$. Notre convention de notation est inspirée par l'analogie avec la composée de deux applications : de même qu'une application est notée comme une « flèche » allant de l'espace « source » à l'espace « but », un chemin peut être vu comme une flèche allant du point « origine » au point « extrémité »).

Avec les notations de la figure 2, où a désigne l'origine de λ et c l'extrémité de μ , on voit facilement que la classe d'homotopie de $\mu\lambda$ dans $[I, X]_{a,c}$ ne dépend que des classes d'homotopie respectives de λ et de μ dans $[I, X]_{a,b}$ et $[I, X]_{b,c}$. Nous la noterons :

$$[\mu][\lambda] := [\mu\lambda].$$

PROPOSITION La loi de composition ainsi définie sur la famille d'ensembles $[I, X]_{a,b}$ ($(a, b) \in X^2$) est une loi de groupoïde : autrement dit elle vérifie les trois propriétés d'associativité, existence d'éléments neutres, existence des inverses, énoncées (et démontrées) ci-après.

Associativité : $[\nu][\mu\lambda] = [\nu\mu][\lambda]$ (si $\lambda(1) = \mu(0)$, $\mu(1) = \nu(0)$).

Preuve

Définissons le chemin $\nu \mu \lambda$ par :

$$(\nu \mu \lambda)(t) = \begin{cases} \lambda(3t) & \text{si } t \in [0, 1/3] \\ \mu(3t - 1) & \text{si } t \in [1/3, 2/3] \\ \nu(3t - 2) & \text{si } t \in [2/3, 1] \end{cases}$$

Alors $\nu(\mu\lambda)$ [resp. $(\nu\mu)\lambda$] est l'application de $[0, 1]$ dans X obtenue en composant $\nu \mu \lambda$ avec l'application de $[0, 1]$ dans lui-même dont le graphe est représenté en trait plein [resp. en tirets] sur la figure 3.

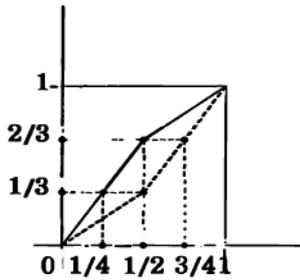


Figure 3.

Il est clair que ces deux applications de $(I, \partial I)$ dans lui-même sont homotopes, de sorte que :

$$\nu(\mu\lambda) \sim (\nu\mu)\lambda .$$

□

Éléments neutres : Soit $[e_a] \in [I, X]_{a,a}$ la classe du chemin constant $e_a(t) = a$; pour tout chemin λ d'origine a [resp. d'extrémité a] on a :

$$\begin{aligned} [\lambda][e_a] &= [\lambda] \\ \text{[resp. } [e_a][\lambda] &= [\lambda] \text{].} \end{aligned}$$

Preuve

λe_a [resp. $e_a \lambda$] est l'application de $[0, 1]$ dans X obtenue en composant λ avec l'application de $[0, 1]$ dans lui-même dont le graphe est représenté en trait plein

[resp. en tirets] sur la figure 4.

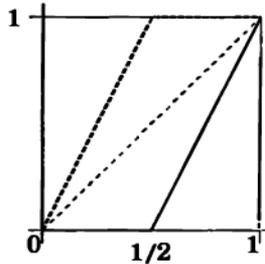


Figure 4.

Il est clair que cette application de $(I, \partial I)$ dans lui-même est homotope à celle représentée en pointillés, de sorte que $\lambda e_a \sim \lambda$ [resp. $e_a \lambda \sim \lambda$]. □

Inverse d'une classe $[\lambda]$ d'origine a et d'extrémité b : c'est la classe du chemin λ^{-1} d'origine b et d'extrémité a défini par $\lambda^{-1}(t) = \lambda(1-t)$ (c'est le chemin λ « parcouru en sens inverse »).

Le chemin $\lambda^{-1}\lambda$ [resp. $\lambda\lambda^{-1}$] est donc l'application de $[0, 1]$ dans X obtenue en composant λ avec l'application de $[0, 1]$ dans lui-même dont le graphe est représenté en trait plein [resp. en tirets] sur la figure 5.

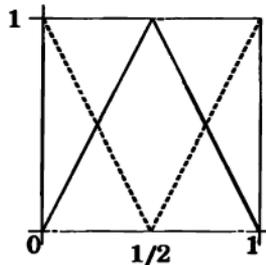


Figure 5.

Cette application de $(I, \partial I)$ dans lui-même est évidemment homotope à l'application constante qui prend la valeur 0 [resp. 1], de sorte que l'on a bien :

$$\begin{aligned} &\lambda^{-1}\lambda \sim e_a \quad \text{dans } [I, X]_{a,a} \\ \text{et } &\lambda\lambda^{-1} \sim e_b \quad \text{dans } [I, X]_{b,b}. \end{aligned}$$

Parenthèse terminologique : « groupoïdes », « monoïdes » et « catégories »

Si, dans la définition de la notion de groupoïde, on oublie l'existence de l'inverse, on obtient la notion de *monoïde*. On remarquera l'étroite ressemblance de cette notion avec celle de *catégorie* (cf. chap. 2, § 3.2) : dans les deux cas on a des ensembles de « flèches » (les morphismes de la catégorie), chacun de ces ensembles étant doublement indexé (par l'« origine » et l'« extrémité » dans le cas qui nous occupe ici, par l'« objet-source » et l'« objet-but » dans le cas d'une catégorie), avec une loi permettant de composer deux flèches si l'extrémité de la première coïncide avec l'origine de la seconde.

La seule différence est de nature syntaxique : alors que les indices des ensembles de flèches d'un monoïde sont eux-mêmes éléments d'un ensemble (l'ensemble X dans le cas qui nous occupe ici), on ne peut en général pas parler de « l'ensemble » des objets d'une catégorie (par exemple la « catégorie des ensembles » a pour objets « tous les ensembles », et des paradoxes célèbres ont conduit les fondateurs de la théorie des ensembles à s'interdire de parler de « l'ensemble de tous les ensembles »... !).

1.1. Définition du groupe fondamental

Choisissons dans l'espace topologique X un « point de base » a . En restreignant la construction du paragraphe 1.0 aux classes d'homotopie de lacets de base a , on obtient une loi de composition interne partout définie sur $[l, X]_{a,a}$ ($= [(l, \partial l), (X, a)]$ dans les notations du paragraphe 0). Il résulte de 1.0 qu'il s'agit d'une loi de groupe. Le groupe ainsi défini est noté $\pi_1(X, a)$ et appelé **groupe fondamental de X basé en a** .

Attention ! Ce groupe n'est généralement pas commutatif.

Par exemple, si X est le plan privé de deux points :

$$X = \mathbb{R}^2 \setminus \{a_1, a_2\} \quad (a_1 \neq a_2).$$

on démontre que le groupe $\pi_1(X, a)$ est engendré par les classes $[\lambda_1], [\lambda_2]$ de deux lacets comme ceux de la figure 6 i), ces classes n'étant sujettes à aucune relation (autres que celles qui résultent des axiomes d'un groupe) : on résume cette affirmation en disant que $\pi_1(X, a)$ est le *groupe libre* (non commutatif) engendré par $[\lambda_1]$ et $[\lambda_2]$.

Le lecteur qui soupçonnerait $\pi_1(X, a)$ d'être commutatif peut toujours essayer d'imaginer une homotopie entre le lacet de la figure 6 ii) (qui représente la classe $[\lambda_2][\lambda_1]$) et celui de la figure 6 iii) (qui représente la classe $[\lambda_1][\lambda_2]$) !

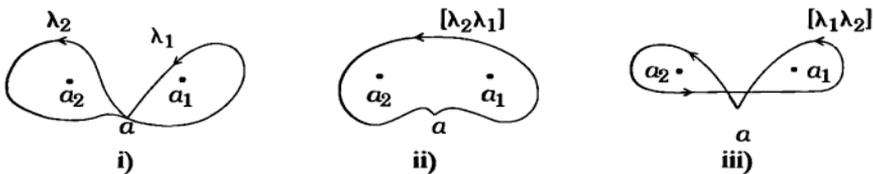


Figure 6.

Changement de point-base Pour comparer $\pi_1(X, a)$ et $\pi_1(X, b)$, supposons X connexe par arcs et donnons-nous un chemin μ d'origine a et d'extrémité b . Alors la « conjugaison par $[\mu]$ » définit évidemment un *isomorphisme de groupes* :

$$\begin{array}{ccc} \pi_1(X, a) & \xrightarrow{\cong} & \pi_1(X, b) \\ [\lambda] & \mapsto & [\mu \lambda \mu^{-1}] \end{array}$$

qui est illustré par la figure 7.



Figure 7. Changement de point-base.

Remarque $\pi_1(X, a) = [(S^1, u_0), (X, a)]$.

Envoyant 0 et 1 sur le même point $a \in X$, tout lacet λ de base a se factorise à travers l'espace topologique quotient de $[0, 1]$ par la relation consistant à identifier le point 0 au point 1. L'homéomorphisme évident entre cet espace quotient et le cercle S^1 établit donc une bijection entre l'ensemble des [classes d'homotopie de] lacets de X , de base a , et l'ensemble des [classes d'homotopie d']applications de (S^1, u_0) dans (X, a) , où $u_0 = (1, 0)$ est le « point-base » du cercle trigonométrique.

Plus généralement, l'ensemble des classes d'homotopie d'applications de (S^n, u_0) dans (X, a) (où u_0 est un point marqué de la sphère S^n) peut être muni d'une loi de composition interne qui en fait un groupe noté $\pi_n(X, a)$, appelé **n -ième groupe d'homotopie de X , de point-base a** . Pour $n > 1$, ce groupe est commutatif.

1.2. Propriétés fonctorielles de π_1

Soit $f : (X, a) \rightarrow (Y, b)$ une application continue d'espaces topologiques pointés. L'image par f d'un lacet λ de (X, a) est un lacet $f \circ \lambda$ de (Y, b) dont la classe d'homotopie dans $\pi_1(Y, b)$ ne dépend que de la classe λ dans $\pi_1(X, a)$ (cf. § 0.2). On obtient ainsi une application :

$$\begin{aligned} \pi_1(f) : \pi_1(X, a) &\rightarrow \pi_1(Y, b) \\ [\lambda] &\mapsto [f \circ \lambda] \end{aligned}$$

(fréquemment notée f_* pour alléger l'écriture).

PROPOSITION $\pi_1(f)$ est un homomorphisme de groupes.

Preuve

Ceci est évident. □

Il résulte encore de 0.2 que f_* ne dépend que de la classe d'homotopie de f dans $[(X, a), (Y, b)]$, et que la correspondance $f \mapsto f_* = \pi_1(f)$ vérifie les propriétés « fonctorielles » :

$$\begin{aligned} (g \circ f)_* &= g_* f_* \\ (\mathbb{I}_{X, a})_* &= \mathbb{I}_{\pi_1(X, a)}. \end{aligned}$$

Autrement dit, cette correspondance définit un foncteur covariant de la catégorie des espaces topologiques pointés (ou, si l'on préfère, de la catégorie d'homotopie correspondante) dans la catégorie des groupes.

Corollaire i).— Toute équivalence d'homotopie (entre deux espaces topologiques pointés $(X, a), (Y, b)$) induit un isomorphisme de leurs groupes fondamentaux :

$$\pi_1(X, a) \cong \pi_1(Y, b).$$

Exemple $\pi_1(\mathbb{R}^2 \setminus \{0\}, a) = \pi_1(S^1, a)$
 (= \mathbb{Z} si l'on en croit la remarque finale de 0.4).

Corollaire ii).— Toute rétraction $r : X \rightarrow Y$ induit un épimorphisme

$$r_* : \pi_1(X, a) \rightarrow \pi_1(Y, a) \text{ (où } a \in Y \subset X).$$

En effet, la définition d'une rétraction ($r \circ i = \mathbb{I}_Y$) implique par fonctorialité $r_* \circ i_* = \mathbb{I}_{\pi_1(Y, a)}$, de sorte que r_* est surjectif.

EXERCICE TEST Existe-t-il une rétraction du disque unité D sur son bord S^1 ?

1.3. Simple connexité

Un espace topologique X est dit **simplement connexe** si tout couple de points a_1, a_2 de X peut être « joint » par un chemin λ (d'origine a_1 et d'extrémité a_2) unique à homotopie près. Cela implique deux choses :

- i) X est connexe par arcs (existence de λ) ;
- ii) $\pi_1(X, a)$ est *trivial* (unicité, à homotopie près, du chemin joignant a à a !).

Réciproquement, on voit facilement que i) et ii) impliquent la simple connexité : s'il existait, dans X supposé connexe par arcs, deux chemins *non homotopes* λ, λ' d'origine a_1 et d'extrémité a_2 , le choix de chemins μ_1, μ_2 joignant un point a à a_1 et a_2 respectivement (Fig. 8) permettrait de construire deux lacets non homotopes de base a , à savoir $\mu_2^{-1} \lambda \mu_1$ et $\mu_2^{-1} \lambda' \mu_1$ (ici intervient à nouveau, comme pour le changement de point-base, le fait que les classes d'homotopie de chemins forment un « groupoïde »).

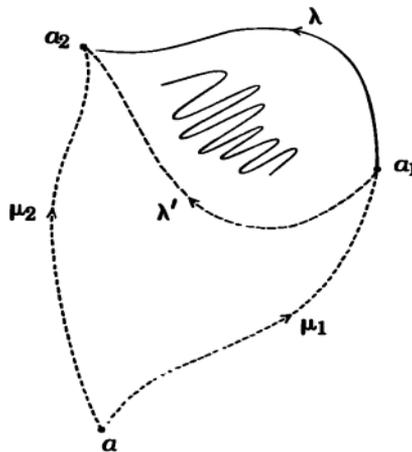


Figure 8.

Exemples d'espaces simplement connexes

- i) Tout espace contractile est simplement connexe (cela résulte immédiatement de 1.2).
- ii) La sphère S^n , pour $n \geq 2$, est simplement connexe.

EXERCICE Le démontrer, en utilisant le fait que $S^n \setminus \{u_0\}$ est homéomorphe à \mathbb{R}^n (par projection stéréographique de centre u_0). Comme \mathbb{R}^n est contractile, tout lacet de S^n dont l'image évite un point u_0 est donc homotope au lacet constant. Malheureusement, il existe des applications continues de I dans S^2 (par exemple) qui sont surjectives (courbes de Peano) ! Pour éviter cette pathologie on pourra démontrer le lemme suivant.

Lemme Tout chemin d'une variété de dimension $n \geq 2$ est homotope à un chemin qui évite un ensemble fini de points donnés à l'avance (en décomposant le chemin donné en une succession finie de « tronçons » inclus dans des domaines de cartes, on se ramène à démontrer le lemme dans les ouverts de \mathbb{R}^n , $n \geq 2$).

Remarque L'exemple ii) est un exemple d'espace simplement connexe qui n'est pas contractile. Pour $n \geq 2$, le fait que S^n ne soit pas contractile peut se démontrer, comme dans le cas $n = 1$, en exhibant un « invariant d'homotopie » non trivial – mais il ne peut pas s'agir dans ce cas du groupe fondamental (cf. chap. 7, § 4.6).

2. Revêtements

Un **revêtement** d'un espace topologique X est un espace fibré de base X à fibre discrète. Le nombre de points de la fibre (si celui-ci est fini) est appelé **degré** du revêtement.

2.0. Exemples de revêtements de degré fini

Proposition i) Soit $P : \mathbb{C} \rightarrow \mathbb{C}$ une application polynomiale de degré $m \geq 1$.

Notons $C_P = \{z \mid P'(z) = 0\}$ l'ensemble (évidemment fini) des points critiques de P , et $V_P = P(C_P)$ l'ensemble des valeurs critiques.

Posons $X = \mathbb{C} \setminus V_P$. Alors la restriction à $Z := P^{-1}(X)$ de l'application P définit une application :

$$p : Z \rightarrow X$$

qui est un revêtement de degré m .

Preuve

Pour tout $u \in X$, $Z_u := p^{-1}(u)$ est l'ensemble des racines de l'équation $P(z) = u$.

Lemme Le nombre d'éléments de l'ensemble Z_u est égal à m .

(c'est clair d'après d'Alembert, car par construction de X l'équation $P(z) = u$ pour $u \in X$ n'a que des racines simples).

Pour construire une trivialisations de (Z, p) au voisinage d'un point $u_0 \in X$, numérotions les éléments de Z_{u_0} :

$$Z_{u_0} = \{z_1^0, \dots, z_m^0\}.$$

Comme Z est un espace topologique séparé, les points z_1^0, \dots, z_m^0 admettent m voisinages disjoints V_1, \dots, V_m , qui d'après le théorème d'inversion locale peuvent être choisis assez petits pour que p envoie chaque V_i difféomorphiquement sur un

ouvert U_i de X . Quitte à remplacer chaque U_i par $U = \bigcap_i U_i$, et chaque V_i par son intersection avec $p^{-1}(U)$, on obtient ainsi des difféomorphismes :

$$p_i := p|_{V_i} : V_i \rightarrow U.$$

Pour tout $u \in U$, la fibre Z_u contient les m points $z_i(u) = p_i^{-1}(u)$ ($i = 1, 2, \dots, m$), et n'en contient pas d'autres d'après le lemme. L'application :

$$\begin{array}{ccc} U \times \{1, 2, \dots, m\} & \rightarrow & Z_U = p^{-1}(U) \\ u, i & \mapsto & z_i(U) \end{array}$$

est donc bijective, et définit la trivialisat on cherch ee de Z_U : c'est en effet un diffeomorphisme, commutant avec les projections  videntes, entre Z_U et $U \times \{1, 2, \dots, m\}$ (union disjointe de m « copies » de U). \square

Cas particulier $P(z) = z^m$: on obtient dans ce cas un revêtement de degré m de $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$:

$$\begin{array}{ccc} \mathbb{C}^* & \rightarrow & \mathbb{C}^* \\ z & \mapsto & u = z^m. \end{array}$$

La fibre au-dessus de u est l'ensemble des racines m -ièmes de u . Ces racines se permutent circulairement quand u fait un tour autour de 0.

En restreignant cet exemple   l'ensemble des nombres complexes de module 1, on obtient un revêtement de degré m du cercle trigonométrique S^1 , ayant S^1 pour espace total :

$$\begin{array}{ccc} S^1 & \rightarrow & S^1 \\ e^{i\theta} & \mapsto & e^{im\theta} \end{array}$$

Digression topologique : homéomorphismes locaux et revêtements Une lecture attentive de la démonstration précédente montre que nous avons démontré un résultat plus général :

PROPOSITION i) Soit $p : Z \rightarrow X$ un homéomorphisme local tel que Z soit séparé, et tel que le nombre d'éléments de l'ensemble $p^{-1}(u)$ soit un nombre fini m indépendant de $u \in X$. Alors p est un revêtement de degré m .

La figure 9 montre un exemple simple d'homéomorphisme local qui n'est pas un revêtement. Pour $x > 0$, la fibre Z_x consiste en un seul point, mais ce point tend vers l'infini quand $x \rightarrow 0$, et pour $x \leq 0$, $Z_x = \emptyset$.

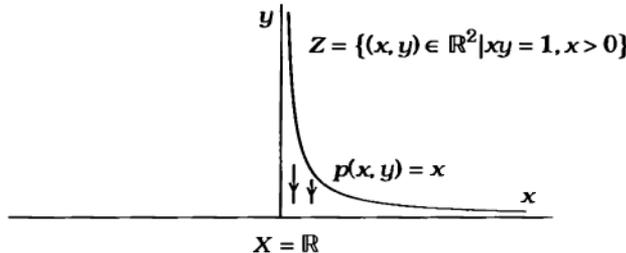


Figure 9. Exemple d'homéomorphisme local qui n'est pas un revêtement.

Le fait qu'un point de la fibre puisse tendre vers l'infini alors que x reste dans un compact signifie que l'application p n'est pas propre (rappelons qu'une application est dite « propre » si l'image réciproque de tout compact est compacte). La proposition suivante montre que la « non-propreté » est la principale raison qui puisse empêcher un homéomorphisme local d'être un revêtement (de degré fini).

PROPOSITION ii) Soit $p : Z \rightarrow X$ un homéomorphisme local propre tel que Z soit séparé. Supposons de plus que X soit localement compact, et connexe. Alors p est un revêtement (de degré fini).

Preuve

Il s'agit d'adapter la preuve de la proposition i), sans savoir à l'avance que le nombre d'éléments de $p^{-1}(u)$ est constant. Pour $u_0 \in X$, notons z_1^0, \dots, z_m^0 les éléments de Z_{u_0} (en nombre fini par hypothèse de propreté). On construit comme précédemment les homéomorphismes $p_i : V_i \rightarrow U$, $i = 1, 2, \dots, m$, et le problème est de démontrer que si U a été choisi assez petit, $p^{-1}(U)$ coïncide avec $\bigcup_i V_i$. Commençons par choisir U assez petit pour que \bar{U} soit compact (c'est à cette fin que nous avons supposé X localement compact). $p^{-1}(\bar{U})$ est alors compact par hypothèse de propreté. En lui enlevant l'ouvert $\bigcup_i V_i$ on obtient un compact $K = p^{-1}(\bar{U}) \setminus \bigcup_i V_i$, et il ne reste qu'à remplacer U par $U' = U \setminus p(K)$. \square

EXERCICE Terminer la démonstration ; à quoi sert l'hypothèse de connexité de X ?

Où l'on retrouve le théorème de d'Alembert : tout polynôme P de degré $m \geq 1$ admet au moins une racine complexe.

La proposition ii) nous donne un moyen de redémontrer la proposition i) sans utiliser le théorème de d'Alembert, en remarquant que l'application $P : \mathbb{C} \rightarrow \mathbb{C}$ est propre (ce qui se vérifie par majorations directes, ou bien en invoquant le fait que P s'étend en une application holomorphe donc continue de la sphère de Riemann dans elle-même, envoyant l'infini sur l'infini : cf. étude 4^{bis}).

On obtient « en prime » une preuve du théorème de d'Alembert, sous la forme suivante : toute application polynomiale non constante $P : \mathbb{C} \rightarrow \mathbb{C}$ est surjective.

En effet, de deux choses l'une : ou bien $u \in V_p$ (ensemble des valeurs critiques de P), et alors $P^{-1}(u)$ n'est pas vide ; ou bien $u \notin V_p$ et alors la proposition ii) montre que le nombre d'éléments de $P^{-1}(u)$ ne dépend pas de u ; ce nombre n'est donc jamais nul, sinon $P^{-1}(X)$ serait vide, ce qui est impossible pour un polynôme non constant (P' n'étant pas identiquement nul, il est impossible que tous les points de \mathbb{C} soient critiques !).

[L'idée de cette démonstration est empruntée au livre de J. Milnor : *Topology from the differentiable Viewpoint*, The University Press of Virginia, 1965.]

QUESTION TEST Soit $p : \mathbb{C} \setminus \{+1, -1\} \rightarrow \mathbb{C}$ l'application $z \mapsto u = \frac{z^3}{3} - z$. Est-ce un difféomorphisme local ? Est-ce un revêtement ?

2.0*. Exemple de revêtement de degré infini :
« la surface de Riemann » du logarithme complexe

Soit $\exp : \mathbb{C} \rightarrow \mathbb{C}^* = \mathbb{C} \setminus \{0\}$ l'application qui à z associe e^z .

PROPOSITION \exp est un revêtement, de fibre \mathbb{Z} .

(La preuve, facile, est laissée au lecteur.)

Pour tout $u \in \mathbb{C}^*$, la fibre $\exp^{-1}(u)$ se compose d'une infinité de points de \mathbb{C} (les « déterminations du logarithme complexe ») : notant $\text{Log } u$ un point arbitraire de $\exp^{-1}(u)$, on a :

$$\exp^{-1}(u) = \{\text{Log } u + 2i \pi k \mid k \in \mathbb{Z}\}.$$

Quand u tourne autour de 0, les différentes déterminations du logarithme s'échangent ($k \mapsto k + 1$), de sorte qu'il est impossible de définir la fonction Log globalement sur \mathbb{C}^* .

En restreignant ce revêtement au-dessus du cercle trigonométrique, on obtient un revêtement de degré infini du cercle qui, par un changement de variable évident ($z = i\theta$), s'identifie au paramétrage usuel du cercle par l'angle :

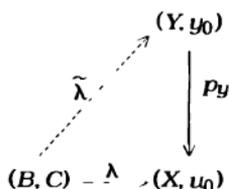
$$\begin{array}{ccc} p & \mathbb{R} & \rightarrow S^1 \\ & \theta & \mapsto e^{i\theta} \end{array}$$

La fibre $p^{-1}(u)$ est l'infinité des « déterminations de l'argument $\arg u$ » :

$$p^{-1}(u) = \{\arg u + 2 \pi k \mid k \in \mathbb{Z}\}.$$

2.1. Relèvement des chemins et des homotopies,
avec application au calcul de $\pi_1(S^1)$

Soit $p_Y : (Y, y_0) \rightarrow (X, u_0)$ un « revêtement pointé », c'est-à-dire un revêtement dont l'espace total est muni d'un point-base y_0 , d'image $p_Y(y_0) = u_0$. « Reléver » dans (Y, y_0) une application $\lambda : (B, C) \rightarrow (X, u_0)$, c'est trouver une application de paires $\tilde{\lambda}$ rendant commutatif le diagramme :



Bien entendu, nous travaillons dans la catégorie des (paires d')espaces topologiques, et « application » signifie application continue.

Théorème de relèvement des chemins.—

Toute application $\lambda : (I, 0) \rightarrow (X, u_0)$ (où $I = [0, 1]$) se relève de façon unique dans (Y, y_0) .

Théorème de relèvement des homotopies.–

Toute application $\Lambda : (I \times I, 0 \times I) \rightarrow (X, u_0)$ se relève de façon unique dans (Y, y_0) .

Corollaire L'extrémité du chemin $\tilde{\lambda}$ qui relève λ ne dépend que de la classe d'homotopie de λ dans $[I, X]_{u_0, u_1}$ (où $u_1 = \lambda(1)$).

En effet, le relèvement d'une homotopie $(\lambda_s)_{s \in [0,1]}$ à extrémités fixes ($\lambda_s(0) = u_0, \lambda_s(1) = u_1$) définit une famille continue $(\tilde{\lambda}_s)$ de chemins de Y dont les extrémités se projettent sur un même point u_1 ; ces extrémités sont donc indépendantes de s par raison de continuité.

Il résulte de ce corollaire que tout revêtement est muni canoniquement d'une famille de bijections $\phi_{[\lambda]}$ entre ses fibres, indexées par les classes d'homotopie de chemins : pour tout $[\lambda] \in [I, X]_{u_0, u_1}$, $\phi_{[\lambda]}$ est l'application de Y_{u_0} dans Y_{u_1} qui, à y_0 , associe l'extrémité du chemin d'origine y_0 qui relève λ ; le fait qu'il s'agisse d'une bijection résulte de ce que $\phi_{[\lambda^{-1}]} = (\phi_{[\lambda]})^{-1}$, conséquence du fait que ϕ est un « homomorphisme de groupoïdes » :

$$\begin{aligned}\phi_{[\lambda^{-1}][\lambda]} &= \phi_{[\lambda^{-1}]} \circ \phi_{[\lambda]} \\ \phi_{[e_{u_0}]} &= \mathbb{1}_{Y_{u_0}}.\end{aligned}$$

Application au calcul de $\pi_1(S^1)$ Soit $p : \mathbb{R} \rightarrow S^1$ ($\theta \mapsto e^{i\theta}$) le revêtement de degré infini du cercle trigonométrique (cf. § 2.0') ; munissons-le du point-base $\theta_0 = 0$, d'image $u_0 = 1$.

D'après le théorème i), tout lacet :

$$\lambda : (I, \partial I) \rightarrow (S^1, u_0)$$

se relève de façon unique en un chemin :

$$\tilde{\lambda} : I \rightarrow \mathbb{R} \text{ tel que } \tilde{\lambda}(0) = 0.$$

Comme l'extrémité $\tilde{\lambda}(1)$ vérifie $e^{i\tilde{\lambda}(1)} = 1$, elle peut s'écrire $\tilde{\lambda}(1) = 2\pi n$, où $n \in \mathbb{Z}$ est un entier appelé indice de λ , qui d'après le corollaire du théorème ii) ne dépend que de la classe d'homotopie de λ .

Or tout chemin de \mathbb{R} est évidemment homotope à un « mouvement rectiligne uniforme ». Le relevé $\tilde{\lambda}$ d'un lacet λ d'indice n est donc homotope à l'application linéaire $t \mapsto 2\pi int$ (cf. Fig. 10), de sorte que λ est homotope au lacet $\lambda_n(t) = e^{2\pi i nt}$, qui parcourt le cercle à vitesse angulaire constante.

Il en résulte que l'application

$$\text{ind} : \pi_1(S^1, u_0) \rightarrow \mathbb{Z}$$

(qui à toute classe de lacet associe son indice) est bijective, et l'on vérifie immédiatement que c'est un isomorphisme de groupes.

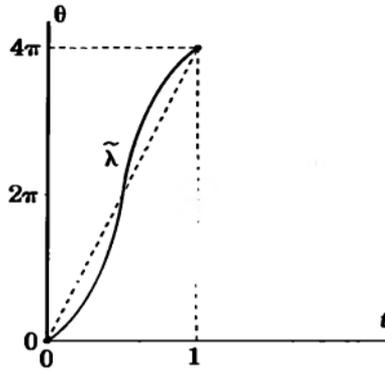


Figure 10. Relevé d'un lacet d'indice 2.

Conclusion $\pi_1(S^1) = \mathbb{Z}$.

N.B. : Lorsque le groupe fondamental est commutatif, omettre de noter le point-base ne peut conduire à aucune ambiguïté (cf. en 1.1 le commentaire illustré par la figure 7).

Preuve des théorèmes de relèvement

En remplaçant le revêtement Y par $Z = \lambda^*Y$ [resp. Λ^*Y] (relèvement de base $B = I$ [resp. $I \times I$] induit par le changement de base λ [resp. Λ]) on ramène le problème à un problème d'extension de section : posant $C = \{0\}$ [resp. $\{0\} \times I$], il s'agit de construire une section $\sigma : B \rightarrow Z$ prenant sur C une valeur assignée à l'avance.

On commencera par remarquer que de tels problèmes d'extension de section sont évidents à résoudre dans le cas d'un revêtement trivial.

Lemme Soit Z un revêtement trivial de base B connexe, et soit C une partie connexe de B . Alors toute section de Z_C (restriction du revêtement Z au-dessus de C) s'étend de façon unique en une section de Z .

Preuve du lemme

C'est évident après trivialisaton, car les sections de $Z = B \times F$ [resp. $Z_C = C \times F$] s'identifient aux applications continues de B [resp. C] dans F , nécessairement constantes puisque la source est connexe et le but discret. □

Revenons à notre problème, avec $Z = \lambda^*Y$ [resp. Λ^*Y]. Comme Z est localement trivial et que sa base B est compacte, on peut subdiviser celle-ci en un nombre fini de « cellules » (segments dans le cas $B = I$, carrés dans le cas $B = I \times I$) au-dessus de chacune desquelles Z est un revêtement trivial. En appliquant le lemme « de proche en proche » aux cellules prises dans un ordre convenable (cf. Fig. 11), on étend à B la section donnée initialement sur $\{0\}$ [resp. sur $\{0\} \times I$].

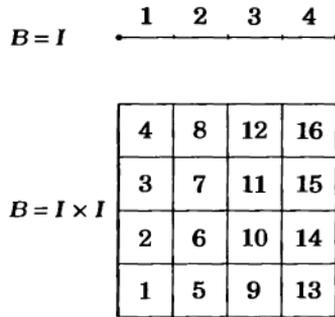


Figure 11. Ordre de succession des cellules de B .

□

2.2. Principe de monodromie, et revêtement universel

La démonstration précédente des théorèmes de relèvement s'étendrait facilement à des espaces tels que l'hypercube $I \times \dots \times I$, par exemple. Mais il est plus intéressant de déduire des théorèmes 2.1 i) et ii) le « principe » très général que voici.

Principe de monodromie

■ *Premier énoncé* : si B est simplement connexe, toute application :

$$\lambda : (B, b_0) \rightarrow (X, u_0)$$

se relève de façon unique dans (Y, y_0) (revêtement dont X est la base).

■ *Deuxième énoncé* : tout revêtement Z dont la base B est simplement connexe est trivial.

Ces deux énoncés sont manifestement équivalents (le premier se déduit du second par changement de base). Prouvons le second.

Preuve

Appliquant au revêtement Z le corollaire des théorèmes de relèvement, on obtient une famille de bijections entre ses fibres :

$$\phi_{[b_1 b_0]} : Z_{b_0} \rightarrow Z_{b_1},$$

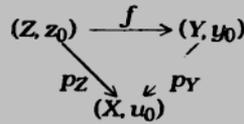
où l'on a noté $[b_1 b_0]$ l'unique classe d'homotopie de chemins d'origine b_0 et d'extrémité b_1 dans l'espace B (simplement connexe). Fixant l'origine b_0 et faisant varier l'extrémité b , on vérifie facilement que la bijection $\phi_{[b b_0]}$ « dépend continûment de b », en ce sens que l'application $\phi : B \times Z_{b_0} \rightarrow Z$ définie par $\phi(b, z_0) = \phi_{[b b_0]}(z_0)$ est une application continue (grâce à la loi « d'homomorphisme de groupoïdes », il suffit de le vérifier pour b assez proche de b_0 , ce qui permet d'utiliser une trivialisat-ion locale de Z).

Il s'ensuit facilement que ϕ est un isomorphisme de revêtements de base B , ce qui prouve notre énoncé. □

L'énoncé qui va être formulé maintenant est une simple paraphrase du principe de monodromie (premier énoncé).

Propriété universelle des revêtements simplement connexes

Soit $(Z, z_0) \xrightarrow{p_Z} (X, u_0)$ un revêtement pointé dont l'espace total Z est simplement connexe. Alors pour tout revêtement pointé $(Y, y_0) \xrightarrow{p_Y} (X, u_0)$ il existe un unique morphisme de revêtements :



EXERCICE Expliciter ce morphisme dans le cas où Y [resp. Z] est le revêtement de degré m [resp. de degré ∞] de C^* ou de S^1 introduit en 2.0 [resp. 2.0']. Pourquoi est-il nécessaire de « pointer » les revêtements pour avoir l'unicité ?

Corollaire *Un revêtement pointé simplement connexe est unique à isomorphisme près.*

DÉFINITION On appelle **revêtement universel** de (X, u_0) un revêtement pointé simplement connexe de (X, u_0) .

L'existence du revêtement universel (sous des hypothèses très générales) sera démontrée en 2.3 ci-après.

2.3. Classification des revêtements de base X par les sous-groupes de $\pi_1(X, u_0)$

Soit X un espace topologique connexe par arcs, et soit $Y \xrightarrow{p} X$ un revêtement de X . Particularisant aux lacets d'origine u_0 le corollaire des théorèmes de relèvement du paragraphe 2.1, on obtient une action du groupe $\pi_1(X, u_0)$ sur Y_{u_0} . Il résulte immédiatement de sa définition que cette action est transitive si et seulement si Y est connexe par arcs, ce que nous supposons désormais.

Étudions le stabilisateur d'un point $y_0 \in Y_{u_0}$: c'est par définition le sous-groupe :

$$H_{y_0} = \{ [\lambda] \in \pi_1(X, u_0) \mid \phi_{[\lambda]}(y_0) = y_0 \} ;$$

autrement dit, c'est l'ensemble des classes d'homotopie de lacets de (X, u_0) dont les relevés d'origine y_0 sont des lacets de Y : c'est donc l'image de l'homomorphisme canonique $p_* = \pi_1(p)$:

$$H_{y_0} = \text{Im} (p_* : \pi_1(Y, y_0) \rightarrow \pi_1(X, u_0)) .$$

Remarquons que cet homomorphisme est injectif : en effet si un lacet $\tilde{\lambda}$ de Y a pour image dans X un lacet λ homotope au lacet constant, il résulte du théorème de relèvement des homotopies que $\tilde{\lambda}$ est homotope au lacet constant.

Nous avons prouvé que $\pi_1(Y, y_0)$ s'identifie (par p_*) à un sous-groupe de $\pi_1(X, u_0)$, qui n'est autre que le stabilisateur H_{y_0} de y_0 .

Il en résulte notamment que l'action de $\pi_1(X, u_0)$ sur Y_{u_0} est simplement transitive (trivialité de H_{y_0}) si et seulement si Y est simplement connexe (trivialité de $\pi_1(Y, y_0)$), c'est-à-dire si Y est un revêtement universel.

Exemple : calcul du groupe fondamental du plan projectif Considérons l'application canonique de la sphère S^n sur l'espace projectif réel \mathbb{P}^n :

$$S^n \xrightarrow{p} \mathbb{P}^n$$

qui au point x de $S^n = \{x \in \mathbb{R}^{n+1} \mid x_0^2 + x_1^2 + \dots + x_n^2 = 1\}$ associe la droite vectorielle de \mathbb{R}^{n+1} passant par x . Il est clair qu'il s'agit d'un revêtement de degré 2 (deux points diamétralement opposés de la sphère sont envoyés sur le même point de \mathbb{P}^n). Pour $n \geq 2$, on sait que S^n est simplement connexe (cf. § 1.3 b)), de sorte que ce revêtement est universel.

Conclusion $\pi_1(\mathbb{P}^n) = \mathbb{Z}/2\mathbb{Z}$ (pour $n \geq 2$).

En effet $\mathbb{Z}/2\mathbb{Z}$ est le seul groupe qui agisse de façon simplement transitive sur un ensemble à deux éléments.

EXERCICE Dessiner ce qui se passe dans le cas du plan projectif ($n = 2$) : représenter un lacet λ dont la classe engendre $\pi_1(\mathbb{P}^2)$, et « dessiner une homotopie » entre $\lambda\lambda$ et le lacet constant.

Remarque fondamentale Revenons au cas général d'un revêtement Y connexe par arcs, muni d'un point-base y_0 . Tout point $y \in Y$ peut alors s'écrire $y = \phi_{[\lambda]}(y_0)$, où $[\lambda]$ est une classe de chemins d'origine $u_0 = p(y_0)$ qui est déterminée par y à multiplication à droite près par H_{y_0} (dire que $\phi_{[\lambda]}(y_0) = \phi_{[\lambda']}(y_0)$, c'est dire que λ et λ' ont même extrémité et que $[\lambda']^{-1}[\lambda] \in H_{y_0}$).

Cette simple remarque est lourde de conséquences. Tout d'abord, elle suggère une méthode générale de construction d'un revêtement pointé ayant pour stabilisateur un sous-groupe $H \subset \pi_1(X, u_0)$ quelconque donné à l'avance : il suffit de définir Y comme l'ensemble des « classes à droite », modulo H_{y_0} , de classes d'homotopie de chemins d'origine et d'extrémité quelconque dans X ; bien entendu il faudrait s'assurer que cet ensemble, muni d'une topologie convenable, est bien un revêtement ayant la propriété annoncée... ce qui sera vrai sous l'hypothèse très générale que voici.

DÉFINITION Un espace topologique X est dit **localement simplement connexe** si tout point admet un système fondamental de voisinages simplement connexes (par exemple toute variété topologique est localement simplement connexe, car localement contractile).

Détaillons le cas du revêtement universel (le cas général s'en déduit facilement par passage au quotient par la relation d'équivalence « multiplication à droite par H »).

PROPOSITION Soit X un espace topologique connexe et localement simplement connexe. Alors X admet un revêtement universel, dont l'espace total peut être défini comme l'ensemble des classes d'homotopie de chemins d'origine u_0 et d'extrémité fixée quelconque dans X .

Remarque La construction fait intervenir un « point marqué » u_0 de X , et le revêtement universel ainsi construit aura pour point marqué la classe du chemin constant égal à u_0 .

Preuve de la proposition

Pour tout $u \in X$, posons $\tilde{X}_u^{u_0} = [I, X]_{u_0, u}$, et $\tilde{X}^{u_0} = \coprod_{u \in X} \tilde{X}_u^{u_0}$. On obtient ainsi un

ensemble projeté $\tilde{X}^{u_0} \xrightarrow{p} X$, dont il reste à construire la topologie. Pour tout ouvert simplement connexe $U \subset X$, considérons sur $\tilde{X}_U^{u_0} = p^{-1}(U)$ la relation de « U -proximité » suivante : $[\lambda']$ est « U -proche » de $[\lambda]$ si $[\lambda'] = [\mu\lambda]$, où μ est un chemin de U (Fig. 12).

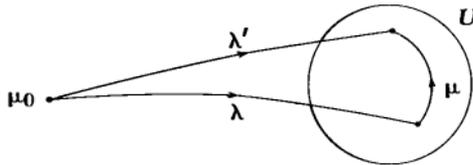


Figure 12. Chemins « U -proches ».

Il s'agit évidemment d'une relation d'équivalence, qui partage $\tilde{X}_U^{u_0}$ en « classes de U -proximité » dont chacune est mise en bijection avec U par la projection p . En appelant « ouverts » de \tilde{X}^{u_0} les unions quelconques et intersections finies de classes de U -proximité (U parcourant l'ensemble de tous les ouverts simplement connexes de X), on vérifie – grâce à l'hypothèse de simple connexité locale de X – qu'on obtient bien une topologie qui fait de \tilde{X}^{u_0} un revêtement simplement connexe de X . □

Nous en avons assez dit pour rendre vraisemblable le théorème suivant, qui résume et complète toute la discussion précédente.

Théorème Soit X un espace topologique connexe, localement simplement connexe, avec point-base u_0 . Alors les revêtements pointés de base (X, u_0) peuvent être classifiés, à isomorphisme de revêtements pointés près, par les sous-groupes de $\pi_1(X, u_0)$.

Plus précisément, pour tout sous-groupe $H \subset \pi_1(X, u_0)$ on peut construire un revêtement pointé (Y, y_0) ayant H pour stabilisateur. De plus, ce revêtement est unique à isomorphisme (de revêtements pointés) près : plus généralement, pour tout couple de revêtements pointés $(Y, y_0), (Y', y'_0)$ dont les stabilisateurs sont inclus l'un dans l'autre ($H \subset H'$), il existe un unique morphisme de revêtements pointés $f : (Y, y_0) \rightarrow (Y', y'_0)$.

Exemple \mathbb{C}^* (ou S^1) n'admet pas d'autres revêtements connexes (à isomorphisme près) que ceux déjà introduits en 2.0, 2.0' : en effet $\pi_1(\mathbb{C}^*) = \pi_1(S^1) = \mathbb{Z}$ n'a pas d'autres sous-groupes que $m\mathbb{Z}$ ($m = 1, 2, \dots$) ou le groupe nul (qui correspond au revêtement universel).

EXERCICE Expliciter tous les morphismes entre ces revêtements.

2.4. Revêtements galoisiens

Soit $Y \xrightarrow{P} X$ un revêtement dont l'espace total et la base sont connexes par arcs. On sait alors que pour tout $u_0 \in X$, $\pi_1(X, u_0)$ agit transitivement sur Y_{u_0} , de sorte que « par changement de point-base » y_0 les stabilisateurs H_{y_0} ($y_0 \in Y_{u_0}$) sont tous conjugués. S'ils sont tous égaux, c'est-à-dire si $H := H_{y_0}$ est un sous-groupe distingué de $\pi_1(X, u_0)$, le revêtement est dit **galoisien**.

Considérons alors, dans le groupoïde des classes d'homotopie de chemins de X , l'action à droite de $\pi_1(X, u_0)$ sur l'ensemble \tilde{X}^{u_0} des classes d'homotopie de chemins d'origine u_0 et d'extrémité quelconque.

Comme H est un sous-groupe distingué, cette action est compatible avec la relation d'équivalence « multiplication à droite par H » dans \tilde{X}^{u_0} .

La remarque fondamentale 2.3, qui nous permet d'identifier Y à l'ensemble des classes à droite de \tilde{X}^{u_0} modulo H , nous fournit ainsi une action de $\pi_1(X, u_0)$ comme groupe d'automorphismes du revêtement. Le noyau de cette action est évidemment H , de sorte que les revêtements galoisiens jouissent de la propriété fondamentale suivante, dont on montre qu'elle les caractérise.

Propriété caractéristique des revêtements galoisiens

Un revêtement est galoisien si et seulement s'il est muni d'un groupe d'automorphismes G qui agit de façon simplement transitive sur chacune de ses fibres.

Ce groupe, isomorphe au quotient de $\pi_1(X, u_0)$ par le sous-groupe distingué H , est appelé **groupe de Galois du revêtement**.

Exemples : tout revêtement universel est galoisien, de groupe de Galois $\pi_1(X, u_0)$.

Si $\pi_1(X, u_0)$ est commutatif, tout revêtement de X est galoisien : par exemple le revêtement de degré m de \mathbb{C}^n ou de S^1 admet comme groupe de Galois le groupe cyclique $\mathbb{Z}/m\mathbb{Z}$, agissant comme groupe de permutations circulaires des racines m -ièmes.

Exemple de revêtement non galoisien Effectuons la construction 2.0 dans le cas du polynôme $P(z) = z - \frac{z^3}{3}$.

On a :

$$C_P = \{+1, -1\}$$

$$V_P = \left\{ -\frac{2}{3}, \frac{2}{3} \right\}.$$

$$p : Z = \mathbb{C} \setminus P^{-1}(V_P) \rightarrow X = \mathbb{C} \setminus V_P.$$

Prenons pour point-base de X le point $u_0 = 0$. La fibre Z_{u_0} se compose de trois points, qui pour $u = u_0$ sont $z_0 = 0, z_+ = +\sqrt{3}, z_- = -\sqrt{3}$. La figure 13 permet de suivre l'évolution des points de la fibre quand u reste réel : quand u se rapproche d'une des valeurs critiques $\pm 2/3$, deux des points de la fibre tendent l'un vers l'autre ; de plus ces deux points s'échangent quand u décrit un petit cercle autour de la valeur critique (en effet, comme le point critique est quadratique, tout se passe localement comme pour l'application $z \mapsto u = z^2$).

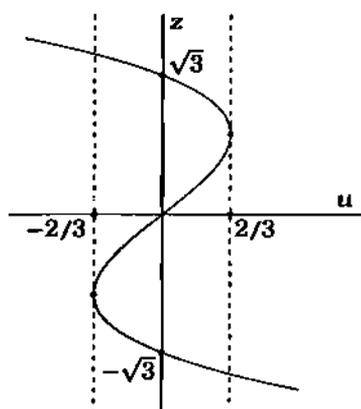


Figure 13.

$\pi_1(X, u_0)$ est le groupe libre (non commutatif) engendré par les deux lacets λ_+, λ_- de la figure 14.

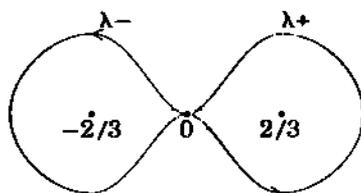


Figure 14.

Son action peut être décrite ainsi :

λ_+ échange z_0 et z_+ , et laisse z_- invariant ;

λ_- échange z_0 et z_- , et laisse z_+ invariant.

Il est clair que le revêtement n'est pas galoisien (λ_+ appartient au stabilisateur de z_- mais pas à celui de z_+).

EXERCICE (de « combinatoire des groupes ») Décrire explicitement les stabilisateurs respectifs de z_0, z_+ et z_- (on pourra commencer par « déblayer le terrain » en travaillant modulo le sous-groupe engendré par λ_+^2 et λ_-^2 , qui stabilise chacun des trois points).

EXERCICE TEST sur le paragraphe 2 Que pouvez-vous dire de la façon dont varient les racines complexes de l'équation $z^2 + 10^{-8}z^5 + u = 0$ quand u parcourt le disque unité épointé $0 < |z| < 1$?

3. Orientations

3.0. Orientations d'un espace vectoriel de dimension finie

Soit E un espace vectoriel réel de dimension finie n . Étant données deux bases $e = (e_1, \dots, e_n)$ et $e' = (e'_1, \dots, e'_n)$ de E , le déterminant de la matrice de changement de base est un nombre réel non nul. Selon qu'il est positif ou négatif, on dira que les deux bases e et e' ont la même orientation (en abrégé or $e = \text{or } e'$) ou qu'elles sont d'orientation opposée (en abrégé or $e = -\text{or } e'$). « Avoir la même orientation » est une relation d'équivalence qui partage l'ensemble des bases de E en deux classes. Se donner une **orientation** de E , c'est choisir l'une de ces deux classes ; les bases appartenant à la classe ainsi choisie seront dites « orientées positivement », celles de l'autre classe seront dites « orientées négativement ».

Cas particuliers remarquables

$n = 1$. Orienter une droite vectorielle équivaut à choisir un « sens de parcours » sur cette droite (choisir, parmi les deux classes de vecteurs non nuls, ceux qui seront dits « orientés positivement »).

$n = 2$. Orienter un plan vectoriel équivaut à choisir un « sens de rotation autour de l'origine », c'est-à-dire à choisir un isomorphisme entre $\pi_1(E \setminus \{0\})$ et \mathbb{Z} : il y a deux tels isomorphismes, qui se déduisent l'un de l'autre par l'automorphisme $n \mapsto -n$ dans \mathbb{Z} ; le choix d'une base de E , autrement dit d'un isomorphisme entre E et \mathbb{R}^2 , détermine un tel isomorphisme entre $\pi_1(E \setminus \{0\})$ et $\pi_1(\mathbb{R}^2 \setminus \{0\}) = \mathbb{Z}$ qui ne dépend que de l'orientation de la base.

$n = 3$. La distinction entre « trièdres orientés positivement » et « trièdres orientés négativement » dans notre espace usuel à trois dimensions correspond dans la vie pratique à la distinction entre deux sortes de tire-bouchons : ceux qu'on achète chez les quincailliers et ceux qu'on achète dans les boutiques de farces attrapes.

QUESTION TEST

Pourquoi, quand on se regarde dans un miroir, voit-on sa droite à sa gauche et pas sa tête à ses pieds ?

Remarque : comment orienter un sous-espace vectoriel ?

Dans un espace vectoriel E de dimensions n convenons d'appeler *base adaptée* à un sous-espace vectoriel F une base de E dont les $n - p$ derniers vecteurs forment une base de F ($p = \text{codim}F$). Si l'on ne considère que des bases de E adaptées à F , le déterminant d'un changement de base est le produit de deux déterminants : celui du changement de base induit sur le sous-espace F , et celui du changement de base induit sur l'espace quotient E/F .

Travailler avec des bases adaptées permet ainsi de mettre en relation les orientations des trois espaces vectoriels $E, F, E/F$ de façon telle que les orientations de deux d'entre eux déterminent une orientation du troisième. Ainsi par exemple, si F est un *hyperplan vectoriel* d'un espace vectoriel orienté E , orienter F équivaut à orienter l'espace à une dimension E/F , c'est-à-dire à se donner un « sens de traversée » de l'hyperplan F dans E .

3.1. Orientations d'un fibré vectoriel

Soit $E = (E \xrightarrow{\pi_E} M)$ un fibré vectoriel de rang n , de base M . « Orienter » le fibré E , c'est associer à tout $a \in M$ une orientation de l'espace vectoriel $E_a = \pi_E^{-1}(a)$ « dépendant continûment de a » en un sens que nous allons préciser. Il va s'avérer que c'est toujours possible *localement* ; si c'est possible globalement le fibré vectoriel est dit **orientable**.

Orientations locales de E

Plaçons-nous sur un ouvert $U \subset M$ tel que le fibré induit E_U soit trivial. Se donner une trivialisatation de E_U , c'est se donner un n -uplet $e = (e_1, \dots, e_n)$ de sections de E_U partout linéairement indépendantes, autrement dit une « base $e(a)$ de E_a dépendant continûment de $a \in U$ » (nous dirons plus brièvement « une base e de E_U »).

Notant $\text{or } e(a)$ l'orientation de E_a définie par la base $e(a)$, nous obtenons la correspondance

$$\text{or } e : a \mapsto \text{or } e(a)$$

que nous appellerons **orientation de E_U définie par la base e** .

Remarquons que pour toute base e' de E_U la matrice du changement de base a pour déterminant une fonction continue de a dont le signe est localement constant sur U (car le déterminant ne s'annule pas) ; si U est connexe il en résulte que de deux choses l'une :

- ou bien $\text{or } e' = \text{or } e$ sur tout U ;
- ou bien $\text{or } e' = -\text{or } e$ sur tout U .

En considérant toutes les bases possibles de E_U on obtient donc en tout et pour tout deux orientations de E_U , comme dans le cas 3.0.

Orientations de E_a « dépendant continûment de a »

Voici ce que signifiera pour nous cette expression :

- i) si a parcourt un ouvert connexe U sur lequel E_U est trivial, il s'agira d'une correspondance de la forme $\text{or } e$ définie comme ci-dessus par une base e de E_U ;
- ii) si a parcourt un ouvert quelconque de M , il s'agira d'une correspondance qui est *localement du type i*) (« localement » voulant dire « en restriction à tout ouvert U du type i »).

On peut paraphraser cette définition en disant que les orientations [resp. orientations locales] du fibré E sont les sections [resp. sections locales] d'un revêtement de degré 2 de M appelé revêtement des orientations de E , dont voici la définition.

PROPOSITION-DÉFINITION Soit $\text{Or } E_a$ l'ensemble à deux éléments des orientations de E_a . L'ensemble $\text{Or } E := \coprod_{a \in M} \text{Or } E_a$ est muni naturellement d'une topologie qui en fait un revêtement de degré 2 de M , le **revêtement des orientations du fibré vectoriel E** .

Preuve

C'est une conséquence facile des considérations qui précèdent. On définit la topologie de $\text{Or } E$ en prenant comme base d'ouverts les ensembles de la forme

Or $e = \{ \text{or } e(a) \mid a \in U \}$, où e est une base de E_U pour un U choisi comme ci-dessus. \square

3.2. Comment reconnaître qu'un fibré vectoriel est orientable

Si M est connexe par arcs, l'orientabilité d'un fibré vectoriel E de base M se reconnaît à l'action du groupe fondamental $\pi_1(M, a)$ sur le revêtement des orientations de E : ou bien cette action est triviale, ce qui signifie que le revêtement des orientations est trivial ; il admet alors deux sections, qui sont les deux orientations possibles de M . Ou bien il existe un lacet $\lambda \in \pi_1(M, a)$ qui échange les deux orientations de E_a ; alors le fibré vectoriel E n'est pas orientable (son revêtement des orientations n'admet aucune section).

Quelques conditions suffisantes d'orientabilité

- i) Tout fibré vectoriel *trivial* est orientable.
- ii) Tout fibré vectoriel à base simplement connexe est orientable (appliquer le principe de monodromie au revêtement des orientations de E).
- iii) Si E est orientable, le fibré λ^*E induit par tout changement de base est aussi orientable.

Exemple de fibré vectoriel non orientable Considérons dans le plan projectif réel une droite projective $\mathbb{P}^1 \subset \mathbb{P}^2$. Alors son fibré normal $T_{\mathbb{P}^1} \mathbb{P}^2$ n'est pas orientable (en effet, un lacet parcourant \mathbb{P}^1 échange les deux « sens de traversée » de la droite \mathbb{P}^1 dans \mathbb{P}^2 : comparer à l'exemple introductif du chapitre 6).

EXERCICE Démontrer par le même raisonnement que le fibré normal à un hyperplan projectif de \mathbb{P}^n ($n \geq 2$) n'est pas orientable.

À propos de ces exemples, on notera au passage que pour les fibrés vectoriels de rang 1 l'orientabilité équivaut à la trivialité.

Remarque sur l'orientabilité des sous-fibrés vectoriels et fibrés vectoriels quotients Soit F un sous-fibré vectoriel de E , et soit E/F le fibré vectoriel quotient. Alors la donnée d'une orientation de deux des trois fibrés $E, F, E/F$ détermine une orientation du troisième (en restriction à un ouvert U où nos fibrés sont triviaux, cette orientation s'obtient grâce à la remarque du paragraphe 3.0).

3.3. Orientations des variétés

On appelle **orientation** [resp. **revêtement des orientations**] d'une variété différentielle M une orientation [resp. le revêtement des orientations] de son fibré vectoriel tangent TM . La variété M est dite **orientable** si son fibré tangent est orientable.

EXEMPLES

- i) Tout ouvert de \mathbb{R}^n est évidemment orientable, car son fibré tangent est un produit ; on dispose dans ce cas d'une orientation canonique, celle de la base canonique de \mathbb{R}^n .
- ii) Soit $S \subset M$ une sous-variété donnée par un système d'équations globales. Alors S est orientable si et seulement si TM_S l'est (appliquer la remarque du paragraphe 3.2 aux trois fibrés vectoriels $E = TM_S$, $F = TS$, $E/F = T_S M$, en se souvenant que le fibré normal à une sous-variété définie par des équations globales est trivial dont orientable).

Corollaire *L'espace projectif réel \mathbb{P}^n est orientable si et seulement si n est impair.*

En effet, on a vu en 3.2 (exercice) que le fibré normal $T_{\mathbb{P}^{n-1}}\mathbb{P}^n$ n'est pas orientable ; sachant que $\mathbb{P}^1 \approx S^1$ est orientable, on en déduit que \mathbb{P}^2 ne l'est pas, donc que \mathbb{P}^3 l'est, etc.

EXERCICE

i) Si la démonstration précédente vous a facilement convaincu, vous vous êtes laissés escroquer. Voyez-vous à quel endroit ?

Réparez l'escroquerie, en utilisant ce que vous savez sur $\pi_1(\mathbb{P}^n)$ (cf. § 2.3).

ii) En combinant la remarque du paragraphe 3.2 à la « condition suffisante d'orientabilité » du paragraphe 3.2 iii), montrer que le produit de deux variétés orientables est orientable.

EXERCICE TEST Pour quelles valeurs de n le revêtement des orientations de \mathbb{P}^n est-il isomorphe à son revêtement universel ?

3.4. Critère combinatoire d'orientabilité

Soit E un fibré vectoriel de base M , et soit $(U_\alpha)_{\alpha \in A}$ un recouvrement de M par des ouverts tels que E_{U_α} soit trivial. Soit $\tau = (\tau_\alpha)_{\alpha \in A}$ une famille de trivialisations des E_{U_α} . Pour tout couple d'indices α, β on notera $\varepsilon_{\alpha, \beta}^\tau$ la fonction localement constante sur $U_\alpha \cap U_\beta$, à valeur ± 1 , définie par le signe du déterminant du changement de base associé au changement de trivialisations $\tau_\beta \tau_\alpha^{-1}$ (cf. § 3.0).

Notant or τ_α l'orientation de E_{U_α} définie par la trivialisations τ_α , on a :

$$\text{or } \tau_\beta = \varepsilon_{\alpha, \beta}^\tau \text{ or } \tau_\alpha \quad \text{sur } U_\alpha \cap U_\beta.$$

Remarque La collection de fonctions $(\varepsilon_{\alpha, \beta}^\tau)_{(\alpha, \beta) \in A \times A}$ vérifie le système de relations :

$$\varepsilon_{\alpha, \beta}^\tau \varepsilon_{\beta, \gamma}^\tau = \varepsilon_{\alpha, \gamma}^\tau \quad (*)$$

sur $U_\alpha \cap U_\beta \cap U_\gamma$.

Un cas particulier remarquable est celui où $\varepsilon_{\alpha, \beta}^\tau = +1$ pour tous α, β , ce qui veut dire :

$$\text{or } \tau_\beta = \text{or } \tau_\alpha \quad \text{sur } U_\alpha \cap U_\beta.$$

On dit alors que τ est une **famille cohérente de trivialisations locales de E** . Dans ce cas les orientations locales $(\text{or } \tau_\alpha)_{\alpha \in A}$ se recollent sur tout M pour définir une orientation de E .

Exemple $E = TM$ (fibré tangent à une variété différentielle). Si l'on considère la famille de trivialisations locales $(\tau_\alpha)_{\alpha \in A}$ de TM définie par un atlas $(x_\alpha)_{\alpha \in A}$ de M , la fonction $\varepsilon_{\alpha, \beta}^\tau$ est le signe du déterminant jacobien du changement de carte $x_\beta \circ x_\alpha^{-1}$. On dit que $(x_\alpha)_{\alpha \in A}$ est un atlas orienté de M si $\varepsilon_{\alpha, \beta}^\tau = 1$ pour tous α, β .

PROPOSITION

- i) La donnée d'une orientation de E équivaut à la donnée d'une famille cohérente de trivialisations locales de E .
- ii) Pour toute famille $\tau = (\tau_\alpha)_{\alpha \in A}$ de trivialisations locales de E telles que les ouverts U_α soient connexes, on reconnaît l'orientabilité de E à l'existence d'un « jeu de signes » $(\sigma_\alpha = \pm 1)_{\alpha \in A}$ tel que :

$$\varepsilon_{\alpha\beta}^\tau = \sigma_\alpha \sigma_\beta \quad \text{sur } U_\alpha \cap U_\beta. \quad (**)$$

Preuve

La première partie de la proposition est évidente. Prouvons la seconde.

Dire qu'un « jeu de signes » $(\sigma_\alpha)_{\alpha \in A}$ vérifie $(**)$ c'est dire que la formule :

$$\text{or } E_\alpha := \sigma_\alpha \text{ or } \tau_\alpha(a)$$

définit une orientation de E_α qui ne dépend pas du choix de α tel que $a \in U_\alpha$. Les orientations locales ainsi définies se recollent donc en une orientation globale.

Inversement, la donnée d'une orientation globale $\text{or } E : M \rightarrow \text{Or } E$ détermine sur les U_α un « jeu de fonctions localement constantes » :

$$\begin{aligned} \sigma_\alpha(a) &= +1 & \text{si } \text{or } E(a) = \text{or } \tau_\alpha(a) \\ &= -1 & \text{si } \text{or } E(a) = -\text{or } \tau_\alpha(a); \end{aligned}$$

comme les U_α ont été supposés connexes, ces fonctions sont en fait constantes et forment « un jeu de signes » qui vérifie les relations $(**)$. \square

Corollaire de la première partie de la proposition *Tout fibré vectoriel complexe de rang n est muni d'une orientation canonique. en tant que fibré vectoriel réel de rang $2n$. En particulier, toute variété analytique complexe de dimension n admet, en tant que variété réelle de dimension $2n$, une orientation canonique.*

Preuve

En effet, un calcul élémentaire montre que tout automorphisme \mathbb{C} -linéaire de \mathbb{C}^n , considéré comme automorphisme \mathbb{R} -linéaire de \mathbb{R}^{2n} , a un déterminant positif, de sorte que toute famille de trivialisations \mathbb{C} -linéaires est cohérente. \square

Autre corollaire *Le produit de deux variétés orientables est orientable (cf. aussi exercice 3.3 ii)).*

En effet, l'atlas produit de deux atlas orientés est évidemment orienté.

Résultat de la « question test » du paragraphe 3.0

Si la question vous a plongé dans un abîme de perplexité, ou si elle vous a fait éclater de rire, vous êtes sur la bonne voie.

Si vous avez réussi à formuler une réponse, est-elle aussi belle que l'était la question ?

Si oui, j'aimerais la connaître ! Sinon, peut-être prenez-vous trop au sérieux les mathématiques ?

ÉTUDE 7

... Où une relecture de l'intermède nous ramène au tore \mathbb{T}^2

Motivation de l'étude

Dans l'intermède sur les indices de champs de vecteurs, nous avons commencé par étudier des champs de vecteurs sur des ouverts de \mathbb{R}^2 . Le fibré tangent à \mathbb{R}^2 étant un produit ($T\mathbb{R}^2 = \mathbb{R}^2 \times \mathbb{R}^2$), la donnée d'un champ de vecteurs non nul sur un ouvert U de \mathbb{R}^2 équivaut à la donnée d'une application $v : U \rightarrow \mathbb{R}^2 \setminus \{0\}$. À tout lacet λ de U correspond donc un lacet $v \circ \lambda$ de $\mathbb{R}^2 \setminus \{0\}$, dont la classe d'homotopie est caractérisée par un « indice » $n \in \mathbb{Z}$ ($\mathbb{R}^2 \setminus \{0\}$ a même type d'homotopie que S^1). Cet indice ne dépend que de la classe d'homotopie de $v \circ \lambda$, ce qui prouve les lemmes 1 et 2 de l'intermède, § 1.2.

Si, au lieu de champs de vecteurs de \mathbb{R}^2 , on considère des champs de vecteurs sur une surface M (variété à deux dimensions), on ne peut plus raisonner comme précédemment car le fibré tangent TM n'est plus un produit. Néanmoins, si l'ouvert U de M sur lequel on travaille est inclus dans un domaine de carte, le fibré TU est trivial. Choissant une trivialisatation de ce fibré, on peut à nouveau identifier les champs de vecteurs v non nuls sur U à des applications de U dans $\mathbb{R}^2 \setminus \{0\}$, ce qui permet comme précédemment de définir l'indice de v le long d'un lacet λ de U . Mais attention : l'indice ainsi défini dépend du choix de la trivialisatation ! Par exemple, on a vu sur l'exemple de la sphère S^2 que l'indice d'un champ de vecteurs non nul le long de l'équateur n'est pas le même selon qu'on le calcule par la trivialisatation « stéréographique nord » ou « stéréographique sud ».

Pour bien cerner la difficulté, nous allons faire quelques réductions. Le fibré qui nous intéresse est λ^*TU (fibré induit de TU par le changement de base λ) ou plus exactement $\lambda^*T'U$, où $T'U$ désigne le complémentaire de la section nulle dans TU .

Comme $\lambda(0) = \lambda(1)$, on peut considérer λ comme une application de S^1 dans U , de sorte que $\lambda^*T'U$ est un fibré de base S^1 et de fibre $\mathbb{R}^2 \setminus \{0\}$.

Comme la notion d'indice est invariante par homotopie, nous pouvons encore simplifier les données en remplaçant $\mathbb{R}^2 \setminus \{0\}$ par S^1 qui a même type d'homotopie : plus précisément, notre fibré de départ $T'U$ peut être remplacé par SU , fibré des directions de demi-droites tangente à U (l'application $T'U \rightarrow SU$, de passage au quotient par la relation d'équivalence entre vecteurs tangents $v \sim v' \iff \exists c > 0, v' = cv$, est une équivalence d'homotopie respectant la structure fibrée). Il résulte de notre hypothèse sur U que tous les fibrés considérés ici sont triviaux. Le fibré induit de SU par le changement de base λ peut être considéré comme un fibré trivial de base S^1 et de fibre S^1 ; son espace total est donc homéomorphe au tore \mathbb{T}^2 .

EXERCICE préliminaire S désigne désormais le cercle trigonométrique. $p : \mathbb{R} \rightarrow S$ désigne le paramétrage du cercle par l'angle. À l'aide du principe de monodromie, on montrera qu'à tout automorphisme φ de $\mathbb{T} = S \times S$ correspond un automorphisme $\tilde{\varphi}$

de $\mathbb{R} \times \mathbb{R}$ tel que le diagramme

$$\begin{array}{ccc} \mathbb{R} \times \mathbb{R} & \xrightarrow{\tilde{\varphi}} & \mathbb{R} \times \mathbb{R} \\ \downarrow p \times p & & \downarrow p \times p \\ \mathbb{T} & \xrightarrow{\varphi} & \mathbb{T} \end{array}$$

soit commutatif. Le mot « automorphisme » peut s'entendre ici aussi bien dans la catégorie des espaces topologiques (homéomorphisme de l'espace dans lui-même) que dans celle des variétés différentielles ou analytiques.

Le tore vu comme « fibré en cercles sur le cercle »

On considère maintenant \mathbb{T} comme fibré de base S , avec la première projection $p_1 : S \times S \rightarrow S$. Pour deviner les réponses aux questions suivantes, on pourra s'aider de dessins dans le plan $\mathbb{R} \times \mathbb{R}$, obtenus en « relevant » les figures de $S \times S$ qui nous intéressent ; il sera commode de représenter la base du fibré \mathbb{T} par le segment $[0, 1]$ (en considérant 0 et 1 comme « identifiés ») et la fibre par la droite \mathbb{R} considérée « modulo les translations de $2k\pi$, $k \in \mathbb{Z}$ ».

i) Deux sections σ_0, σ_1 du fibré \mathbb{T} sont dites **isotopes** si on peut les interpoler par une famille continue de sections $(\sigma_t)_{t \in [0,1]}$ (c'est la même notion que celle d'*homotopie*, sauf que les applications interpolantes sont astreintes à être, comme σ_0 et σ_1 , des sections du fibré). Montrer que les sections de \mathbb{T} sont caractérisées à isotopie près par un entier $n \in \mathbb{Z}$, leur « indice ».

ii) Soit α un automorphisme du fibré \mathbb{T} (au-dessus de \mathbb{I}_S). Les images par α des sections constantes du fibré \mathbb{T} sont toutes isotopes (pourquoi ?). Leur indice sera appelé **indice de l'automorphisme** α , et noté $\text{ind} \alpha$.

Quelle relation y a-t-il entre l'indice d'une section quelconque de \mathbb{T} et celui de sa transformée par un automorphisme α respectant l'orientation ? Même question pour un automorphisme α inversant l'orientation.

iii) Deux automorphismes α_0, α_1 du fibré \mathbb{T} sont dits **isotopes** si on peut les interpoler par une famille continue $(\alpha_t)_{t \in [0,1]}$ d'automorphismes du fibré \mathbb{T} . Montrer que les automorphismes de \mathbb{T} respectant l'orientation sont classifiés à isotopie près par leur indice : plus précisément on a un *isomorphisme de groupes* :

$$\text{ind} : [\text{Aut}^+ \mathbb{T}] \rightarrow \mathbb{Z}$$

où $[\text{Aut}^+ \mathbb{T}]$ désigne le groupe des classes d'isotopie d'automorphismes de \mathbb{T} respectant l'orientation.

Préciser de même la structure du groupe $[\text{Aut} \mathbb{T}]$ de toutes les classes d'isotopie d'automorphismes (respectant ou non l'orientation).

Application : indices des sections d'un fibré trivial en cercles sur le cercle

Soit \mathbb{T} un fibré trivial de base S^1 et de fibre S^1 . Que faut-il entendre par « l'indice d'une section de \mathbb{T} » relativement à une trivialisatation τ de \mathbb{T} ?

Montrer que cet indice dépend du choix de τ , mais que la différence des indices de deux sections σ, σ' n'en dépend pas, à condition toutefois de ne considérer que des trivialisations ayant même orientation. Que devient cette différence pour une trivialisatation d'orientation opposée ?

Ce qui précède justifie le lemme admis au paragraphe 3.2 de l'intermède.

Épilogue : classification des tores complexes

Cherchons à classier, à *isomorphisme analytique complexe près*, les variétés analytiques complexes de la forme \mathbb{C}/Ω , où Ω est un réseau du plan (cf. étude 4^{ter}).

En utilisant l'« exercice préliminaire » ci-dessus et le résultat de l'étude 4^{bis} (problème B) sur la classification des automorphismes de \mathbb{C} , on montrera que :

deux telles variétés \mathbb{C}/Ω et \mathbb{C}/Ω' sont isomorphes si et seulement si les réseaux Ω et Ω' se déduisent l'un de l'autre par une similitude plane directe.

Introduction au calcul intégral sur les variétés

Pour savoir calculer sur les variétés il faut avoir compris comment les calculs dans \mathbb{R}^n se transforment par changements de coordonnées. S'agissant de calcul intégral, il faut donc avoir compris les changements de variables d'intégration. Commençons par rappeler la formule du changement de variable pour une intégrale simple, sur un intervalle de \mathbb{R} :

$$\int_a^b f(x) dx = \int_\alpha^\beta f(\varphi(\xi)) \varphi'(\xi) d\xi \quad (0)$$

Dans cette formule, φ est un difféomorphisme entre deux intervalles ouverts de \mathbb{R} contenant respectivement les points α, β et $a = \varphi(\alpha)$, $b = \varphi(\beta)$: on ne suppose pas nécessairement que $a < b$ ni que $\alpha < \beta$, la convention $\int_b^a = -\int_a^b$ permettant d'écrire (0) sans faire cette hypothèse ; c'est grâce à cette convention qu'on a pu écrire $\varphi'(\xi)$ et non pas $|\varphi'(\xi)|$ dans le membre de droite de (0).

En introduisant la forme différentielle de degré 1 :

$$\omega = f(x) dx,$$

on peut récrire la formule (0) sous forme plus compacte grâce à la notion d'« image réciproque d'une forme différentielle » (Chap. 5, § 4.2) :

$$\int_a^b \omega = \int_\alpha^\beta \varphi^* \omega. \quad (0)'$$

Le changement de variable d'intégration se comprend donc beaucoup mieux si l'on considère que ce ne sont pas les *fonctions* que l'on intègre mais les *formes différentielles*.

Nous connaissons déjà les formes différentielles de degré 1, et avons déjà eu l'occasion de les intégrer sur des variétés de dimension 1, ou sur des chemins dans des variétés de dimension supérieure (cf. étude 3^{bis}, exercices finaux des paragraphes 0 et 1). Ce chapitre va définir les *formes différentielles de degré p*, ainsi que leur intégration sur des « chaînes de dimension p ».

0. L'algèbre des formes extérieures sur un espace vectoriel

Ce paragraphe expose les préliminaires algébriques qui nous permettront, au paragraphe suivant, de définir les formes différentielles de degré quelconque.

0.0. Formes extérieures

Soit E un espace vectoriel de dimension n sur \mathbb{R} (ou sur un corps commutatif quelconque). On appelle **forme p -linéaire alternée sur E** (ou « **forme extérieure de degré p** », ou plus brièvement « **p -forme** » sur E) une application :

$$\omega : \underbrace{E \times \cdots \times E}_{p \text{ fois}} \rightarrow \mathbb{R},$$

linéaire par rapport à chaque facteur, et telle que $\omega(v_1, \dots, v_p) = 0$ toutes les fois que deux des v_i sont égaux. De ces deux propriétés on déduit facilement que ω est *antisymétrique*, c'est-à-dire change de signe quand on intervertit deux des v_i (remplacer v_i et v_j par $v_i + v_j$, et utiliser la linéarité).

Cas particuliers

- $p = 1$ Une 1-forme sur E est tout simplement une forme linéaire sur E . L'espace des 1-formes est donc l'espace vectoriel dual E^* .
 - $p = 2$ Un exemple important de 2-forme est fourni par le « produit extérieur de deux 1-formes » :
- pour $u_1, u_2 \in E^*$, on définit la 2-forme $u_1 \wedge u_2 : E \times E \rightarrow \mathbb{R}$ par

$$(u_1 \wedge u_2)(v_1, v_2) = \det \begin{pmatrix} u_1(v_1) & u_1(v_2) \\ u_2(v_1) & u_2(v_2) \end{pmatrix}.$$

Géométriquement, le nombre $(u_1 \wedge u_2)(v_1, v_2)$ peut s'interpréter comme l'aire du parallélogramme orienté de \mathbb{R}^2 construit sur les vecteurs images de v_1, v_2 par l'application linéaire :

$$u = (u_1, u_2) : E \rightarrow \mathbb{R}^2,$$

avec la convention que l'aire est comptée positivement ou négativement selon l'orientation de la base de \mathbb{R}^2 que constituent les deux vecteurs $(u(v_1), u(v_2))$ supposés linéairement indépendants (s'ils ne le sont pas, l'aire est nulle).

On définirait de même, plus généralement, le produit extérieur de p 1-formes $u_1, \dots, u_p \in E^*$:

$$(u_1 \wedge \cdots \wedge u_p)(v_1, \dots, v_p) = \det(u(v_j))$$

(volume du parallélépipède orienté de \mathbb{R}^p construit sur les vecteurs $u(v_1), \dots, u(v_p)$).

En fait, nous allons généraliser aux formes de degrés quelconques la notion de produit extérieur, de telle façon que la notion de produit extérieur de p 1-formes ne fasse qu'illustrer l'*associativité* de l'opération de produit extérieur de deux formes de degrés quelconques.

0.1. Produit extérieur

Dans le groupe \mathfrak{S}_{p+q} des permutations de $p+q$ objets, notons $\mathfrak{S}_{(p,q)}$ le sous-ensemble des « imbrications » :

$$\mathfrak{S}_{(p,q)} = \{ \sigma \in \mathfrak{S}_{p+q} \mid \sigma(1) < \dots < \sigma(p); \sigma(p+1) < \dots < \sigma(p+q) \}$$

(c'est l'opération élémentaire que fait quelqu'un qui bat un paquet de cartes, le divisant en deux paquets qu'il imbrique ensuite l'un dans l'autre).

Le **produit extérieur** d'une p -forme ω par une q -forme ω' est défini par la formule :

$$(\omega \wedge \omega')(v_1, \dots, v_{p+q}) = \sum_{\sigma \in \mathfrak{S}_{(p,q)}} (-1)^{\text{sgn } \sigma} \omega(v_{\sigma(1)}, \dots, v_{\sigma(p)}) \omega'(v_{\sigma(p+1)}, \dots, v_{\sigma(p+q)})$$

où $\text{sgn } \sigma$ désigne la signature de la permutation σ .

Propriétés du produit extérieur

(Cf. par exemple Godement, *Cours d'algèbre*, Hermann, 1963, exercice 12, § 23.)

- Bilinéarité : (évidente).
- Associativité : $\alpha \wedge (\beta \wedge \gamma) = (\alpha \wedge \beta) \wedge \gamma$.
- Antisymétrie : $\beta \wedge \alpha = (-1)^{pq} \alpha \wedge \beta$ où $p = \deg \alpha$, $q = \deg \beta$ (degré de la forme considérée).

0.2. Structure de l'algèbre des formes extérieures

Notons $\Lambda^p E^*$ l'espace des formes extérieures de degré p sur E . La collection d'espaces vectoriels $(\Lambda^p E^*)_p$, munie de l'opération « produit extérieur », constitue ce qu'on appelle l'*algèbre des formes extérieures* sur E . Cette algèbre est *graduée* (par le degré p), le degré d'un produit extérieur étant la somme des degrés.

On convient que $\Lambda^0 E^* = \mathbb{R}$ (le corps des scalaires), de sorte que notre algèbre graduée a un élément neutre, l'élément unité de \mathbb{R} (le « produit extérieur » par une 0-forme étant la multiplication usuelle par le scalaire qu'est cette 0-forme).

Les espaces vectoriels $\Lambda^p E^*$ sont nuls pour $p > n = \dim E$. Pour $0 \leq p \leq n$ ils sont de dimension finie $\binom{n}{p}$, et l'on peut en construire des bases de la façon suivante : soit u_1, \dots, u_n une base de E^* ; alors pour tout $p = 1, 2, \dots, n$ la famille des p -formes :

$$(u_{i_1} \wedge \dots \wedge u_{i_p})_{i_1 < \dots < i_p}$$

est une base de $\Lambda^p E^*$.

En particulier, l'espace vectoriel $\Lambda^n E^*$ des formes de degré maximum est à une dimension, engendré par $u_1 \wedge u_2 \wedge \dots \wedge u_n$.

0.3. Formes volumes

On vient de voir que tout élément de $\Lambda^n E^*$ est multiple de $u_1 \wedge \dots \wedge u_n$, où $u = (u_1, \dots, u_n)$ est un système de coordonnées linéaires sur E (base de E^*).

Géométriquement $u_1 \wedge \dots \wedge u_n$ est la forme qui à un n -uplet de vecteurs v_1, \dots, v_n de E associe le volume du parallélépipède orienté de \mathbb{R}^n , image par l'isomorphisme $u : E \rightarrow \mathbb{R}^n$ du parallélépipède construit sur v_1, \dots, v_n .

Pour cette raison, un élément non nul de $\wedge^n E^*$ est appelé *forme volume* sur E . En particulier, pour $E = \mathbb{R}^n$ et si u_1, \dots, u_n désigne le système de coordonnées canoniques sur \mathbb{R}^n , $u_1 \wedge \dots \wedge u_n$ est la *forme volume canonique* de \mathbb{R}^n . Plus généralement, on a la proposition suivante.

PROPOSITION

Tout espace vectoriel euclidien orienté E admet une forme volume canonique, caractérisée par la propriété que $\omega(v_1, \dots, v_n) = 1$ pour toute base orthonormée de E orientée positivement.

Si (u_1, \dots, u_n) est le système de coordonnées linéaires de E défini par une telle base, la forme ω peut être définie par $\omega = u_1 \wedge \dots \wedge u_n$. Le fait qu'elle ne dépende pas du choix de la base vient de ce que les matrices de changement de base ont un déterminant égal à 1 (elles appartiennent au groupe spécial orthogonal $SO(n)$).

EXERCICE Démontrer que la forme volume canonique d'un espace vectoriel euclidien orienté E est donnée par la formule :

$$\omega(v_1, \dots, v_n) = \pm \sqrt{\det(\langle v_i, v_j \rangle)} \quad (*)$$

où $(\langle v_i, v_j \rangle)$ désigne la matrice constituée par les produits scalaires des vecteurs v_1, \dots, v_n , et le signe devant la racine est + ou - selon que (v_1, \dots, v_n) est une base de E orientée positivement ou négativement (si les vecteurs ne sont pas linéairement indépendants, le déterminant est nul).

Indication Le choix d'une base orthonormée de E permet de se ramener à démontrer la formule pour la forme volume canonique de \mathbb{R}^n , ce qui revient à démontrer le lemme suivant.

Lemme Pour tout n -uplet (v_1, \dots, v_n) de vecteurs de \mathbb{R}^n ,

$$(\det(v_{ij}))^2 = \det(\langle v_i, v_j \rangle).$$

L'intérêt de la formule (*) est de donner une expression *intrinsèque* de la forme volume, évitant d'avoir à choisir une base de E .

1. L'algèbre graduée des formes différentielles sur une variété

1.0. Formes différentielles

Soit M une variété différentielle. En appliquant la construction du paragraphe 0 à l'espace tangent $E = T_a M$, on obtient l'algèbre des formes extérieures sur $T_a M$. De même que $T_a M$ peut être considéré comme la fibre d'un fibré vectoriel sur M , l'espace vectoriel des p -formes sur $T_a M$ peut être considéré comme la fibre d'un fibré vectoriel sur M que nous appellerons le *fibré des p -formes* sur M (pour $p = 1$ c'est le fibré cotangent T^*M). Les sections (différentiables) de ce fibré sont les **formes différentielles de degré p** .

L'ensemble des formes différentielles de degré p sur un ouvert U de M [resp. l'ensemble des *germes en a* de formes différentielles de degré p sur M] sera noté $\Omega^p(U)$ [resp. $\Omega_{M,a}^p$].

Si $x = (x_1, \dots, x_n)$ est une carte locale de M définie sur U [resp. le germe en a d'une telle carte], une forme différentielle de degré p s'écrit d'une façon et d'une seule :

$$\omega = \sum_{\substack{i=(i_1, \dots, i_p) \\ i_1 < \dots < i_p}} \varphi_i dx_{i_1} \wedge \dots \wedge dx_{i_p}, \tag{**}$$

où $\varphi_i \in C^\infty(U)$ [resp. $C^\infty_{M,a}$].

N.B. : Les énoncés très généraux qui suivent s'appliqueront aussi bien aux germes de fonctions ou de formes différentielles qu'aux fonctions ou formes différentielles définies sur un ouvert U de M (non nécessairement muni d'une carte : ce pourra par exemple être M tout entier). De plus, ils seront, en un sens évident, compatibles avec l'opération de restriction à un ouvert plus petit.

Pour cette raison, nous emploierons des notations telles que C^∞_M, Ω^p_M , qu'on pourra lire à volonté comme signifiant $C^\infty_{M,a}, \Omega^p_{M,a}$ ou bien $C^\infty(U), \Omega^p(U)$ pour U ouvert de M .

1.1. Produit extérieur

Le produit extérieur défini en 0.1 passe de façon évidente aux formes différentielles, définissant pour tous p, q une application bilinéaire :

$$\begin{aligned} \Omega^p_M \times \Omega^q_M &\xrightarrow{\wedge} \Omega^{p+q}_M \\ \omega, \omega' &\mapsto \omega \wedge \omega' \end{aligned}$$

vérifiant les mêmes propriétés d'associativité et d'« antisymétrie » qu'en 0.1. Munie de cette loi de produit, la collection des $(\Omega^p_M)_{p=0,1,\dots,n}$, que nous noterons pour abrégé Ω^*_M , forme une *algèbre graduée*. En particulier l'application :

$$\Omega^0_M \times \Omega^p_M \xrightarrow{\wedge} \Omega^p_M$$

est la loi évidente de multiplication d'une forme différentielle par une fonction, qui fait de Ω^p_M un module sur l'anneau $\Omega^0_M = C^\infty_M$ (de même qu'une forme extérieure de degré zéro sur un espace vectoriel est un scalaire, une forme différentielle de degré zéro sur M est une fonction différentiable).

1.2. Différentiation extérieure

PROPOSITION

Il existe sur Ω^*_m un unique endomorphisme \mathbb{R} -linéaire gradué $d : \Omega^*_M \rightarrow \Omega^*_M$, de degré 1 (c'est-à-dire envoyant Ω^p_M dans Ω^{p+1}_M) :

$$C^\infty_M = \Omega^0_M \xrightarrow{d} \Omega^1_M \xrightarrow{d} \Omega^2_M \xrightarrow{d} \dots \xrightarrow{d} \Omega^n_M \xrightarrow{d} 0$$

doté des propriétés suivantes :

- i) $d \circ d = 0$;
- ii) $d|_{\Omega^0_M}$ est la différentiation usuelle des fonctions ;
- iii) $d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^{\deg \alpha} \alpha \wedge d\beta$.

Cet endomorphisme est appelé **différentiation extérieure** de l'algèbre Ω^*_M .

Preuve

Nous nous contenterons d'esquisser la démonstration, insistant surtout sur ce qu'il faut en retenir pour les usages pratiques : une expression explicite de la différentielle extérieure en coordonnées locales.

Toute forme différentielle ω peut localement s'écrire sous la forme 1.0 (**). Appliquant à cette formule les règles i), ii) et iii) on trouve par un calcul sans difficultés :

$$d\omega = \sum_{\substack{i=(i_1, \dots, i_p) \\ i_1 < \dots < i_p}} d\varphi_i \wedge dx_{i_1} \wedge \dots \wedge dx_{i_p},$$

c'est-à-dire plus explicitement :

$$d\omega = \sum_{\substack{i=(i_1, \dots, i_p) \\ i_1 < \dots < i_p}} \sum_j \partial_{x_j} \varphi_i dx_j \wedge dx_{i_1} \wedge \dots \wedge dx_{i_p}.$$

Cela montre que s'il existe un endomorphisme d vérifiant i), ii) et iii), celui-ci est *unique* et donné *localement* par les expressions ci-dessus. Reste à vérifier que l'endomorphisme d ainsi défini dans les ouverts de cartes vérifie effectivement i), ii) et iii), ce qui n'est pas difficile. Par argument d'unicité cet endomorphisme ne dépend pas du choix de la carte $x = (x_1, \dots, x_n)$, et s'étend donc aux ouverts quelconques de M . \square

1.3. Images réciproques**PROPOSITION**

À tout morphisme $f : M \rightarrow N$ de variétés différentielles est associé « de façon fonctorielle » un morphisme d'algèbres différentielles graduées :

$$f^* : \Omega_N^* \rightarrow \Omega_M^*$$

caractérisé par le fait qu'en degré 0 il envoie la fonction $\varphi \in C_N^\infty$ sur $f^*\varphi := \varphi \circ f \in C_M^\infty$.

« Morphisme d'algèbres différentielles graduées » signifie ceci :

- i) f^* est une application linéaire de Ω_N^p dans Ω_M^p ;
- ii) $f^*(\omega \wedge \omega') = f^*\omega \wedge f^*\omega'$;
- iii) $f^*(d\omega) = d(f^*\omega)$.

Dire que la correspondance $f \mapsto f^*$ est « fonctorielle » veut dire que :

$$\begin{aligned} (\mathbb{1}_M)^* &= \mathbb{1}_{\Omega_M^*} \\ (g \circ f)^* &= f^* \circ g^* \end{aligned}$$

il s'agit d'un *foncteur contravariant* de la catégorie des variétés différentielles dans la catégorie des algèbres différentielles graduées.

Expression de l'image réciproque en coordonnées locales

Définissons f^* en degré zéro par :

$$f^* = C_N^\infty \rightarrow C_M^\infty \\ \psi \mapsto \psi \circ f$$

Alors, par la propriété iii), l'action de f^* est connue sur les différentielles de fonctions. On en déduit à l'aide de ii) que si $\omega \in \Omega_N^p$ est donnée en coordonnées locales par

$$\omega = \sum_j \psi_j dy_{j_1} \wedge \cdots \wedge dy_{j_p}$$

son image réciproque $f^*\omega$ s'écrit localement :

$$\begin{aligned} f^*\omega &= \sum_j f^*\psi_j d(f^*y_{j_1}) \wedge \cdots \wedge d(f^*y_{j_p}) \\ &= \sum_j (\psi_j \circ f) d(y_{j_1} \circ f) \wedge \cdots \wedge d(y_{j_p} \circ f), \end{aligned}$$

que rien n'empêche de réexprimer dans la base $(dx_{i_1} \wedge \cdots \wedge dx_{i_p})_i$ associée à un système de coordonnées locales (x_1, \dots, x_m) de M , en utilisant la formule de différentiation d'une fonction composée.

Cas particulier important

Si $\dim M = \dim N = m$, la forme différentielle de degré maximum :

$$\omega = dy_1 \wedge \cdots \wedge dy_m$$

a pour image réciproque

$$f^*\omega = \det \left(\frac{\partial y_j}{\partial x_i} \right) dx_1 \wedge \cdots \wedge dx_m,$$

où $\left(\frac{\partial y_j}{\partial x_i} \right) = (\partial_{x_i} (y_j \circ f))$ est la matrice jacobienne de l'expression de f dans les coordonnées locales x (source), y (but).

Expression intrinsèque de f^*

Soit $\omega \in \Omega_N^p(V)$. Alors $f^*\omega \in \Omega_M^p(U)$ est la forme différentielle ainsi définie sur $U = f^{-1}(V)$: pour $a \in U$, $b = f(a) \in V$, $\omega(b)$ est une p -forme sur T_bN ; $(f^*\omega)(a)$ est alors la p -forme sur T_aM qui au p -uplet v_1, \dots, v_p d'éléments de T_aM associe le nombre :

$$\omega(b)(T_a f(v_1), \dots, T_a f(v_p)).$$

Cas particulier : restriction d'une forme $\omega \in \Omega_M^p$ à une sous-variété $S \subset M$

On appelle *restriction de ω à la sous-variété S* l'image réciproque de ω par l'inclusion canonique $i : S \rightarrow M$. On la note :

$$\omega|_S = i^* \omega .$$

Explicitement, $(\omega|_S)(a)$ est la p -forme sur $T_a S$ définie de façon évidente en restreignant à $T_a S$ la p -forme $\omega(a)$ donnée sur $T_a M$.

Preuve de la proposition

En définissant f^* par l'expression intrinsèque ci-dessus il est facile de vérifier les propriétés i) et ii). La propriété iii) se vérifie plus facilement sur l'expression en coordonnées locales, dont il est facile par ailleurs d'établir l'équivalence avec l'expression intrinsèque. La vérification des propriétés fonctorielles ne présente aucune difficulté. \square

1.4. Formes différentielles fermées et formes différentielles exactes

Une forme différentielle ω est dite *fermée* si $d\omega = 0$. Elle est dite *exacte* si elle peut s'écrire $\omega = d\xi$. De l'identité $d \circ d = 0$ il résulte que toute forme différentielle exacte est fermée. La réciproque est fautive en général, mais vraie localement comme l'affirme le lemme suivant.

Lemme de Poincaré Sur un pavé de \mathbb{R}^n (ouvert produit de n intervalles de \mathbb{R}) toute forme différentielle fermée est exacte.

Nous ne donnerons pas ici la démonstration générale de ce lemme.

EXERCICE Expliciter, en coordonnées locales, ce que signifie pour une forme différentielle de degré 1 le fait d'être fermée [resp. exacte]. Reconnaître ainsi dans le lemme de Poincaré en degré 1 l'exercice 2 du chapitre 3.

Exemple de forme différentielle fermée qui n'est pas exacte : la forme « $d\theta$ » sur le cercle S^1 , ou sur $\mathbb{R}^2 \setminus \{0\}$ (cf. étude 3, exemple 3 et étude 4^{bis}, paragraphe 1). On prendra garde au fait que la fonction θ n'est pas bien définie globalement sur le cercle mais seulement sur le revêtement universel de celui-ci, ce qui explique que « $d\theta$ » – malgré la notation – ne soit pas exacte sur le cercle.

Nous aborderons au paragraphe 4, sous le nom de *cohomologie*, l'étude générale des obstructions à ce qu'une forme différentielle fermée soit exacte.

2. Formes volumes

2.1. Définition et principaux exemples

On appelle *forme volume* sur une variété différentielle M une forme de degré maximal, non nulle en tout point de M .

Exemple \mathbb{R}^n est muni d'une forme volume canonique, la forme $dx_1 \wedge \cdots \wedge dx_n$ (où x_1, \dots, x_n sont les fonctions coordonnées canoniques).

Remarque La donnée d'une forme volume ω sur M détermine une orientation de M .

En effet, pour toute base (v_1, \dots, v_m) de T_aM , le nombre $\omega(v_1, \dots, v_m)$ est non nul, et l'on peut convenir que la base est orientée positivement si ce nombre est positif.

Inversement, la proposition du paragraphe 0.3 a pour conséquence immédiate la proposition suivante.

PROPOSITION Toute variété riemannienne orientée est munie d'une forme volume canonique.

Rappelons (cf. chap. 5, commentaire 4.3) que l'on appelle « variété riemannienne » une variété différentielle dont les espaces tangents T_aM sont munis d'une structure euclidienne « dépendant différenciablement de a ».

DÉFINITION Plus précisément, on appelle **variété riemannienne** le couple $(M, \langle \cdot, \cdot \rangle)$ constitué par une variété différentielle M et une fonction numérique $\langle \cdot, \cdot \rangle \in C^\infty(TM \times TM)$ dont la restriction $\langle \cdot, \cdot \rangle_a$ à chaque fibre $T_aM \times T_aM$ soit une forme bilinéaire définie positive.

Expression, en coordonnées locales, de la forme volume canonique d'une variété riemannienne orientée

Donnons-nous, sur un ouvert de la variété riemannienne M , un système de coordonnées locales (t_1, \dots, t_m) orienté positivement, c'est-à-dire tel qu'en tout point a les vecteurs tangents $\partial_{t_1}, \dots, \partial_{t_m}$ forment une base orientée positivement de T_aM .

Alors, la forme volume canonique de M est donnée dans cet ouvert par :

$$\omega = \sqrt{\det \langle \partial_{t_i}, \partial_{t_j} \rangle} dt_1 \wedge \dots \wedge dt_m \quad (*)$$

(c'est la formule (*) de l'exercice 0.3).

Exemple : courbe lisse de \mathbb{R}^n , donnée par un paramétrage $t \mapsto (x_1(t), \dots, x_n(t))$. Prenant t comme coordonnée locale sur cette courbe, on trouve par la formule (*) :

$$\omega = \sqrt{x_1'(t)^2 + \dots + x_n'(t)^2} dt.$$

La « forme volume canonique » s'interprète dans ce cas comme « l'élément de longueur infinitésimal » le long de la courbe considérée.

2.2. Calcul de la forme volume canonique d'une sous-variété orientée

$M \subset \mathbb{R}^n$ donnée par un système d'équations

(globales ou locales) $s_1 = 0, \dots, s_p = 0$

Pour une sous-variété donnée de cette façon il est plus commode de raisonner sur le fibré normal que sur le fibré tangent. L'espace \mathbb{R}^n étant canoniquement orienté, la donnée d'une orientation de M équivaut, d'après la remarque du chapitre 7 paragraphe 3.2, à la donnée d'une orientation du fibré normal, et l'on peut toujours supposer que (ds_1, \dots, ds_p) est un système de coordonnées orienté positivement des espaces normaux.

Or la structure euclidienne de \mathbb{R}^n permet d'identifier l'espace normal $T_a \mathbb{R}^n / T_a M$ à $T_a^\perp M$, supplément orthogonal de $T_a M$ dans $T_a \mathbb{R}^n$. L'« espace orthogonal » $T_a^\perp M$ est ainsi un espace vectoriel euclidien orienté.

Lemme Pour toute base orthonormée (v_1, \dots, v_p) de $T_a^\perp M$, orienté positivement, on a :

$$(ds_1 \wedge \dots \wedge ds_p)(v_1, \dots, v_p) = \sqrt{\det(\text{grad } s_i, \text{grad } s_j)}$$

Preuve

En posant $v_{i,j} = ds_i(v_j)$ ce lemme est une conséquence immédiate du lemme du paragraphe 0.3. \square

Corollaire La forme volume canonique sur $T_a^\perp M$ est donnée pour tout $a \in M$ par la formule :

$$\nu = \frac{ds_1 \wedge \dots \wedge ds_p}{\sqrt{\det(\text{grad } s_i, \text{grad } s_j)}}$$

Plus précisément, cette formule définit sur \mathbb{R}^n une forme différentielle ν dotée des propriétés suivantes :

- i) appliquée à un n -uplet de vecteurs (v_1, \dots, v_n) de $T_a \mathbb{R}^n$ ($a \in M$), ν s'annule dès que l'un des v_i appartient à $T_a M$, et peut donc être considérée comme une p -forme sur $T_a \mathbb{R}^n / T_a M$;
- ii) appliquée à un p -uplet de vecteurs de $T_a^\perp M$, ν coïncide avec la forme volume canonique sur $T_a^\perp M$.

De ces propriétés de la forme différentielle ν on déduit immédiatement la proposition suivante.

PROPOSITION

Pour une forme différentielle ω de degré $m = n - p$ sur \mathbb{R}^n , les propriétés suivantes sont équivalentes :

- i) $\nu \wedge \omega = dx_1 \wedge \dots \wedge dx_n$;
- ii) $\omega|_M$ est la forme volume canonique sur M .

Preuve

Il suffit de faire agir $\nu \wedge \omega$ sur un n -uplet $(v_1, \dots, v_p, v_{p+1}, \dots, v_n)$, où (v_1, \dots, v_p) [resp. (v_{p+1}, \dots, v_n)] est une base orthonormée, orientée positivement, de $T_a^\perp M$ [resp. $T_a M$]. \square

Exemple : forme volume canonique d'une hypersurface orientée de \mathbb{R}^n Désignant par s une équation de l'hypersurface M , choisie de façon que l'orientation de M corresponde au « sens de traversée » $ds > 0$ (cf. chap. 7, § 3.0), on a :

$$\nu = \frac{ds}{\|\text{grad } s\|} = \nu_1 dx_1 + \dots + \nu_n dx_n,$$

où les $\nu_i = \partial_i s / \|\text{grad } s\|$ sont les composantes du vecteur unitaire orthogonal à M en a , dirigé dans le sens de traversée défini par l'orientation de M .

Alors la forme différentielle :

$$\omega = \sum_{j=1}^n (-1)^{j-1} v_j \, dx_1 \wedge \cdots \wedge \widehat{dx_j} \wedge \cdots \wedge dx_n$$

vérifie évidemment la propriété i) de la proposition, et définit donc (en restriction à M) la forme volume canonique sur M .

3. Intégration des formes différentielles

Le but de ce paragraphe est d'énoncer (sans démonstration) le *théorème de Stokes*, qui permet de ramener l'intégrale d'une forme différentielle exacte de degré p à une intégrale en dimension $p - 1$. On trouve dans la littérature de nombreux énoncés du théorème de Stokes, se traduisant tous par la même formule mais différant par la façon dont sont définis les objets géométriques sur lesquels on intègre. J'ai choisi ici de définir l'intégration sur les *chaînes singulières* d'une variété différentielle M . Pour définir une « chaîne » de dimension p on commencera par en définir les « maillons » élémentaires, qu'on appellera *cellules* de dimension p .

3.1. Intégration sur une cellule singulière

Considérons, dans un espace affine euclidien orienté E de dimension p , un compact Δ d'intérieur non vide, et une application différentiable :

$$\sigma : \Delta \rightarrow M ;$$

« différentiable » signifie ici que σ est la restriction à Δ d'une application différentiable sur un ouvert de E contenant Δ .

Pour toute forme différentielle ω de degré p sur M , $\sigma^* \omega$ est une forme différentielle de degré p sur un ouvert de E contenant Δ ; étant de degré maximum ($p = \dim E$), elle peut donc s'écrire :

$$\sigma^* \omega = f \cdot \omega_E,$$

où ω_E est la forme volume canonique sur E , et $f \in C^\infty(\Delta)$.

DÉFINITION On appelle *intégrale de ω sur σ* le nombre :

$$\int_{\sigma} \omega := \int_{\Delta} f \, \mu_E,$$

où μ_E désigne la mesure de Lebesgue sur l'espace euclidien E .

Remarque Si $E = \mathbb{R}^p$ les équations précédentes peuvent s'écrire respectivement :

$$\sigma^* \omega = f \, dx_1 \wedge \cdots \wedge dx_p$$

et

$$\int_{\sigma} \omega = \int_{\Delta} f \, dx_1 \cdots dx_p.$$

EXERCICE : Invariance par « changement de paramétrage ». Montrer que si Δ et Δ' sont deux compacts d'espaces affines euclidiens orientés, se déduisant l'un de l'autre par un difféomorphisme $\varphi : \Delta' \rightarrow \Delta$ respectant l'orientation, on a :

$$\int_{\sigma} \omega = \int_{\sigma'} \omega \quad (\text{où } \sigma' = \sigma \circ \varphi)$$

(utiliser la formule de changement de variable dans les intégrales, et le caractère fonctoriel de l'image réciproque des formes différentielles).

Nous allons maintenant restreindre nos hypothèses sur le compact Δ , de façon à pouvoir définir le bord de σ et énoncer le théorème de Stokes.

DÉFINITION On appelle **cellule singulière** de dimension p sur M la donnée :

- i) d'un espace affine euclidien orienté E_{Δ} de dimension p , et d'un polyèdre convexe $\Delta \subset E_{\Delta}$ de même dimension ;
- ii) d'une application différentiable $\sigma : \Delta \rightarrow M$.

Nous résumerons toutes ces données par la seule lettre σ , résumant la première donnée en disant que l'espace source Δ de σ est un « **polyèdre convexe orienté de dimension p** ».

Exemple $p = 1$, $\Delta = [0, 1] \subset E_{\Delta} = \mathbb{R}$ La cellule singulière σ est alors un « chemin » différentiable de M .

Remarque Cette définition n'impose pas à σ d'envoyer Δ difféomorphiquement sur son image. C'est pour cette raison que σ est appelée cellule *singulière* (son image dans M peut ne pas ressembler du tout à ce qu'on aurait envie d'appeler une cellule). En particulier, σ peut fort bien n'être nulle part une immersion, c'est-à-dire avoir une application tangente de rang partout inférieur à p . Dans ce cas l'image réciproque $\sigma^* \omega$ de toute forme différentielle $\omega \in \Omega_M^p$ est identiquement nulle, de sorte que $\int_{\sigma} \omega = 0$.

Nous dirons qu'une telle cellule singulière σ est *dégénérée*.

3.2. Bord d'une cellule singulière. Théorème de Stokes

Rappels sur les polyèdres

■ **Polyèdre convexe** : compact convexe Δ d'un espace affine E , n'ayant qu'un nombre fini de « points extrémaux » (appelés *sommets* du polyèdre). (Rappelons qu'un « point extrémal » d'un fermé convexe Δ est un point $a \in \Delta$ qui n'appartient à aucun segment ouvert inclus dans Δ .)

Un polyèdre convexe Δ est l'enveloppe convexe de l'ensemble de ses sommets. On appelle *dimension* de Δ la dimension du plus petit sous-espace affine E_{Δ} qui le contient (qui n'est autre, par raison de convexité, que le sous-espace affine engendré par l'ensemble de ses sommets).

Un polyèdre Δ de dimension p a au moins $p + 1$ sommets. S'il en a exactement $p + 1$ on l'appelle **simplexe de dimension p** : un simplexe de dimension p est donc l'enveloppe convexe de $p + 1$ points affinement indépendants.

■ **Orienté** d'un polyèdre convexe Δ : c'est la donnée d'une orientation de E_{Δ} .

■ **Face d'un polyèdre convexe** : partie de Δ de la forme $\Delta \cap F$, où F est un « hyperplan frontière » de Δ dans E_{Δ} .

Par « hyperplan frontière » d'un polyèdre convexe Δ de dimension p , nous entendons un $(p - 1)$ -plan $F \subset E_\Delta$ tel que :

- i) Δ soit situé « d'un seul côté » de F (c'est-à-dire ne rencontre qu'une seule des deux composantes convexes de E_Δ/F) ;
- ii) le polyèdre $\Delta \cap F$ soit de dimension $p - 1$.

Un polyèdre convexe n'a qu'un nombre fini de faces.

■ **Orientation des faces « par la normale extérieure »** : la donnée d'une orientation de E_Δ détermine une orientation de chacun des hyperplans frontières de Δ par la « convention de la normale extérieure » (on oriente l'espace normal à F dans E_Δ dans le sens de « sortie de Δ », et l'on applique la remarque du paragraphe 3.0 du chapitre 7).

DÉFINITION Soit $\sigma : \Delta \rightarrow M$ une cellule singulière de dimension p . On appelle **face** de la cellule singulière σ une cellule singulière de dimension $p - 1$, de la forme :

$$\sigma_F = \sigma|_{\Delta_F} : \Delta_F \rightarrow M,$$

où Δ_F est une face du polyèdre Δ , munie de l'orientation déduite de celle de Δ par la convention de la normale extérieure.

Théorème de Stokes L'intégrale, sur une cellule singulière σ , d'une forme différentielle exacte $d\chi$ est donnée par la formule :

$$\int_{\sigma} d\chi = \int_{\partial\sigma} \chi$$

où $\int_{\partial\sigma}$ est une abréviation pour $\sum_F \int_{\sigma_F}$. **somme des intégrales sur les faces σ_F de la cellule singulière σ .**

Cas particulier de la dimension 1

L'intégrale, sur un chemin σ de M , d'une forme différentielle exacte df (f = fonction sur M) est donnée par :

$$\int_{\sigma} df = f(b) - f(a),$$

où a et b désignent respectivement l'origine et l'extrémité du chemin σ .

Cet énoncé ne fait que traduire dans le langage des formes différentielles les propriétés bien connues des « primitives » de fonctions d'une variable. Pour cette raison on appelle **primitive** d'une forme différentielle ω de degré 1 une fonction f telle que $\omega = df$.

On remarquera que cet énoncé ne peut être considéré comme un cas particulier de l'énoncé général qu'à condition de convenir qu'une cellule orientée à une dimension a pour « faces » son origine et son extrémité affectées respectivement des signes $-$ et $+$ (les signes remplacent ici l'orientation, que nous n'avons pas définie en dimension 0).

3.3. Introduction à la notion de « chaîne singulière »

Convenons provisoirement d'appeler *chaîne singulière de dimension p sur M* toute combinaison linéaire formelle, finie, de cellules singulières de dimension p :

$$c = \sum_i n_i \sigma_i \quad (n_i \in \mathbb{Z}) \quad (1)$$

Ayant défini en 3.1 l'intégration sur une cellule singulière nous en déduisons, par linéarité, la notion d'intégration sur une chaîne singulière (1) :

$$\int_c := \sum_i n_i \int_{\sigma_i}. \quad (2)$$

En particulier, le second membre de la formule de Stokes n'est autre que l'intégrale sur le « bord » $\partial\sigma$ de la cellule singulière σ , défini par la somme formelle des faces de σ , si $\dim \sigma \geq 2$:

$$\partial\sigma = \sum_F \sigma_F, \quad (3)$$

et par

$$\partial\sigma = b - a \quad \text{si} \quad \dim \sigma = 1.$$

Ainsi définie sur les cellules, la notion de **bord** s'étend par linéarité aux chaînes (1) :

$$\partial c = \sum_i n_i \partial \sigma_i, \quad (4)$$

permettant d'énoncer le **théorème de Stokes pour les chaînes régulières** :

$$\int_c d\chi = \int_{\partial c} \chi. \quad (5)$$

EXERCICE Démontrer de deux façons différentes que l'intégrale sur un chemin λ d'une forme différentielle *fermée* ω de degré 1 ne dépend que de la classe d'homotopie (à extrémités fixes) de λ .

• *Première méthode* : appliquer le théorème de Stokes en dimension 2, en intégrant sur la cellule singulière (de source $\Delta = [0, 1] \times [0, 1]$) que définit l'homotopie entre les deux chemins.

• *Deuxième méthode* : se ramener par le lemme de Poincaré du paragraphe 1.4 au cas d'une forme différentielle exacte, ce qui permet d'appliquer le théorème de Stokes de dimension 1.

3.4. Le complexe des chaînes singulières sur une variété, d'après De Rham

Parmi les « chaînes singulières » introduites en 3.3, beaucoup sont manifestement inutiles du point de vue de l'intégration, par exemple celles dont toutes les cellules σ_i sont « dégénérées » (cf. § 3.1), ou bien des chaînes comme $c = \sigma - \sigma'$ où σ et σ' sont deux cellules singulières déduites l'une de l'autre par un « changement de paramétrage » (cf. exercice du § 3.1), etc.

Plus généralement, convenons d'appeler **négligeable** toute combinaison linéaire (1) sur laquelle l'intégrale (2) de toute forme différentielle ω soit nulle. Dans le groupe abélien que constitue l'ensemble de toutes les expressions (1), celles qui sont négligeables forment évidemment un sous-groupe.

DÉFINITION Le groupe quotient du groupe des expressions de la forme (1) par le sous-groupe des expressions « négligeables » est appelé **groupe des chaînes singulières de dimension p sur M** , et noté $C_p(M)$.

Par construction même, la définition (2) de l'intégrale passe au quotient par cette relation d'équivalence, et il résulte immédiatement de la formule de Stokes (5) que l'homomorphisme bord passe lui aussi au quotient (si c est négligeable, il en est de même de ∂c). De plus, compte tenu de l'identité $d(d\omega) = 0$ (cf. § 1.2), la formule de Stokes appliquée deux fois montre que l'opérateur ∂ appliqué deux fois à une expression (1) quelconque a un résultat négligeable. Autrement dit, la suite d'homomorphismes de groupes abéliens :

$$0 \xrightarrow{\partial} C_n(M) \xrightarrow{\partial} C_{n-1}(M) \xrightarrow{\partial} \cdots \xrightarrow{\partial} C_1(M) \xrightarrow{\partial} C_0(M) \xrightarrow{\partial} 0$$

vérifie la propriété :

$$\partial \circ \partial = 0 ;$$

on résume cette propriété en disant que cette suite d'homomorphismes est un **complexe de groupes abéliens**.

4. Introduction à l'homologie et à la cohomologie

4.1. Homologie et cohomologie : dualité de De Rham

En vis-à-vis du « complexe des chaînes singulières » $C_*(M)$ qui vient d'être défini en 3.4, considérons le complexe $\Omega^*(M)$ des formes différentielles sur M , qui a été défini en 1.2. C'est un complexe de \mathbb{R} -espaces vectoriels (les morphismes d sont \mathbb{R} -linéaires, et l'on a $d \circ d = 0$, et il en résulte que l'image de $d_{p-1} : \Omega^{p-1}(M) \rightarrow \Omega^p(M)$ (espace des **formes différentielles exactes de degré p**) est un sous-espace vectoriel du noyau de $d_p : \Omega^p(M) \rightarrow \Omega^{p+1}(M)$ (espace des **formes différentielles fermées de degré p**). Le \mathbb{R} -espace vectoriel quotient $\text{Ker } d_p / \text{Im } d_{p-1}$ est appelé **p -ième espace de cohomologie de De Rham de M** , et noté $H_{DR}^p(M)$. Ses éléments sont appelés **classes de cohomologie de degré p** (de M) ; se donner une classe de cohomologie c'est donc se donner globalement sur M une forme différentielle fermée, modulo la relation d'équivalence suivante dite **relation de cohomologie** : deux formes différentielles fermées sont « cohomologues » si leur différence est une différentielle exacte.

Procédant de façon similaire pour le complexe de groupes abéliens $C_*(M)$, on définira dans chacun des groupes $C_p(M)$ les deux sous-groupes suivants :

- le **groupe des bords** $B_p(M) = \text{Im } \{\partial_{p-1} : C_{p-1}(M) \rightarrow C_p(M)\}$;
- le **groupe des cycles** $Z_p(M) = \text{Ker } \{\partial_p : C_p(M) \rightarrow C_{p+1}(M)\}$;

$B_p(M)$ est un sous-groupe de $Z_p(M)$, et le groupe quotient est appelé **p -ième groupe d'homologie de M** . On le note :

$$H_p(M) = Z_p(M)/B_p(M)$$

Ses éléments sont appelés **classes d'homologie de dimension p** de M ; se donner une classe d'homologie c'est donc se donner un cycle de M , modulo la relation d'équivalence suivante dite **relation d'homologie** : deux cycles sont « homologues » si leur différence est un bord.

La formule de Stokes nous dit que les applications ∂_p et d_p sont « transposées » l'une de l'autre par la forme bilinéaire « intégration ». Elle implique notamment que :

- i) l'intégrale sur un cycle d'une forme différentielle exacte est nulle ;
- ii) l'intégrale sur un bord d'une forme différentielle fermée est nulle.

Ces deux résultats se résument en un seul : *l'intégrale sur un cycle γ d'une forme différentielle fermée ω ne dépend que de la classe d'homologie de γ et de la classe de cohomologie de ω .*

Remarque Il y a une légère disparité entre nos façons de traiter l'homologie et la cohomologie : alors que les espaces de cohomologie sont ici des \mathbb{R} -espaces vectoriels, les groupes d'homologie ne sont que des groupes abéliens (\mathbb{Z} -modules). En fait, dans la mesure où l'on ne s'intéresse à l'homologie que pour intégrer des formes différentielles, il serait naturel de travailler avec les « chaînes singulières à coefficients réels » (en modifiant les définitions du paragraphe 3.4 de façon à permettre aux coefficients n_i de la formule (1) du paragraphe 3.3 d'être des nombres réels arbitraires). On obtient ainsi les *espaces d'homologie à coefficients réels*, que nous noterons $H_p(M, \mathbb{R})$.

Théorème de dualité de De Rham¹ *La forme bilinéaire « intégration » fait de $H_{DR}^p(M)$ l'espace vectoriel dual de $H_p(M, \mathbb{R})$.*

Deux conséquences du théorème de dualité de De Rham

- i) La propriété d'avoir une intégrale nulle sur tout cycle caractérise, parmi les formes différentielles fermées, celles qui sont exactes.
- ii) L'annulation de l'intégrale de toute forme différentielle fermée caractérise, parmi les cycles, ceux qui sont des bords.

L'affirmation i) exprime l'injectivité de l'application canonique :

$$\alpha : H_{DR}^p(M) \rightarrow H_p(M, \mathbb{R})^* ,$$

tandis que l'affirmation ii) exprime l'injectivité de l'application canonique :

$$\tilde{\alpha} : H_p(M, \mathbb{R}) \rightarrow H_{DR}^p(M)^* .$$

Dans le cas où l'espace vectoriel $H_p(M, \mathbb{R})$ est de dimension finie (donc égal à son bidual), $\tilde{\alpha}$ n'est autre que la transposée de α , et les affirmations i) et ii) épuisent le contenu du théorème de dualité de De Rham (dire que la transposée d'une application linéaire est injective, c'est dire que cette application est surjective).

¹ Ce théorème est vrai sous une hypothèse très générale sur M . qu'on exprime souvent en disant que « M est paracompacte ». Pour un espace topologique localement compact (comme le sont toutes les variétés), être « paracompact » équivaut à être « dénombrable à l'infini », c'est-à-dire union d'une famille dénombrable de compacts.

EXERCICE Soit $M = \mathbb{R}^2 \setminus A$, où A est un ensemble éventuellement infini mais *discret* (tous ses points sont isolés). Pour tout $a \in A$ on choisit dans M un lacet γ_a tournant une fois dans le sens positif le long d'un petit cercle de centre a (de rayon assez petit pour n'entourer aucun autre point de A).

En considérant les intégrales sur des formes différentielles fermées de M convenablement choisies, montrer que les classes d'homologie des γ_a ($a \in A$) sont linéairement indépendantes dans $H_1(M, \mathbb{R})$.

En admettant (ce qui est vrai) que ces classes d'homologie engendrent tout $H_1(M, \mathbb{R})$, que nous apprend le théorème de De Rham sur les formes différentielles fermées de M ?

4.2. Autres points de vue sur l'homologie

À la définition 3.4 du complexe des chaînes, qui fait explicitement appel aux formes différentielles (notion de chaîne « négligeable »), on préfère souvent des définitions ne faisant pas intervenir la structure différentielle, et qui s'appliquent à des classes d'espaces beaucoup plus larges que les variétés.

Je vais en esquisser deux, dont chacune nous aidera à éclairer un aspect différent de la notion d'homologie. Dans chacune de ces définitions on pourra prendre les « coefficients » des chaînes dans un groupe abélien de notre choix, disons \mathbb{Z} pour fixer les idées. Les groupes d'homologie ainsi obtenus sont notés $H_p(M, \mathbb{Z})$ (ils sont les mêmes pour les deux définitions). Dans le cas particulier où M est une variété, le groupe noté $H_p(M)$ en 4.1 coïncide avec $H_p(M, \mathbb{Z})$ « libéré de sa torsion », c'est-à-dire avec le groupe quotient de $H_p(M, \mathbb{Z})$ par le sous-groupe formé par les éléments de γ tels que $n\gamma = 0$ pour un $n \in \mathbb{Z}^*$.

À chacune de ces définitions du complexe des chaînes correspond une notion – duale – de complexe des cochaînes, conduisant à des *groupes de cohomologie* duaux des groupes d'homologie.

a) Point de vue des décompositions cellulaires. On décompose l'espace topologique M à étudier en une union disjointe de « cellules » de diverses dimensions : une « cellule de dimension p » est homéomorphe à l'intérieur d'un polyèdre convexe de \mathbb{R}^p , et la frontière d'une cellule de dimension p est une union de cellules de dimensions inférieures à p (ses faces, les faces de ses faces, etc.).

Cette « décomposition cellulaire » de M étant fixée, on appelle « chaîne de dimension p » une expression formelle comme l'expression (1) du paragraphe 3.3, mais où les σ_i sont des cellules de dimension p de la subdivision donnée, munies chacune d'une orientation (de plus on fait la convention que $-\sigma_i$ représente la cellule σ_i munie de l'orientation opposée). Les « faces » d'une cellule σ de dimension p sont des cellules de dimension $p-1$, dont chacune est munie de l'orientation déduite de celle de σ par convention « de la normale extérieure ». La somme de ces faces ainsi orientées est par définition le *bord* $\partial\sigma$ de σ .

Ce point de vue a l'avantage d'inscrire explicitement la théorie de l'homologie dans le prolongement des problèmes célèbres de « topologie combinatoire » auquel le nom d'Euler reste attaché : par exemple la notion de « caractéristique d'Euler » d'une surface compacte n'a pas été définie initialement par Euler à l'aide de champs de vecteurs comme nous l'avons fait dans l'intermède, mais à l'aide d'une subdivision cellulaire de la surface (en fait les « surfaces » qu'il étudiait étaient polyédrales) ; ce qu'Euler a découvert à ce sujet, c'est que :

pour toute surface compacte M munie d'une subdivision cellulaire, le nombre $\chi(M)$ défini par

$$\chi(M) = e_2(M) - e_1(M) + e_0(M)$$

(où $e_p(M)$ désigne le nombre de cellules de dimension p de la subdivision) était indépendant du choix de la subdivision.

Activité suggérée Tester cette affirmation pour diverses subdivisions de la sphère, du tore, etc., et constater que la « caractéristique d'Euler » ainsi définie coïncide avec celle définie dans l'intermède.

Un prolongement naturel de cette activité, auquel le lecteur pourra s'essayer quand il saura calculer quelques groupes d'homologie (après lecture des paragraphes 4.3 et 4.4 ci-après), est de tester sur ces exemples cette autre propriété de la caractéristique d'Euler (qu'on prend aujourd'hui comme définition de celle-ci) :

$$\chi(M) = \sum_p (-1)^p b_p(M),$$

où

$$b_p(M) = \text{rang}_{\mathbb{Z}} H_p(M)$$

(« p -ième nombre de Betti » de M).

a') (variante de a) : point de vue des triangulations. C'est le cas particulier de a) où l'on impose à toutes les cellules d'être des « simplexes » (cf. § 3.2 pour la signification de ce mot).

b) Point de vue des « chaînes singulières » sur les espaces topologiques. Le point de vue a) a l'avantage d'être très constructif et explicite (c'est de la « topologie combinatoire »), mais pose des problèmes théoriques délicats : quels sont les espaces qui admettent une subdivision cellulaire ? Comment démontrer que les groupes d'homologie ainsi définis ne dépendent pas du choix de la subdivision ?

Le point de vue moderne des « chaînes singulières » évite ces inconvénients, et présente l'avantage de s'appliquer aux espaces topologiques les plus généraux. Il reprend l'idée des paragraphes 3.1, 3.2 et 3.3 en remplaçant les « cellules singulières » par des **simplexes singuliers** :

$$\sigma : \Delta_p \rightarrow M;$$

Δ_p désigne le **simplexe standard à p dimensions** défini par :

$$\Delta_p = \left\{ x \in \mathbb{R}^p \mid x_i \geq 0, \sum_{i=1}^p x_i \leq 1 \right\};$$

l'application σ est supposée continue, sans autre hypothèse. Le groupe des « chaînes singulières de dimension p » de M est défini comme l'ensemble des combinaisons linéaires finies de simplexes singuliers de dimension p . En particulier, le bord d'un simplexe singulier σ de dimension p est la chaîne singulière de dimension $p-1$ définie par :

$$\partial\sigma = \sum_{i=1}^p (-1)^{i-1} \sigma \circ F_i.$$

où F_i est l'injection de Δ_{p-1} dans Δ_p définie par :

$$F_i(x_1, \dots, x_{p-1}) = (x_1, \dots, x_{i-1}, 0, x_i, \dots, x_{p-1}).$$

Étendu par linéarité aux chaînes, l'homomorphisme « bord » ainsi défini vérifie de façon évidente $\partial \circ \partial = 0$, et le « complexe des chaînes singulières » ainsi défini conduit à une « bonne » théorie de l'homologie sans qu'il soit nécessaire de passer préalablement au quotient par quelque relation d'équivalence que ce soit (d'une façon que j'ai toujours trouvée miraculeuse, tous les cycles que l'on aurait envie de négliger se trouvent automatiquement être des bords !).

4.3. Homologie en dimensions 0 et 1

Homologie en dimension 0 Pour tout espace topologique M , $H_0(M, \mathbb{Z})$ est le groupe abélien libre engendré par l'ensemble des composantes connexes par arcs de M .

C'est très facile à voir dans le point de vue du paragraphe 4.2.b) ; car :

- i) se donner un simplexe singulier de dimension 0 équivaut à se donner un point de M ;
- ii) les simplexes singuliers de dimension 1 sont les chemins de M (applications continues de $\Delta_1 = [0, 1]$ dans M) ;
- iii) le bord d'un simplexe singulier σ de dimension 1 est $\sigma(1) - \sigma(0)$.

Homologie en dimension 1 Pour tout espace topologique M connexe par arcs, $H_1(M, \mathbb{Z})$ peut être identifié au groupe fondamental de M « rendu commutatif », c'est-à-dire au groupe quotient de $\pi_1(M, a)$ par le sous-groupe engendré par les « commutateurs » $\mu \lambda \mu^{-1} \lambda^{-1}$.

Idée D'après la description des chaînes de dimension 1 et de leurs bords qui vient d'être donnée, le groupe des cycles singuliers de dimension 1 de M est engendré par les *lacets* de M , à point-base arbitraire.

L'affirmation ci-dessus revient à dire que deux tels lacets sont *homologues* si et seulement s'ils se déduisent l'un de l'autre par une *homotopie de lacets à point-base variable* (de sorte que, d'après le chap. 7, § 1.1, Fig. 7, tout lacet λ est homologue à tous ses conjugués $\mu \lambda \mu^{-1}$).

Remarque Comparer les deux méthodes de résolution de l'exercice du paragraphe 3.3 était déjà une façon de prendre conscience d'un lien entre la notion d'homologie en dimension 1 et celle d'homotopie.

Exemples

- i) *Homologie du plan moins deux points* Il résulte ainsi du paragraphe 1.1 du chapitre 7 (commentaire accompagnant la figure 6), que $H_1(\mathbb{R}^2 \setminus \{a_1, a_2\}, \mathbb{Z}) \approx \mathbb{Z} \oplus \mathbb{Z}$ (groupe abélien libre à deux générateurs engendré par deux « petits cercles » autour de a_1 et a_2 ; comparer à l'exercice du paragraphe 4.1).
- ii) *Homologie du tore \mathbb{T}^2* . On a $H_1(\mathbb{T}^2) \approx \pi_1(\mathbb{T}^2) \approx \mathbb{Z} \oplus \mathbb{Z}$ (dans cet exemple, le groupe fondamental était déjà commutatif ; le calcul de ce groupe est un exercice facile d'application de 7, § 2.3, en raisonnant sur le revêtement universel).

On montrerait de la même façon que :

$$H_1(\mathbb{T}^n) \approx \pi_1(\mathbb{T}^n) \approx \mathbb{Z}^n$$

4.4. Homologie en dimension maximale, pour les variétés compactes orientées

Une variété M de dimension m n'a d'homologie non triviale qu'en dimension $p = 0, 1, \dots, m$. Si M est compacte et orientée, un théorème de Poincaré affirme l'existence d'un isomorphisme canonique (l'**isomorphisme de Poincaré**) entre

$$H_p(M, \mathbb{Z}) \quad \text{et} \quad H_{m-p}(M, \mathbb{Z})$$

(Poincaré démontrait ce résultat en utilisant le point de vue du paragraphe 4.2 a) ou a'), grâce à la notion de « polyèdre dual » ou de « triangulation duale ».)

En particulier, $H_m(M, \mathbb{Z})$ est canoniquement isomorphe à $H_0(M, \mathbb{Z})$, c'est-à-dire à \mathbb{Z} en supposant la variété M connexe.

Le générateur canonique de $H_m(M, \mathbb{Z})$ est appelé **classe fondamentale de M** . La façon la plus simple de se le représenter est d'utiliser le point de vue du paragraphe 4.2 a) : on munit M d'une subdivision cellulaire, et l'hypothèse d'orientabilité de M permet d'orienter toutes les cellules de dimension maximale « de façon concordante » (pour reprendre une expression déjà utilisée dans l'intermède, paragraphe 3.3) ; la « concordance » des orientations se traduit par le fait que la somme des cellules ainsi orientées est un cycle, et la classe d'homologie de ce cycle est le générateur cherché de $H_m(M, \mathbb{Z})$.

4.5. Intégrale d'une forme différentielle de degré maximal sur une variété compacte orientée

Toute forme différentielle ω de degré maximal m est évidemment fermée, et l'on peut donc en définir l'intégrale sur une classe d'homologie.

Si M est une variété compacte orientée, on définit donc $\int_M \omega$ comme étant l'intégrale de ω sur la classe fondamentale de M .

Remarquons que, par dualité de De Rham, la cohomologie $H_{DR}^m(M)$ est un \mathbb{R} -espace vectoriel à une dimension. Comme générateur de cet espace vectoriel, on peut prendre la classe d'une *forme volume* sur M (§ 2) : il suffit pour s'en convaincre de vérifier que l'intégrale d'une forme volume sur la classe fondamentale de M ne peut pas être nulle, ce qui résulte immédiatement du fait que les intégrales d'une telle forme sur les cellules orientées qui composent la classe fondamentale sont toutes de même signe (le signe de $\omega(v_1, \dots, v_m)$, où (v_1, \dots, v_m) est n'importe quelle base $T_a M$ orientée positivement).

4.6. En guise de conclusion...

La richesse de la notion d'homologie, qui se traduit par un foisonnement de définitions différentes du complexe des chaînes (donnant en fin de compte les mêmes groupes d'homologie) a conduit Eilenberg et Steenrod à proposer une *approche axiomatique de l'homologie*. Selon ce point de vue, qui pose au départ une liste d'axiomes de l'homologie, les multiples constructions du complexe des chaînes ne sont que des artifices techniques pour démontrer l'existence d'une théorie de l'homologie vérifiant les axiomes. Une fois cette existence reconnue on peut oublier les chaînes, car les

axiomes à eux seuls suffisent en théorie à caractériser de façon unique les groupes d'homologie (et en pratique à les calculer effectivement, pour les espaces usuels).

Le lecteur intéressé est invité à se reporter à la littérature, et en premier lieu au livre d'Eilenberg et Steenrod (*Foundations of Algebraic Topology*, Princeton University Press, 1952).

Contentons-nous ici de noter, parmi les axiomes de l'homologie, le fait que H_p est un foncteur de la catégorie des espaces topologiques dans la catégorie des groupes, et que ce foncteur se factorise en fait à travers la catégorie d'homotopie (Chap. 7, § 0). Il en résulte notamment que deux espaces ayant même type d'homotopie ont des groupes d'homologie et de cohomologie isomorphes, ce qui entraîne beaucoup de conséquences spectaculaires dont voici quelques échantillons.

i) La sphère S^n n'est pas contractile (cf. chap. 7, remarque du § 1.3).

En effet, pour $n \geq 1$, $H_n(S^n) = \mathbb{Z}$ (S^n est une variété compacte orientée de dimension n). Comme l'homologie du point est triviale en toutes dimensions autres que 0, il en résulte que S^n ne peut pas avoir le type d'homologie d'un point.

ii) La boule B^{n+1} ne peut se rétracter sur son bord S^n .

En effet, B^{n+1} , ayant le même type d'homotopie qu'un point, a une homologie triviale en toutes dimensions autres que 0. Comme ce n'est pas le cas de S^n , l'inexistence d'une rétraction $r : B^{n+1} \rightarrow S^n$ en découle par le même argument de functorialité que celui utilisé au chapitre 7, paragraphe 1.2 à propos du foncteur π_1 .

Corollaire Toute application continue de la boule fermée dans elle-même a au moins un point fixe.

Preuve

S'il existait une application sans point fixe $f : B^{n+1} \rightarrow B^{n+1}$, on en déduirait une rétraction $r : B^{n+1} \rightarrow S^n$ par la construction symbolisée sur la figure 1. \square

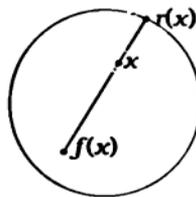


Figure 1.

Signalons pour finir le fait que le lemme de Poincaré du paragraphe 1.4 est vrai non seulement pour les pavés de \mathbb{R}^n mais pour toute variété contractile : en effet le lemme de Poincaré consiste en l'affirmation de la trivialité de la cohomologie en toutes dimensions autres que 0, trivialité qui est assurée pour tout espace contractile.

ÉTUDE 8

Mécanique hamiltonienne et géométrie symplectique

1 Champs de vecteurs hamiltoniens

Toute notre étude se situera sur un ouvert connexe de $\mathbb{R}^n \times \mathbb{R}^n$, dont les coordonnées « canoniques » seront notées $(p, q) = (p_1, \dots, p_n; q_1, \dots, q_n)$.

Un champ de vecteurs v sur cet ouvert est dit **hamiltonien** s'il existe une fonction H telle que :

$$\begin{cases} v_{p_i} &= -\partial_{q_i} H \\ v_{q_i} &= \partial_{p_i} H \end{cases} \quad i = 1, \dots, n. \quad (1)$$

La fonction H (unique à l'addition d'une constante près) est alors appelée **fonction de Hamilton** de v , et l'on dit que v est le **champ hamiltonien** de H (notation abrégée : $v = v_H$).

1.1 Cas $n=1$

La propriété pour v d'être hamiltonien dans un ouvert du plan $\mathbb{R} \times \mathbb{R}$ implique évidemment :

$$\partial_p v_p + \partial_q v_q = 0. \quad (*)$$

Cette condition (*) signifie géométriquement que le flot de v conserve les aires (cf. remarque ci-après). Réciproquement il est facile de vérifier que tout champ de vecteurs défini sur un ouvert U de $\mathbb{R} \times \mathbb{R}$ où il vérifie (*) est *localement hamiltonien*, c'est-à-dire qu'il existe une famille de fonctions (H_α) définies sur des ouverts U_α recouvrant U telles que $v|_{U_\alpha} = v_{H_\alpha}$.

EXERCICE Calculer le champ de vecteurs de $\mathbb{R}^2 \setminus \{0\}$ dont les fonctions de Hamilton locales H_α sont les diverses déterminations de l'« angle » ($H_\alpha = \text{Arctan } p/q$ ou $-\text{Arctan } q/p$ à une constante additive c_α près). Tracer ses trajectoires intégrales et esquisser la façon dont un élément d'aire choisi aussi simplement que possible se déforme par le flot.

Remarque Plus généralement, si v est un champ de vecteurs sur un ouvert de \mathbb{R}^m , la condition :

$$\text{div } v = 0,$$

où $\text{div } v := \partial_1 v_1 + \dots + \partial_m v_m$ (« divergence de v »), équivaut à dire que le jacobien du flot de v est de déterminant égal à 1 pour tout temps. Géométriquement, cela signifie que le flot de v conserve les volumes, résultat souvent utilisé en mécanique quand on étudie les « fluides incompressibles ».

1.2 Cas général

Pour $n > 1$, la propriété pour v d'être (localement) hamiltonien est beaucoup plus forte que la condition $\text{div } v = 0$: elle implique non pas une mais n^2 conditions :

$$\partial_{p_i} v_{p_j} + \partial_{q_i} v_{q_j} = 0 \quad (i, j = 1, \dots, n) \quad (**)$$

(d'où découle en particulier $\text{div } v = 0$ en faisant $i = j$ et en sommant sur i). Analysons ces conditions.

Réécriture des conditions (**): un calcul immédiat montre que :

$$(**) \Rightarrow d(\varpi(v)),$$

où $\varpi(v)$ désigne la forme différentielle de degré 1 :

$$\varpi(v) = \sum_{i=1}^n (v_{p_i} dq_i - v_{q_i} dp_i). \quad (2)$$

Corollaire Les conditions (**) sont nécessaires et suffisantes pour que v soit localement hamiltonien. En effet, d'après le lemme de Poincaré la forme différentielle $\varpi(v)$ est fermée si et seulement si elle est localement exacte, c'est-à-dire s'il existe localement au voisinage de tout point une fonction H telle que :

$$\varpi(v) = dH. \quad (1)'$$

Un calcul immédiat montre que cette dernière égalité n'est qu'une réécriture du fait que $v = v_H$ (équation (1)).

Remarque Il est clair qu'en tout point a de $\mathbb{R}^n \times \mathbb{R}^n$ la correspondance qui au vecteur v_a associe le covecteur $\varpi(v_a)$ est un isomorphisme entre l'espace tangent et l'espace cotangent à $\mathbb{R}^n \times \mathbb{R}^n$ en a . Cet isomorphisme (dont la dépendance en a est triviale donc différentiable !) permet d'identifier les champs de vecteurs aux formes différentielles de degré 1, et tout ce qui précède peut se résumer ainsi :

Conclusion Un champ de vecteurs sur un ouvert U de $\mathbb{R}^n \times \mathbb{R}^n$ est localement hamiltonien [resp. hamiltonien] si et seulement si la forme différentielle qui lui correspond par l'isomorphisme ϖ est fermée [resp. exacte].

2 Forme canonique et coordonnées symplectiques

2.1 Forme canonique

Nous avons vu au paragraphe 1.1 que dans le cas $n = 1$, v est localement hamiltonien si et seulement si son flot conserve les aires. Dans le cas général, on a la proposition suivante.

PROPOSITION

v est localement hamiltonien si et seulement si le flot de v conserve la forme différentielle de degré 2, dite **forme canonique** sur $\mathbb{R}^n \times \mathbb{R}^n$:

$$\omega = \sum_{i=1}^n dp_i \wedge dq_i. \quad (3)$$

Preuve

Soit ψ^t le flot (local) de v à l'instant t . Il s'agit de démontrer que v est localement hamiltonien si et seulement si la forme différentielle $\psi^{t*}\omega$ est indépendante du temps, c'est-à-dire :

$$\frac{d}{dt} \psi^{t*} \omega = 0.$$

Cela résultera immédiatement du lemme ci-après. □

Lemme $\frac{d}{dt} \psi^{t*} \omega = \psi^{t*} d(\varpi(v))$, où $\varpi(v)$ est la forme différentielle introduite en 1.2.

Preuve

Comme ψ^{t*} est un morphisme (commutant avec d) de l'algèbre graduée des formes différentielles, on a :

$$\psi^{t*} \omega = \sum_{i=1}^n d(\psi^{t*} p_i) \wedge d(\psi^{t*} q_i).$$

Il ne reste plus qu'à calculer l'action de d/dt sur le second membre en se souvenant qu'il s'agit d'une dérivation de l'algèbre des formes différentielles :

$$\frac{d}{dt}(\alpha \wedge \beta) = \frac{d\alpha}{dt} \wedge \beta + \alpha \wedge \frac{d\beta}{dt},$$

que cette dérivation commute avec d , et que par définition du flot ψ^t :

$$\begin{aligned} \frac{d}{dt}(\psi^{t*} p_i) &= \psi^{t*} v_{p_i} \\ \frac{d}{dt}(\psi^{t*} q_i) &= \psi^{t*} v_{q_i}. \end{aligned}$$

Le détail du calcul est laissé au lecteur. □

Remarque : lien entre la forme canonique ω et l'isomorphisme ϖ de 1.2. Un calcul immédiat montre que la 1-forme $\varpi(v)$ vérifie :

$$(\varpi(v))(w) = \omega(v, w). \quad (4)$$

Autrement dit, l'application ϖ de la remarque 1.2 est l'application linéaire de l'espace tangent dans l'espace cotangent associée à la forme bilinéaire ω (forme bilinéaire sur l'espace tangent).

L'affirmation que ϖ est un isomorphisme signifie que la forme bilinéaire ω est non dégénérée.

Parenthèse culturelle : la notion générale de « variété symplectique ». On appelle *variété symplectique* un couple (M, ω) où M est une variété différentielle de dimension paire $2n$ et ω une forme différentielle de degré 2 sur M , non dégénérée, et fermée ($d\omega = 0$). Un théorème général de Darboux (1842-1917) affirme que toute variété symplectique peut être munie localement de systèmes de coordonnées $(p, q) = (p_1, \dots, p_n; q_1, \dots, q_n)$, dites « coordonnées canoniques », telles que :

$$\omega = \sum_{i=1}^n dp_i \wedge dq_i.$$

Toutes les idées que nous développons ici pourraient être formulées dans le cadre général des variétés symplectiques.

2.2 La grande idée de Jacobi

Revenons à notre approche pédestre où M est un ouvert de $\mathbb{R}^n \times \mathbb{R}^n$, les coordonnées « canoniques » (p, q) étant les coordonnées usuelles de $\mathbb{R}^n \times \mathbb{R}^n$.

Considérons dans M un autre système de coordonnées :

$$(P, Q) = (P_1, \dots, P_n; Q_1, \dots, Q_n)$$

choisi de façon que l'expression de la forme ω dans ces coordonnées soit la même que dans les coordonnées (p, q) , à savoir :

$$\omega = \sum_{i=1}^n dP_i \wedge dQ_i.$$

Autrement dit, (P, Q) est un autre système de coordonnées « canoniques » de (M, ω) – nous dirons plutôt **coordonnées symplectiques**, continuant à réserver l'épithète « canonique » aux coordonnées usuelles de $\mathbb{R}^n \times \mathbb{R}^n$.

Il résulte de la remarque du paragraphe 1.2 que l'isomorphisme ω est donné dans les coordonnées (P, Q) par le même type de formule que dans les coordonnées (p, q) , de sorte que le champ hamiltonien d'une fonction H s'exprime encore en coordonnées (P, Q) par les formules (1), avec $\partial_{p_i}, \partial_{q_i}$ remplacés par $\partial_{P_i}, \partial_{Q_i}$.

Cette remarque est la base de la grande idée de Jacobi² en mécanique hamiltonienne : pour faciliter l'intégration d'un champ de vecteurs hamiltonien $v = v_H$, on cherche un système de coordonnées symplectiques (P, Q) dans lequel la fonction de Hamilton H s'exprime aussi simplement que possible. Par exemple, si H ne dépend que de P_1, \dots, P_n et pas des Q_i , les équations de Hamilton s'écrivent :

$$\begin{aligned} \frac{dP_i}{dt} &= v_{P_i} = -\frac{\partial H}{\partial Q_i} = 0 \\ \frac{dQ_i}{dt} &= v_{Q_i} = \frac{\partial H}{\partial P_i}; \end{aligned}$$

le premier groupe d'équations nous dit que P_1, \dots, P_n sont des *constantes du mouvement* ; reportant les valeurs de ces constantes dans les seconds membres du deuxième groupe d'équations, on voit que les Q_i sont *fonctions affines du temps*.

Un cas idéalement simple est celui où l'on a trouvé un système de coordonnées symplectiques tel que $H = P_1$. Alors les équations de Hamilton s'écrivent :

$$\begin{aligned} \frac{dP_i}{dt} &= 0 \\ \frac{dQ_i}{dt} &= 0 \text{ si } i \neq 1 \\ &= 1 \text{ si } i = 1. \end{aligned}$$

de sorte que toutes les coordonnées Q_i sont constantes sauf Q_1 qui est de la forme $t + Cte$!

La mise en œuvre effective de la stratégie de Jacobi est un chapitre très riche de la Mécanique, qu'il n'est pas question de traiter en quelques pages ! Nous nous contenterons, au paragraphe 4, d'en illustrer quelques aspects dans le cas très simple du problème de Kepler à deux dimensions.

² Carl Gustav Jacobi, 1804-1851

3 Symétrie et lois de conservation

3.0 Crochet de Lie de deux champs de vecteurs

Oublions provisoirement la structure symplectique pour revenir à l'étude générale des champs de vecteurs sur une variété différentielle M . On sait que les champs de vecteurs peuvent être identifiés localement aux dérivations qu'ils définissent sur l'algèbre des fonctions (« dérivations directionnelles », cf. chap. 5, § 3.2). Mais pour la clarté des raisonnements qui vont suivre il sera préférable de ne pas noter une dérivati on directionnelle par la même lettre v que le vecteur qui la définit : nous la noterons L_v , la lettre L étant l'initiale du nom du mathématicien norvégien Sophus Lie (1842-1899).

Soient donc v et w deux champs de vecteurs sur M , et soit :

$$[L_v, L_w] := L_v L_w - L_w L_v$$

le commutateur des dérivations correspondantes.

Il est facile de voir que le commutateur de deux dérivations est une dérivati on : il s'agit là d'un résultat d'algèbre pure – vrai sur une algèbre quelconque – dont la vérification est élémentaire.

Dans notre contexte, où l'algèbre considérée est celle des fonctions différentiables, nous allons montrer que le « commutateur de deux dérivations directionnelles est une dérivati on directionnelle ».

Lemme Il existe sur M un unique champ de vecteurs, noté $[v, w]$, tel que :

$$L_{[v, w]} = [L_v, L_w].$$

Ce champ de vecteurs est appelé **crochet de Lie** des deux champs de vecteurs v et w .

Évidemment, le crochet de Lie est antisymétrique :

$$[v, w] = -[w, v]$$

et il vérifie l'**identité de Jacobi** :

$$[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0.$$

Preuve

Dans le domaine d'une carte x , on peut écrire :

$$\begin{aligned} L_v &= \sum_i v_i^x \partial_{x_i} \\ L_w &= \sum_i w_i^x \partial_{x_i} \end{aligned}$$

où les v_i^x , w_i^x sont les coordonnées des vecteurs v, w dans la carte x . Compte tenu de la relation bien connue $[\partial_{x_i}, \partial_{x_j}] = 0$, un calcul immédiat donne :

$$[L_v, L_w] = L_u,$$

où u est le champ de vecteurs dont les coordonnées dans la carte x sont définies par :

$$u_i^x = \sum_{j=1}^m v_j^x (\partial_{x_j} w_i^x).$$

Bien que cette formule n'ait de sens que dans le domaine de la carte x , le champ de vecteurs u qu'elle y définit ne dépend pas du choix de la carte (car la dérivation correspondante a été définie de façon intrinsèque). Ce champ de vecteurs est donc bien défini sur tout M . \square

PROPOSITION Pour deux champs de vecteurs v, w de M les propriétés suivantes sont équivalentes :

- i) $[v, w] = 0$;
- ii) w est invariant par le flot de v .

Preuve

Soit ψ_v^t le flot de v au temps t , et soit w_t l'image du champ de vecteurs w par ce flot. La proposition résulte immédiatement du lemme ci-après. \square

Lemme $\frac{d}{dt} w_t = [w_t, v]$.

Preuve

Il s'agit de démontrer que pour toute fonction f (indépendante du temps) on a :

$$\frac{d}{dt} (L_{w_t} f) = [L_{w_t}, L_v] f.$$

Or, par définition de w_t , on a :

$$L_{w_t} = \phi_v^{-t*} L_w \phi_v^{t*}.$$

D'autre part, la définition de ψ_v^t comme flot de v se traduit par les relations suivantes entre opérateurs sur les fonctions dépendant du temps :

$$\left[\frac{d}{dt}, \phi_v^{t*} \right] = L_v \phi_v^{t*} = \phi_v^{t*} L_v$$

(ainsi que les relations analogues pour ϕ_v^{-t*} avec L_v remplacé par $-L_v$). Le lemme en résulte par calcul algébrique immédiat. \square

Remarque Les propriétés i) et ii) de la proposition impliquent aussi les suivantes (qui leur sont équivalentes) :

- iii) le flot de v transforme les courbes intégrales de ω en courbes intégrales de ω ;
- iv) les flots de v et de ω commutent :

$$\phi_v^t \phi_\omega^t = \phi_\omega^t \phi_v^t .$$

3.1 Crochet de Poisson de deux fonctions sur une variété symplectique

Revenons à la géométrie symplectique, avec notre « forme canonique » ω sur $\mathbb{R}^n \times \mathbb{R}^n$ (cf. § E.2 ; nous pourrions aussi nous placer sur une « variété symplectique »). Nous allons montrer que l'opération « crochet de Lie », si on la restreint aux champs hamiltoniens, se traduit sur les fonctions de Hamilton par une opération appelée « crochet de Poisson ».

PROPOSITION-DÉFINITION

On appelle **crochet de Poisson** de deux fonctions H, K la fonction $\{H, K\} := L_{v_H} K$. Le crochet de Poisson $\{ \cdot, \cdot \}$ est antisymétrique et vérifie l'identité de Jacobi. De plus,

$$v_{\{H, K\}} = [v_H, v_K].$$

Preuve

Dans les coordonnées canoniques (p, q) le crochet de Poisson est défini par :

$$\{H, K\} = \sum_{i=1}^n \left(\frac{\partial H}{\partial p_i} \frac{\partial K}{\partial q_i} - \frac{\partial H}{\partial q_i} \frac{\partial K}{\partial p_i} \right).$$

Toutes les assertions de la proposition peuvent être vérifiées sur cette expression. □

Remarque Il résulte de sa définition que le crochet de Poisson peut encore s'écrire :

$$\begin{aligned} \{H, K\} &= L_{v_H} K = dK(v_H) = (\varpi(v_K))(v_H) \\ &= \omega(v_K, v_H) \end{aligned}$$

(cf. remarque du § 2.1).

EXERCICE TEST Pour deux champs de vecteurs quelconques v et w sur une variété, la valeur en a du crochet de Lie :

$$[v, w](a)$$

ne dépend-elle que des valeurs de v et w en a ? Même question pour deux champs de vecteurs hamiltoniens sur une variété symplectique.

3.2 Fonctions en involution : symétrie et lois de conservation

Deux fonctions H et K sont dites **en involution** si $\{H, K\} = 0$.

D'après la définition du crochet de Poisson, cela équivaut à dire que K est une constante du mouvement pour le champ hamiltonien de H (et vice versa par antisymétrie).

Cela équivaut aussi à :

$$[v_H, v_K] = 0 \quad (\text{cf. Prop. du § 3.1}),$$

avec toutes les conséquences qui en résultent par la proposition du paragraphe E.3.0.

Exemples

i) Considérons sur $\mathbb{R}^2 \times \mathbb{R}^2$ le moment angulaire :

$$K(p, q) = q_1 p_2 - q_2 p_1.$$

Son champ hamiltonien s'écrit :

$$v_K = -q_2 \partial_{q_1} + q_1 \partial_{q_2} - p_2 \partial_{p_1} + p_1 \partial_{p_2}.$$

Le flot de v_K est donc le groupe à un paramètre d'automorphismes linéaires de $\mathbb{R}^2 \times \mathbb{R}^2$ défini par :

$$\phi_{v_K}^t : (p, q) \mapsto (R(t)p, R(t)q), \quad \text{où } R(t) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$$

c'est le groupe des rotations $SO(2)$ agissant « de façon diagonale » sur $\mathbb{R}^2 \times \mathbb{R}^2$.

Dire qu'une fonction H sur $\mathbb{R}^2 \times \mathbb{R}^2$ est invariante par rotation équivaut à dire que le moment angulaire K est une constante du mouvement pour v_H , comme nous l'avons déjà remarqué dans l'étude 3^{bis} (§ 1.2) à propos du problème de Kepler à deux dimensions.

ii) Si maintenant on se place sur $\mathbb{R}^3 \times \mathbb{R}^3$, un calcul analogue peut être fait séparément pour chacune des trois composantes J_1, J_2, J_3 du vecteur « moment angulaire » \vec{J} :

$$J_1 = q_2 p_3 - q_3 p_2 \quad (\text{et permutations circulaires}) :$$

H est invariant par rotation autour du i -ième axe ($i = 1, 2, 3$) si et seulement si J_i est une constante du mouvement pour v_H ; en particulier, si H est invariant par le groupe des rotations $SO(3)$ tout entier (agissant de façon diagonale sur $\mathbb{R}^3 \times \mathbb{R}^3$), chacune des trois composantes de \vec{J} est une constante du mouvement.

EXERCICE Calculer les crochets de Poisson $\{J_i, J_j\}$. À quelle propriété du groupe $SO(3)$ correspond le fait qu'ils ne soient pas nuls ?

3.3 Systèmes de fonctions en involution

Soit $F = (F_1, \dots, F_\ell)$ un ℓ -uplet de fonctions à différentielles linéairement indépendantes sur une variété symplectique M . Comme les dF_i sont les images par l'isomorphisme ω des champs hamiltoniens v_{F_i} , l'hypothèse d'indépendance des covecteurs dF_i équivaut à l'indépendance des vecteurs v_{F_i} .

Examinons ce que signifie la propriété pour (F_1, \dots, F_ℓ) d'être **en involution**, c'est-à-dire :

$$\{F_i, F_j\} = 0 \quad \text{pour tous } i, j = 1, \dots, \ell$$

(propriété équivalente, d'après la proposition du paragraphe 3.1, à :

$$[v_{F_i}, v_{F_j}] = 0 \quad i, j = 1, \dots, \ell).$$

Plaçons-nous d'abord en un point $a \in M$, et posons :

$$E = T_a M, \quad E^* = T_a^* M.$$

Dans E et E^* considérons les sous-espaces vectoriels à ℓ dimensions :

V engendré par $v_1 = v_{F_1}(a), \dots, v_\ell = v_{F_\ell}(a)$,

W engendré par $dF_1(a), \dots, dF_\ell(a)$.

Évidemment, l'isomorphisme ω envoie V sur W .

Avec ces notations l'involutivité ponctuelle en a :

$$\{F_i, F_j\}(a) = 0 \quad i, j = 1, \dots, \ell$$

équivaut par définition du crochet de Poisson à :

i) $dF_j(v_i) = 0 \quad i, j = 1, \dots, \ell$ c'est-à-dire $W \subset V^\perp$, où $V^\perp \subset E^*$ est l'espace des covecteurs (formes linéaires sur E) qui s'annulent sur V .

Cette condition d'involutivité peut aussi s'écrire :

$$\omega(V) \subset V^\perp,$$

ce qui équivaut à :

ii) $\omega|_V = 0$: la forme bilinéaire ω s'annule en restriction à V .

En comparant les dimensions : $\dim V = \dim W = \ell$, $\dim V^\perp = 2n - \ell$, on en déduit que $\ell \leq n$:

sur une variété symplectique de dimension $2n$, il ne peut pas exister plus de n fonctions en involution à différentielles linéairement indépendantes.

Faisant varier le point a nous déduisons de i) la traduction suivante de l'hypothèse d'involutivité :

i)' les champs de vecteurs $v_{F_1}, \dots, v_{F_\ell}$ sont tangents aux sous-variétés de niveau de $F = (F_1, \dots, F_\ell)$.

Un cas remarquable est le cas $\ell = n$: pour un n -uplet de fonctions F_1, \dots, F_n à différentielles linéairement indépendantes, la propriété d'involutivité équivaut à la suivante :

ii)' la restriction de ω aux variétés de niveau de $F = (F_1, \dots, F_n)$ est nulle.

Un tel n -uplet est appelé **système complet de fonctions en involution**.

Exemples Soit $(P, Q) = (P_1, \dots, P_n; Q_1, \dots, Q_n)$ un système de coordonnées symplectiques sur M . Alors le n -uplet P_1, \dots, P_n est un système complet de fonctions en involution. Il en est de même, plus généralement, du n -uplet :

$$P_{i_1}, \dots, P_{i_k}, \quad Q_{j_1}, \dots, Q_{j_{n-k}},$$

où $\{1, 2, \dots, n\} = \{i_1, \dots, i_k\} \cup \{j_1, \dots, j_{n-k}\}$.

4 Exemples de coordonnées symplectiques : les coordonnées « actions-angles »

4.0 Retour au problème de Kepler Reprenons les hypothèses et notations de l'étude $\mathfrak{3}^{bis}$, paragraphe 2 (problème de Kepler à deux dimensions).

Le champ hamiltonien v_H admet deux constantes du mouvement en involution H et K : le « **hamiltonien** » $H = \frac{p_1^2 + p_2^2}{2m} + V_r \left(r = \sqrt{q_1^2 + q_2^2} \right)$, dont la valeur s'interprète comme « l'énergie » E ; le « **moment angulaire** » $K = q_1 p_2 - q_2 p_1$.

On a vu que, sous des hypothèses convenables sur le potentiel V , les surfaces de niveau $\{H = E, K = k\}$ sont difféomorphes à des tores \mathbb{T}^2 pour toute une plage U de valeurs de (E, k) : par exemple si V est voisin du potentiel newtonien $-c/r$, la « **plage** » U peut être prise de la forme :

$$U : \begin{cases} k > 0 & (\text{ou, si l'on préfère, } k < 0) \\ E_{min}(k) < E < 0, \end{cases}$$

où $E_{min}(k)$ est la valeur minimale du « **potentiel effectif** » représenté sur la figure 7 de l'étude $\mathfrak{3}^{bis}$, paragraphe 2.2.

Plus précisément, si l'on désigne par \mathcal{U} l'ouvert image réciproque de U par l'application :

$$F = (H; K) : \mathbb{R}^4 \rightarrow \mathbb{R} \times \mathbb{R}.$$

l'application $F|_{\mathcal{U}}$ est une fibration triviale de base U , ayant pour fibre modèle le tore \mathbb{T}^2 .

Nous nous proposons de montrer que l'ouvert \mathcal{U} peut être muni d'un système de coordonnées symplectiques $(I, J; \theta_I, \theta_J)$, dites **coordonnées actions-angles**, où :

- i) I et J (les coordonnées d'action) sont fonctions de $H(p, q)$ et $K(p, q)$ seulement (en fait on aura $J = K$) et définissent un changement de coordonnées sur la base U ;
- ii) θ_I et θ_J sont des fonctions à valeurs dans $\mathbb{R} \bmod 2\pi\mathbb{Z}$, qui définissent un système de « **coordonnées angulaires** » dans chacun des tores-fibres.

Comme il s'agira de coordonnées symplectiques, les équations du mouvement pourront s'écrire d'après le paragraphe 2.2 :

$$\begin{cases} I & = & \text{Cte} \\ J & = & \text{Cte} \\ \frac{d}{dt}\theta_I & = & \alpha_I \\ \frac{d}{dt}\theta_J & = & \alpha_J \end{cases}$$

où les α_I, α_J sont des constantes sur chaque tore-fibre, données par :

$$\alpha_I = \partial H / \partial I \quad , \quad \alpha_J = \partial H / \partial J.$$

Le flot de H sera donc, en restriction à chaque tore, le flot d'un champ de vecteurs constant dans les coordonnées angulaires θ_I, θ_J , comme annoncé dans l'étude $\mathfrak{5}^{bis}$, paragraphe 2 ii).

4.1 Construction générale des coordonnées d'action

Considérons la situation générale suivante. Sur un ouvert \mathcal{U} de $\mathbb{R}^n \times \mathbb{R}^n$, muni de la forme canonique $\omega = \sum_{i=1}^n dp_i \wedge dq_i$, soit $F = (F_1, \dots, F_n)$ un système complet de fonctions en involution définissant une fibration triviale :

$$F = (F_1, \dots, F_n) : \mathcal{U} \rightarrow U$$

ayant pour base un ouvert simplement connexe U de \mathbb{R}^n et pour fibre-modèle le tore \mathbb{T}^n .

Considérons alors sur U la forme différentielle de degré 1 :

$$\sigma = \sum_{i=1}^n p_i dq_i.$$

Sa différentielle extérieure $d\sigma$ coïncide avec ω , dont la restriction aux tores-fibres $F^{-1}(u)$ ($u \in U$) est nulle d'après 3.3 ii). Autrement dit, la forme différentielle σ est fermée en restriction aux tores-fibres $F^{-1}(u)$. D'après le théorème de Stokes, l'intégrale $\int_{\gamma(u)} \sigma|_{F^{-1}(u)}$ sur un cycle $\gamma(u)$ de $F^{-1}(u)$ ne dépend donc que de la classe d'homologie de ce cycle.

Soit donc $\gamma_1, \dots, \gamma_n$ une base du groupe d'homologie $H_1(\mathbb{T}^n)$ (cf. chap. 8, § 4.3). Par trivialisations du fibré $F : \mathcal{U} \rightarrow U$, on en déduit une base $\gamma_1(u), \dots, \gamma_n(u)$ de $H_1(F^{-1}(u))$, dépendant continûment de u .

Pour $i = 1, \dots, n$, posons :

$$I_i(u) = \frac{1}{2\pi} \int_{\gamma_i(u)} \sigma|_{F^{-1}(u)}.$$

On obtient ainsi n fonctions différentiables sur U , dont on supposera qu'elles forment un système de coordonnées sur U (localement, c'est-à-dire pour U assez petit, cela revient à supposer que dl_1, \dots, dl_n sont linéairement indépendants).

Posons enfin $I_i^F = I_i \circ F$ ($i = 1, \dots, n$). On obtient ainsi sur \mathcal{U} un système complet de fonctions en involution I_1^F, \dots, I_n^F (en involution car $I^F := (I_1^F, \dots, I_n^F)$ a les mêmes variétés de niveau que $F = (F_1, \dots, F_n)$ qui est un système en involution). Ce sont les coordonnées « d'action » cherchées.

EXERCICE TEST Énoncer (et démontrer) la propriété générale des systèmes de fonctions en involution que nous venons d'utiliser.

L'« intégrale d'action » vue comme intégrale indéfinie

On a vu que pour tout $u \in U$ la forme différentielle de degré 1 $\sigma|_{F^{-1}(u)}$ est fermée. Localement sur $F^{-1}(u)$ elle admet donc une primitive (cf. Chap. 8, § 1.1, Exercice), que l'on peut définir sans ambiguïté au voisinage d'un « point-base » donné $b_u \in F^{-1}(u)$ en convenant qu'elle vaut 0 en ce point : cette primitive est donnée par la formule intégrale :

$$S_u(x) = \int_{b_u}^x \sigma|_{F^{-1}(u)}.$$

Bien définie pour x voisin de b_u (et plus précisément dans tout voisinage simplement connexe de b_u), cette formule devient ambiguë quand x parcourt tout le tore car l'intégrale dépend de la classe d'homotopie du chemin choisi pour aller de b_u à x sur le tore.

S_u est donc une « fonction multiforme », c'est-à-dire une fonction définie sur un revêtement universel de l'espace $F^{-1}(u)$ sur lequel on prétendait la définir : quand x parcourt le cycle γ_i de $F^{-1}(u)$, la fonction S_u est modifiée par l'ajout de $2\pi I_i(u)$.

Nous supposons dans la suite que le point-base b_u a été choisi de façon à dépendre différemment de u (ce qui est toujours possible puisqu'il s'agit de choisir une section différentiable d'un fibré trivial). Il en résulte que la fonction :

$$S : (u, x) \mapsto S_u(x)$$

est différentiable.

4.2 Construction générale des coordonnées angulaires

Nous venons de construire une famille de fonctions (S_u) vérifiant pour tout u :

$$dS_u = \sigma | F^{-1}(u).$$

Il en résulte l'égalité :

$$dS = \sigma \pmod{(dF_1, \dots, dF_n)},$$

que l'on peut encore écrire :

$$dS = \sigma \pmod{(dI_1^F, \dots, dI_n^F)},$$

puisque $F = (F_1, \dots, F_n)$ et $I^F = (I_1^F, \dots, I_n^F)$ ont mêmes variétés de niveau.

Autrement dit, il existe sur le revêtement universel de \mathcal{U} (puisque c'est là qu'était définie la fonction S) des fonctions $\theta_1, \dots, \theta_n$ telles que :

$$dS = \sigma + \theta_1 dI_1^F + \dots + \theta_n dI_n^F. \quad (1)$$

La nature de la « multiformité » de ces fonctions θ_i se déduit immédiatement de celle de S : on a vu en 4.1 que quand x parcourt le cycle γ_i d'un tore-fibre, la fonction S est modifiée par l'ajout de $2\pi I_i = 2\pi I_i^F$; sa différentielle dS est donc modifiée par l'ajout de $2\pi dI_i^F$; en lisant le second membre de l'équation (1) on en déduit que θ_i est changé en $\theta_i + 2\pi$, tandis que les fonctions θ_j ($j \neq i$) restent inchangées.

Les fonctions $\theta_1, \dots, \theta_n$ sont donc bien définies sur \mathcal{U} si on les considère comme fonctions à valeurs dans $\mathbb{R}/2\pi\mathbb{Z}$. De plus, elles vérifient les relations :

$$\int_{\gamma_i} d\theta_j = 2\pi \delta_{ij}. \quad (2)$$

4.3 Preuve que $(I_1^F, \dots, I_n^F, \theta_1, \dots, \theta_n)$ est un système de coordonnées symplectiques

Différentiant l'équation (1) on trouve :

$$0 = d\sigma + d\theta_1 \wedge dI_1^F + \dots + d\theta_n \wedge dI_n^F$$

c'est-à-dire :

$$\omega = dI_1^F \wedge d\theta_1 + \dots + dI_n^F \wedge d\theta_n, \tag{3}$$

comme attendu d'un système de coordonnées symplectiques.

Il reste donc seulement à montrer que $(I_1^F, \dots, I_n^F, \theta_1, \dots, \theta_n)$ est bien un système de coordonnées sur \mathcal{Q}_U , autrement dit que l'application :

$$(I^F, \theta) = (I_1^F, \dots, I_n^F, \theta_1, \dots, \theta_n) : \mathcal{Q}_U \rightarrow U \times \mathbb{T}^n \tag{4}$$

est un difféomorphisme.

Qu'il s'agisse d'un difféomorphisme local est un exercice élémentaire d'algèbre linéaire : sur l'espace tangent $T_{\alpha} \mathcal{Q}_U$ on sait déjà que les formes linéaires dI_1^F, \dots, dI_n^F sont linéairement indépendantes, et que la forme bilinéaire ω est non dégénérée ; un calcul élémentaire de déterminant, dans une base de $T_{\alpha} \mathcal{Q}_U$ « adaptée » au sous-espace vectoriel $\text{Ker}(dI_1^F, \dots, dI_n^F)$, permet de déduire de (3) que $d\theta_1, \dots, d\theta_n$ sont linéairement indépendantes en restriction à ce sous-espace vectoriel.

Nous savons donc que l'application (4) est un difféomorphisme local. Pour montrer que c'est un difféomorphisme global on peut raisonner en restriction à chacun des tores-fibres. Or il s'agit d'une application propre (car un tore est un espace compact). D'après la proposition ii) du chapitre 7, paragraphe 2, c'est donc un revêtement de degré fini. Il reste à montrer que ce revêtement est de degré 1, c'est-à-dire (cf. Chap. 7, § 2.3) que le groupe fondamental de sa base agit trivialement sur sa fibre : or c'est précisément ce que nous disent les relations (2).

4.4 L'exemple du problème de Kepler

Pour expliciter les constructions précédentes dans l'exemple du problème de Kepler à deux dimensions, il est commode de passer en coordonnées polaires : aux coordonnées polaires (r, φ) dans le plan des q correspond dans « l'espace de phase » de $\mathbb{R}^2 \times \mathbb{R}^2$ (privé de $\{q = 0\}$) le système de coordonnées symplectiques :

$$(p_r, p_\varphi ; r, \varphi)$$

où p_r et p_φ sont respectivement la composante radiale de l'impulsion p et le moment angulaire :

$$\begin{aligned} p_r &= (p_1 q_1 + p_2 q_2) / r \\ p_\varphi &= q_1 p_2 - q_2 p_1. \end{aligned}$$

EXERCICE Vérifier qu'il s'agit bien de coordonnées symplectiques dans lesquelles la forme $\sigma = p_1 dq_1 + p_2 dq_2$ s'écrit $\sigma = p_r dr + p_\varphi d\varphi$.

Dans ces coordonnées, les deux fonctions en involution H, K de 4.0 s'écrivent :

$$\begin{cases} H &= \frac{1}{2m} \left(p_r^2 + \frac{p_\varphi^2}{r^2} \right) + V(r) \\ K &= p_\varphi. \end{cases}$$

Pour une valeur (E, k) de $F = (H, K)$ choisie dans la « plage » U de 4.0, l'ensemble $F^{-1}(E, k)$ peut être décrit ainsi :

- $p_\varphi = k$
- φ parcourt le cercle $\mathbb{R}/2\pi\mathbb{Z}$
- (r, p_r) parcourt le « cercle » $C_{E,k} : \frac{p_r^2}{2m} + V_k(r) = E$; cf. Fig. 4, de l'étude 3^{bis}, § 0 avec (q, p) remplacés par r, p_r , avec comme potentiel le « potentiel radial effectif » :

$$V_k(r) = V(r) + \frac{k}{2mr^2}.$$

La surface de niveau $F^{-1}(E, k)$ est donc bien un tore, dont l'homologie en dimension 1 peut être engendrée par les deux cycles γ_I et γ_J que voici :

γ_I : φ reste constant tandis que le point (r, p_r) décrit la courbe $C_{E,k}$ dans le sens des flèches de la figure 4 de l'étude 3^{bis} ;

γ_J : (r, p_r) reste constant, tandis que φ parcourt le cercle $\mathbb{R}/2\pi\mathbb{Z}$ dans le sens direct.

À ces deux cycles correspondent les deux « actions » :

$$\begin{aligned} I(E, k) &= \frac{1}{2\pi} \int_{\gamma_I} \sigma | F^{-1}(E, k) = \frac{1}{2\pi} \int_{C_{E,k}} p_r dr \\ J(E, k) &= \frac{1}{2\pi} \int_{\gamma_J} \sigma | F^{-1}(E, k) = \frac{1}{2\pi} \int_0^{2\pi} k d\varphi = k. \end{aligned}$$

Vérifions que ces deux fonctions I, J forment bien un système de coordonnées sur l'ouvert U .

Lemme $\partial I / \partial E > 0$ pour tout E dans l'intervalle $E_{\min}(k) < E < 0$.

Preuve

C'est immédiat sur l'expression intégrale de $\partial I / \partial E$ obtenue en dérivant

$$I(E, k) = \frac{1}{\pi} \int_{r_-(k)}^{r_+(k)} \sqrt{2m(E - V_k(r))} dr \text{ sous le signe d'intégration.} \quad \square$$

Corollaire L'application $(E, k) \mapsto (I(E, k), J(E, k) = k)$ est un difféomorphisme de l'ouvert U sur un ouvert de \mathbb{R}^2 .

Preuve

Le théorème d'inversion locale montre que c'est un difféomorphisme local. Pour voir qu'il envoie *bijectivement* U sur un ouvert de \mathbb{R}^2 , il suffit de raisonner à k fixé, en appliquant le théorème des valeurs intermédiaires à la fonction monotone $E \mapsto I(E, k)$.

Nos « coordonnées d'action » $I_1 = I, I_2 = J$ vérifient donc bien les hypothèses du § 4.1. Nous allons maintenant expliciter géométriquement les coordonnées angulaires correspondantes, que nous noterons θ_I et θ_J . □

Par dérogation aux notations adoptées au § 4.1, nous noterons ici par la même lettre I [resp. J] la fonction des deux variables (E, k) qui vient d'être définie et la fonction des quatre variables (p_1, p_2, q_1, q_2) qui s'en déduit par la substitution :

$$\begin{aligned} E &= H(p, q) \\ k &= K(p, q) \end{aligned}$$

Comme $(I, J, \theta_I, \theta_J)$ est un système de coordonnées symplectiques, nous savons d'avance que le champ de vecteurs v_I [resp. v_J] aura pour flot :

$$\begin{aligned} (v_I) \begin{cases} \dot{\theta}_I = 1 & (\text{où } \dot{\theta}_I = d\theta_I / dt) \\ \dot{\theta}_J = \text{Cte} \end{cases} \\ \text{[resp. } (v_J) \begin{cases} \dot{\theta}_I = \text{Cte} \\ \dot{\theta}_J = 1 & (\text{où } \dot{\theta}_J = d\theta_J / dt). \end{cases} \end{aligned}$$

Or le flot de v_J est très facile à calculer en coordonnées polaires $(p_r, p_\varphi, r, \varphi)$: comme $J = K = p_\varphi$, on a $v_J = \partial_\varphi$, de sorte que les courbes intégrales de v_J sont les cercles concentriques de la figure 2, parcourus à vitesse angulaire uniforme $\dot{\varphi} = 1$.

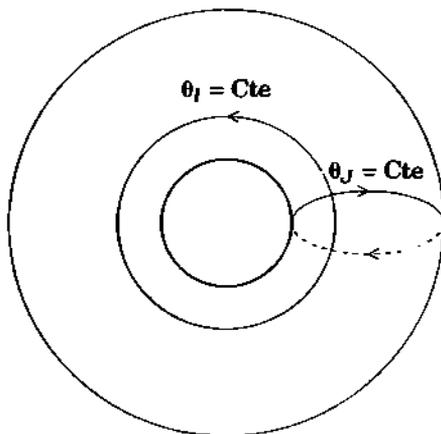


Figure 2. Lignes de niveau des deux coordonnées angulaires θ_I, θ_J , en restriction à l'un des tores $C_{E,k}$ (vu en projection sur le plan des q). Les lignes $\theta_I = \text{Cte}$ (trajectoires intégrales de v_J) sont les « grands cercles » du tore. On a dessiné aussi une ligne $\theta_J = \text{Cte}$, en représentant en trait plein [resp. en tirets] la partie située sur le « dessus » [resp. « dessous »] du tore, c'est-à-dire du côté $p_r > 0$ [resp. $p_r < 0$]. Toutes les autres lignes $\theta_J = \text{Cte}$ se déduisent de celle-là par rotation autour de l'origine, l'angle de rotation étant précisément la différence des valeurs choisies pour θ_J .

Il nous reste à interpréter le flot de v_I , qu'il sera instructif de comparer à celui de v_H .

Lemme $v_I = \frac{\partial I}{\partial E} v_H + \frac{\partial I}{\partial k} v_K$.

Preuve

C'est la traduction, par l'isomorphisme ω , de la relation :

$$dI = \frac{\partial I}{\partial E} dH + \frac{\partial I}{\partial k} dK. \quad \square$$

EXERCICE En explicitant le second membre du lemme (à l'aide de la représentation intégrale de I) montrer que :

$$v_I = \frac{T}{2\pi} v_H - \frac{\Delta\varphi}{2\pi} \partial_\varphi,$$

où T et $\Delta\varphi$ sont les quantités (constantes sur chaque tore) calculées dans l'étude 3^{bts} (période du mouvement radial et variation d'angle du flot de v_H) :

- période radiale $T =$ temps mis par (r, p_r) à parcourir le cycle $C_{E,k}$;
- variation d'angle $\Delta\varphi =$ accroissement de φ quand (r, p_r) parcourt $C_{E,k}$ (dans le cas du potentiel newtonien on a $\Delta\varphi = -2\pi$; sinon il faut corriger par la « précession »).

Interprétation géométrique Notant \dot{r}_I et $\dot{\varphi}_I$ [resp. \dot{r}_H et $\dot{\varphi}_H$] les dérivées par rapport au temps des coordonnées r et φ évoluant selon le flot hamiltonien de I [resp. H], on a :

$$\begin{cases} \dot{r}_I = \frac{T}{2\pi} \dot{r}_H \\ \dot{\varphi}_I = \frac{T}{2\pi} \dot{\varphi}_H - \frac{\Delta\varphi}{2\pi} \end{cases}$$

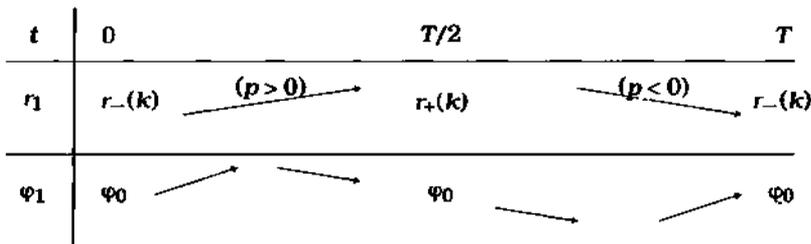
Un « changement d'échelle des temps » de la forme $t = \frac{2\pi}{T} s$ permet de récrire ces relations ainsi :

$$\begin{aligned} \frac{dr_I}{ds} &= \frac{dr_H}{dt} \\ \frac{d\varphi_I}{ds} &= \frac{d\varphi_H}{dt} - \frac{\Delta\varphi}{T} \end{aligned}$$

On voit qu'à ce changement d'échelle des temps près, le flot hamiltonien de I se déduit de celui de H de la façon suivante.

Imaginons notre tore comme un pneu de vélo, d'abord immobile, sur lequel on regarde une coccinelle se promener selon une courbe intégrale de v_H (cf. Fig. 5 de l'étude 3^{bts}) ; alors en faisant tourner le pneu avec une vitesse angulaire constante $-\frac{\Delta\varphi}{T}$ on verra la coccinelle se promener selon une courbe intégrale de v_I .

Ces courbes sont dessinées sur la figure 2. Pour se convaincre sans calculs qu'elles ont bien l'allure indiquée, il suffit de se souvenir que d'après la loi des aires de Kepler la vitesse angulaire $\dot{\varphi}_H$ est d'autant plus petite que r est grand ; comme la constante $\frac{\Delta\varphi}{T}$ qu'on retranche à $\dot{\varphi}_H$ pour obtenir $\dot{\varphi}_I$ n'est autre que la « vitesse angulaire moyenne », il en résulte pour les fonctions r_I et φ_I le tableau de variations suivant sur une période (en prenant comme origine des temps t le temps où r est minimal) :



Pour en savoir plus Les « coordonnées d'action » présentées en 4.1 ne sont pas de simples gadgets de mathématiciens. Elles ont une grande importance en physique, notamment dans l'étude des variations lentes des systèmes hamiltoniens (notion d'invariants adiabatiques).

Sur ce sujet, cf. Landau et Lifschitz, *Mécanique*, Éditions Mir, Moscou, 1966 et Arnold, *Méthodes mathématiques de la mécanique classique* (trad. fr. Mir, 1976).

A

action

 coordonnées d'(-), 251 ;

 intégrale d'(-), 251.

analytique, 12.

atlas

 continu, 111 ;

 différentiable, 112.

B

bord, 233.

C

caractéristique

 lieu (-), 177 ;

 point (-), 143, 144, 173.

carte, 24, 108 ;

 changement de (-), 56, 110.

catégoric, 58 ;

 des sous-variétés, 58.

chaîne singulière, 233, 234, 237.

champ

 de vecteurs, 75, 78, 139 ;

 courbe intégrale d'(-), 139 ;

 hamiltonien, 241.

classe

 (de différentiabilité) C^r , 12, 45, 55 ;

 fondamentale, 239.

cohomologie, 234.

cône dual, 120.

connexe (simplement -), 199.

constante du mouvement, 76.

contact (élément de -), 123, 148, 149, 182, 187.

contour apparent, 71.

contractile, 193.

coordonnées, 16, 18, 19, 24, 108, 135 ;

 homogènes, 114.

cotangent(e)

 application (-), 135 ;

 espace (-), 134 ;

 fibré (-), 140 ;

 vecteur (-), 134.

critique

 courbe (-), 71 ;

 point (-), 27, 31, 71.

crochet

 de Lie, 245 ;

 de Poisson, 247.

D

Darboux, 243.
 De Rham, 233, 235 ;
 dualité de (-). 235.
 dérive, 172.
 dérivée directionnelle, 8, 77, 135.
 développement limité, 12.
 difféomorphisme, 47, 59, 116 ;
 local, 48, 65.
 différentiable, 45, 57 ;
 application (-), 115 ;
 changement de variable (-), 16 ;
 continûment (-), 12 ;
 de classe C^r , 12, 45 ;
 indéfiniment (-), 12.
 différentielle, 45, 134 ;
 forme (-), 7, 141, 223 ;
 d'une fonction, 7, 27 ;
 d'une relation, 14, 15, 16 ;
 implication (-), 10 ;
 itérée, 11.
 dimension, 17, 51, 109.

E

énergie, 87.
 enveloppe, 144, 179 ;
 à la source, 177 ;
 au but, 177.
 équations locales, 52.
 espace projectif, 113, 117, 119 ;
 sous(-), 118.
 étoilé, 193.
 Euler (caractéristique d'-), 100.
 exacte (forme différentielle -), 8, 142, 227.
 extérieur(e)
 différentiation (-), 224 ;
 forme (-), 221 ;
 produit (-), 222.

F

fermée (forme différentielle -), 227.
 fibre, 139, 158, 162.
 fibré
 conormal, 168 ;
 cotangent, 140 ;
 normal, 168 ;
 restriction d'un (-), 162 ;
 trivial, 162 ;
 trivialisation d'un (-), 162 ;
 vectoriel, 163.
 flot, 76, 139.
 foncteur
 contravariant, 63 ;
 covariant, 63.
 fonctionnel(le)
 graphe (-), 43 ;
 relation (-), 4, 41.
 forme
 canonique, 242 ;
 différentielle, 7, 141, 223 ;
 volume, 227.

G

Galois (groupe de -), 210.
 galoisien (revêtement -), 210.
 gradient (champ de -), 85.
 grassmannienne, 114.
 groupe
 d'homotopie, 198 ;
 fondamental, 197.
 groupoïde, 195.

H

Hamilton (fonction de $-$), 77, 241.
 hamiltonien, 250 ;
 champ $(-)$, 241 ;
 système $(-)$, 77.
 hérisson, 92, 102.
 hessienne, 32.
 homologue, 235.
 homotopie, 191 ;
 catégorie d' $(-)$, 192 ;
 de chemins, 194 ;
 équivalence d' $(-)$, 192 ;
 groupe d' $(-)$, 198 ;
 type d' $(-)$, 192.

I

image réciproque
 d'une fonction, 57 ;
 d'une forme différentielle, 141.
 immersion(s), 50, 137.
 implicite(s), 5 ;
 fonctions $(-)$, 27, 28, 29.
 indice, 93, 96, 100, 218.
 induit(e)
 espace $(-)$, 161 ;
 section $(-)$, 161.
 intégrale
 courbe $(-)$, 75 ;
 première, 76 ;
 trajectoire $(-)$, 76.
 inversion locale, 30, 48.
 involution (fonctions en $-$), 247, 249.
 isomorphisme, 59.
 isotope, 218.

J, K

Jacobi, 245 ;
 identité de $(-)$, 245.
 jacobienne, 46.
 jct, 132.
 Kepler (lois de $-$), 88.

L

lacet, 191.
 Legendre, 188.
 legendrienne
 courbe $(-)$, 149 ;
 variété $(-)$, 188.
 Leibniz, 6.
 Lie, 245.
 lisse
 relation $(-)$, 16 ;
 sous-ensemble $(-)$, 22, 51.

M

moment angulaire, 89, 250.
 morphisme, 59.
 Morse, 37.

N

Newton, 76, 88, 107.

O

orientable, 213, 214.
 orientation, 212, 213, 214.
 oscillateur, 86 ;
 harmonique, 86.

P

paramétrage local, 25.

plongement, 59, 118.

Poincaré

isomorphisme de (-), 239 ;

lemme de (-), 227.

Poisson (crochet de -), 247.

produit fibré, 160, 162.

propre, 54.

singulier (point - d'un champ de vecteurs), 81.

sous-variété, 22, 51, 118.

sphère de Riemann, 117.

Stokes (théorème de -), 232.

submersion, 49, 137.

symplectique(s)

coordonnées (-), 244 ;

variété (-), 243.

Q

quadratique non dégénéré (point critique -), 32.

R

redressement local, 51 ;

d'un champ de vecteurs, 81, 139.

relation, 39 ;

fonctionnelle, 41.

rétracté, 192 ;

par déformation, 192.

rétractile, 192 ;

par déformation, 192.

rétraction, 59, 192 ;

par déformation, 192.

revêtement, 200 ;

universel, 207.

S

section

d'un espace projeté, 158 ;

d'un fibré, 139.

simple (racine -), 31.

T

tangent(e)

à l'ordre r , 131, 133 ;

application (-), 61, 133 ;

espace (-), 23, 108, 133 ;

fibré (-), 23, 139 ;

vecteur (-), 23, 133.

Taylor (série de -), 12.

tore, 69, 127, 154, 216.

transverse, 67.

V

variable (changement de -), 41.

variété

différentielle, 108 ;

produit, 116 ;

projective, 120 ;

riemannienne, 142, 228

topologique, 110.