

## Calcul différentiel et intégral 2

Le module appelé “Calcul différentiel et intégral 2” fait partie des modules obligatoires de la troisième année de la licence de mathématiques à l’UHP. Il fait suite à un “Calcul différentiel et intégral 1” qui présente, en principe, les bases de ce calcul différentiel.

Ce second module a une cote mal taillée. Il n’est pas possible d’envisager d’y aborder sérieusement certains sujets, comme la géométrie des surfaces dans l’espace de dimension 3, ou de s’y placer dans un cadre adapté, comme celui des variétés différentielles. En effet les connaissances de base des étudiants sont bien trop fragiles. On ne peut en particulier pas faire comme s’ils avaient assimilé le programme d’algèbre linéaire du L1 et du L2, le programme d’analyse de ces mêmes années, la topologie générale du premier semestre du L3 et le calcul différentiel de ce même semestre.

Le contenu réalise donc un perpétuel compromis entre une première visite de quelques sujets nouveaux, parmi lesquels on peut citer les formes différentielles, et un retour en arrière sur des sujets déjà étudiés, comme la topologie générale et les différentielles.

S’il y avait un thème central dans ce cours ce serait celui du théorème d’inversion locale. Pour autant ce n’est pas un sujet entièrement nouveau, puisque l’énoncé est déjà donné, avec celui des fonctions implicites, au premier semestre. Cependant le second semestre donne l’occasion d’en donner des applications, la plus évidente étant la réduction des immersions et submersions locales et l’équivalence entre diverses définitions locales des sous-variétés. On s’en servira aussi pour interpréter les longueurs et les aires en les ramenant à des volumes. Enfin on pourra encore en faire usage, dans le cadre des espaces de Banach, pour étudier le flot d’une équation différentielle.

Hormis les thèmes de l’inversion locale, de l’existence et de l’unicité des solutions d’une équation différentielle et des formes différentielles, c’est d’avantage une fonction de synthèse et de consolidation d’acquis supposés qu’on a donné au module.

On commence ainsi par un chapitre sur les applications différentiables qui n’est qu’une révision du cours du premier semestre. On s’est placé dans le cadre des espaces normés en pensant aux équations différentielles, mais en insistant sur la dimension finie, voire les dimensions 1, 2 et 3.

On a encore repris le cas des équations linéaires à coefficients constants, étudié en L2 avec la réduction des endomorphismes, dans un esprit algébrique. C’est une occasion de revisiter quelques thèmes de cette seconde année : algèbres, théorème chinois, réduction etc.

Les rudiments de topologie générale ont été revisités, après les premiers chapitres de calcul différentiel. On a fait l’inventaire de ce dont se servait en permanence. Plutôt que de redonner des définitions, on a préféré insister sur le côté pratique, le “comment fait-on?”. D’ailleurs on ne devrait pas poser en licence d’exercice portant uniquement

sur la topologie générale; cette dernière devrait tout le temps être sollicitée avec le calcul différentiel.

Même à propos des formes différentielles, on a repris au début l'algèbre extérieure des premières années qui a conduit à l'introduction du déterminant, en restant dans le cadre des formes  $p$ -linéaires alternées.

On a encore fait la "révision" de quelques sujets dont on n'a jamais parlé — mais dont on aurait dû le faire. Il faut dire que dans les deux modules de "Calcul différentiel et intégral", l'usage est de ne pas faire le moindre soupçon de calcul intégral. On trouvera donc un petit chapitre où on dit en trois pages l'essentiel de ce qu'il faut savoir sur l'intégrale de Lebesgue. On peut d'ailleurs résumer en une ligne : savoir se ramener à des fonctions positives et connaître le théorème de Fubini.

Dans l'attente de remarques et commentaires

Jean-Pierre Ferrier, décembre 2008

## Chapitre 1. Applications différentiables

### 1. Définition.

Une application différentiable est, en chaque point, proche d'une application linéaire; c'est la traduction de la linéarité des petits accroissements en physique, qu'on énonce souvent comme le principe d'addition des petits mouvements.

Soit une application  $U \xrightarrow{f} V$ , où  $U, V$  sont des parties ouvertes dans des espaces normés  $E$  et  $F$ ; soit encore un point  $a \in U$ .

On dit que  $f$  est différentiable en  $a$  si l'on peut trouver une application linéaire continue  $l$  telle que

$$f(a+h) = f(a) + l(h) + o(\|h\|)$$

où  $o(\|h\|) = \|h\|o(1)$  et où  $o(1) \rightarrow 0$  quand  $h \rightarrow 0$ .

Quand  $E = \mathbf{R}$  — autrement dit dans le cas d'une seule variable numérique — les applications linéaires sont de la forme  $h \mapsto h \cdot \vec{l}$  et la différentiabilité s'écrit

$$f(a+h) = f(a) + h \cdot \vec{l} + o(|h|) ;$$

on parle de *dérivabilité*; ici  $\vec{l}$  est le vecteur *dérivée*.

Dans le cas général du début, on demande à l'application linéaire  $l$  d'être continue. Cela signifie que

$$\|l(h)\| \leq C\|h\|$$

pour une constante  $C \geq 0$ ; la plus petite de ces constantes est la norme  $\|l\|$  de  $l$ ; ainsi

$$\|l(h)\| \leq \|l\| \cdot \|h\| .$$

**Théorème.** si  $E$  est de dimension finie, alors  $l$  est automatiquement continue.

**Exercice.** Une application différentiable en  $a$  est continue en  $a$ .

### 2. Notations, exemples.

**Exercice.** une telle application  $l$  est unique.

On dit que  $l$  est la *différentielle* de  $f$  au point  $a$ . On la note  $T_a f$  ou  $df(a)$  ou encore  $f'(a)$ ,  $df_a$ ,  $Df_a$  etc.

Son effet sur le vecteur  $h$  est noté  $df(a)(h)$  ou  $(df(a))(h)$  ou plutôt  $df(a).h$  pour l'esthétique.

Si  $f$  est différentiable en tout point de  $U$ , alors  $df$  définit une application de  $U$  dans  $L(E, F)$  — espace normé des applications linéaires continues de  $E$  dans  $F$ .

Voici quelques exemples.

- . Si  $f$  est constante, alors  $df = 0$ .
- . Si  $f$  est linéaire, alors  $df(a) = f$  en tout point  $a$ .

. Si  $E = \mathbf{R}^2$  et si  $e_1, e_2$  composent sa base canonique, alors

$$\begin{aligned} df(a).h &= df(a).(h_1e_1 + h_2e_2) \\ &= h_1df(a).e_1 + h_2df(a).e_2 \\ &= h_1\frac{\partial f}{\partial x_1}(a) + h_2\frac{\partial f}{\partial x_2}(a) \end{aligned}$$

où nous verrons plus loin comment interpréter les vecteurs  $\frac{\partial f}{\partial x_1}(a)$  et  $\frac{\partial f}{\partial x_2}(a)$  en termes de dérivées partielles. Si on note  $x_1, x_2$  les projections de  $\mathbf{R}^2$  sur ses facteurs, qui sont linéaires, on a  $dx_1(a) = x_1$ , i.e.  $dx_1(a).h = h_1 \dots$  Ainsi

$$df = \frac{\partial f}{\partial x_1}dx_1 + \frac{\partial f}{\partial x_2}dx_2 .$$

Si, toujours à titre d'exemple,  $E = \mathbf{R}^2$  et  $F = \mathbf{R}^3$ , et si l'on note  $f_1, f_2, f_3$  les composantes de  $f$ , alors la différentielle s'identifie à la matrice

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \\ \frac{\partial f_3}{\partial x_1} & \frac{\partial f_3}{\partial x_2} \end{pmatrix}$$

dite **matrice jacobienne**. On laisse le lecteur envisager le cas d'une application de  $E = \mathbf{R}^m$  dans  $F = \mathbf{R}^p$ .

### 3. Composition des applications différentiables.

**Théorème.** Soient des applications  $U \xrightarrow{f} V \xrightarrow{g} W$ , où  $U, V, W$  sont des parties ouvertes d'espaces normés, un point  $a \in U$  et  $b = f(a)$ .

Si  $f$  est différentiable en  $a$  et si  $g$  est différentiable en  $b$ ; alors  $g \circ f$  est différentiable en  $a$  et

$$d(g \circ f)(a) = dg(b) \circ df(a) .$$

Les calculs sont faciles mais tiennent un peu de place; on préfère détailler le point qui suit.

**Exemple.** Si  $f$  est différentiable en  $a$  et si

$$\phi(t) = f(a + th)$$

alors  $\phi$  est dérivable en 0 et

$$\phi'(0) = df(a).h .$$

En effet  $\phi$  est la composée des applications  $t \mapsto a + th$  et de  $f$ .

Sa différentielle est la composée de  $k \mapsto kh$  et de  $df(a+th)$ ; c'est  $k \mapsto df(a+th).kh = kdf(a+th).h$ .

Revenons sur les vecteurs  $\frac{\partial f}{\partial x_i}$ , en prenant le cas de  $E = \mathbf{R}^2$  et  $h = e_1$ . On voit que  $df(a).e_1$  est la dérivée en 0 de  $f(a + te_1) = f(a_1 + t, a_2)$ , ce qui est bien la définition usuelle de la dérivée partielle par rapport à  $x_1$ .

#### 4. Inversion des applications différentiables.

**Théorème.** Soient une application  $U \xrightarrow{f} V$ , où  $U, V$  sont des parties ouvertes d'espaces normés, un point  $a \in U$  et  $b = f(a)$ .

On suppose que  $f$  est bijective et que  $f, f^{-1}$  sont continues (on dit que c'est un homéomorphisme).

Si  $f$  est différentiable en  $a$  et si  $df(a)$  est inversible comme application linéaire continue, alors  $f^{-1}$  est différentiable en  $b$  et

$$df^{-1}(b) = (df(a))^{-1} .$$

Pour la démonstration on part de

$$f(a+h) = f(a) + df(a).h + o(\|h\|) .$$

Si  $k = f(a+h) - f(a)$  est l'accroissement de  $f$  correspondant à  $h$ , on écrit

$$k = df(a).h + o(\|h\|) .$$

On applique  $df(a)^{-1}$ , ce qui donne

$$df(a)^{-1}.k = h + df(a)^{-1}.o(\|h\|) .$$

Comme  $\|df(a)^{-1}.o(\|h\|)\| \leq \|df(a)^{-1}\|.o(\|h\|)$ , il vient

$$h = df(a)^{-1}.k + o(\|h\|) .$$

Il reste deux ajustements à faire. D'abord remplacer  $o(\|h\|)$  par  $o(\|k\|)$  : c'est l'objet d'un petit **exercice**. Puis  $h \rightarrow 0$  par  $k \rightarrow 0$ , en montrant que  $k \rightarrow 0$  implique  $h \rightarrow 0$  : c'est dire que  $f^{-1}$  est continue en  $b$ , comme il est demandé dans l'hypothèse.

#### 5. Accroissements finis.

a) **Théorème.** Soit une application  $[a, b] \xrightarrow{f} V$ , où  $V$  est une partie ouverte d'un espace normé  $E$ .

\* (forme de Lagrange) Si  $f$  est continue sur  $[a, b]$ , si elle est dérivable sur  $]a, b[$  et si

$$\|f'(t)\| \leq M$$

sur cet intervalle, alors

$$\|f(b) - f(a)\| \leq M|b - a| .$$

\* (théorème fondamental de l'Analyse) Si  $f$  est continûment dérivable (on dit de classe  $\mathcal{C}^1$ ) sur  $[a, b]$  et  $E$  complet, alors

$$f(b) - f(a) = \int_a^b f'(t)dt .$$

**N.B.** La formule avec  $f'(c)$  au second membre est fautive pour une fonction à valeurs complexes ou vectorielles : penser à l'application  $t \mapsto e^{it}$  sur  $[0, 1]$ .

b) **Théorème.** Soit une application  $U \xrightarrow{f} V$ , où  $U, V$  sont des parties ouvertes d'espaces normés  $E, F$ , et soit un segment  $[a, a+h]$  inclus dans  $U$ .

\* Si  $f$  est différentiable dans  $U$  et si

$$\|df(x)\| \leq M$$

dans ce domaine, alors

$$\|f(a+h) - f(a)\| \leq M\|h\| .$$

\* Si  $f$  est continûment différentiable (on dit de classe  $\mathcal{C}^1$ ) dans  $U$  et  $F$  complet, alors

$$f(a+h) - f(a) = \left( \int_0^1 df(a+th)dt \right) . h .$$

**N.B.** Pour le b), on applique le a) à la fonction  $\phi(t) = f(a+th)$  sur  $[0, 1]$ ; en fait il suffit de supposer la continuité et la borne sur  $[a, b]$ .

Maintenant la démonstration de a) est sérieuse. Il faut au moins démontrer que  $f' = 0$  implique  $f(b) = f(a)$ . Cela peut se faire en considérant les  $x$  pour lesquels une certaine propriété a lieu sur l'intervalle  $[a, x]$ , puis en prenant la borne supérieure desdits  $x$ ; à faire en **exercice**.

Noter qu'ici *fini* s'oppose à *infinitésimal*.

## 6. Caractérisation de la classe $\mathcal{C}^1$ .

**Théorème.** Soit une application  $f : U \rightarrow V$ , où  $U$  est une partie de  $\mathbf{R}^n$  — on est en dimension finie — et  $V$  une partie ouverte d'un espace normé  $F$ . Les propriétés suivantes sont équivalentes.

(a)  $f$  est continûment différentiable (on dit de classe  $\mathcal{C}^1$ ) sur  $U$ .

(b)  $f$  admet des dérivées partielles continues sur  $U$ .

Pour la démonstration, on applique le théorème des accroissements finis : voir le cours de CDI1.

## 7. Formule de Taylor.

Si  $df$  est différentiable en  $a$  comme application de  $U$  dans  $L(E, F)$ , on dit que  $f$  est deux fois différentiable en  $a$ . Alors

$$d(df)(a)$$

est un élément de  $L(E, L(E, F))$  qu'on note  $d^2f(a)$ .

Si  $k$  est un élément de  $E$ , alors  $d^2f(a).k$  est un élément de  $L(E, F)$ .

Si  $h$  est aussi un élément de  $E$ , alors  $(d^2f(a).k).h$  est un élément de  $F$ .

**Théorème** (Schwarz). L'application bilinéaire  $d^2f(a)$  est symétrique.

On écrit  $d^2f(a).(k, h)$  au lieu de  $(d^2f(a).k).h$ .

En particulier  $d^2f(a).(h)^2$  est  $d^2f(a).h, h$ .

**Théorème.** Soit une application  $U \xrightarrow{f} V$ , où  $U, V$  sont des parties ouvertes d'espaces normés  $E, F$ , et un segment  $[a, a + h]$  inclus dans  $U$ .

\* (forme de Lagrange) Si  $f$  est  $n + 1$  fois différentiable dans  $U$  et si

$$\|d^{n+1}f(x)\| \leq M$$

dans ce domaine, alors

$$f(a + h) = f(a) + \frac{1}{1!}df(a).h + \cdots + \frac{1}{n!}d^n f(a).(h)^n + r_{n,a}(h)$$

où

$$\|r_{n,a}(h)\| \leq \frac{M}{(n+1)!}\|h\|^{n+1} .$$

\* (reste intégral) Si est de classe  $\mathcal{C}^{n+1}$  dans  $U$  et  $F$  complet, on a la même relation où

$$r_{n,a}(h) = \left( \int_0^1 \frac{(1-t)^n}{n!} d^{n+1}f(a+th)dt \right) .(h)^{n+1} .$$

Pour la démonstration, on se ramène au théorème des accroissements finis en posant  $b = a + h$  et en établissant ce qui suit.

**Lemme (formule de Taylor).** Si

$$\Phi(x) = f(x) + \frac{1}{1!}df(x).(b-x) + \frac{1}{n!}d^n f(x).(b-x)^n$$

alors  $r_{n,a}(h) = \Phi(b) - \Phi(a)$  et

$$d\Phi(x) = \frac{1}{n!}d^{n+1}f(x).(b-x)^n .$$

**N.B.** Quand on suppose seulement que  $f$  est  $n$  fois différentiable en  $a$ , on a la forme de Taylor-Young qui dit que  $r_{n,a}(h) = o(\|h\|^n)$ . La démonstration repose aussi sur le lemme indiqué.

## 8. Un exemple.

Nous allons préciser  $d^2f$  pour une application  $f$  des variables  $x_1, x_2$ . On a

$$df(x_1, x_2).h = h_1 \frac{\partial f}{\partial x_1}(x_1, x_2) + h_2 \frac{\partial f}{\partial x_2}(x_1, x_2) .$$

En différentiant en  $x_1, x_2$ , il vient

$$\begin{aligned} d^2f(x_1, x_2).k.h &= h_1 \left( k_1 \frac{\partial^2 f}{\partial x_1^2}(x_1, x_2) + k_2 \frac{\partial^2 f}{\partial x_2 \partial x_1}(x_1, x_2) \right) \\ &+ h_2 \left( k_1 \frac{\partial^2 f}{\partial x_1 \partial x_2}(x_1, x_2) + k_2 \frac{\partial^2 f}{\partial x_2^2}(x_1, x_2) \right) . \end{aligned}$$

Autrement dit

$$d^2 f(x_1, x_2).(k.h) = h_1 k_1 \frac{\partial^2 f}{\partial x_1^2} + h_1 k_2 \frac{\partial^2 f}{\partial x_2 \partial x_1} + h_2 k_1 \frac{\partial^2 f}{\partial x_1 \partial x_2} + h_2 k_2 \frac{\partial^2 f}{\partial x_2^2}$$

et notamment

$$d^2 f(x_1, x_2).(h)^2 = h_1^2 \frac{\partial^2 f}{\partial x_1^2} + h_1 h_2 \left( \frac{\partial^2 f}{\partial x_2 \partial x_1} + \frac{\partial^2 f}{\partial x_1 \partial x_2} \right) + h_2^2 \frac{\partial^2 f}{\partial x_2^2}$$

où la parenthèse de droite vaut

$$2 \frac{\partial^2 f}{\partial x_1 \partial x_2}$$

par Schwartz.

## 9. Problèmes d'extrema.

Ici  $f$  est à valeur réelles. On dit que  $a$  est un *maximum local* (resp. *minimum local*) si  $f(a+h) \leq f(a)$  (resp.  $f(a+h) \geq f(a)$ ) pour  $h$  dans un voisinage de 0; dans l'un et l'autre cas on parle d'*extremum local*.

**Théorème (condition nécessaire d'extremum local).** Si  $f$  est définie au voisinage de  $a$  et différentiable en  $a$ , une condition nécessaire d'extremum local est que  $df(a) = 0$ .

Ici  $df(a)$  est de rang maximum 1; un point où il ne l'est pas, donc ici  $df(a) = 0$ , est dit critique; si  $E = \mathbf{R}^n$ , cela signifie que les dérivées partielles y sont toutes nulles.

**Théorème (condition nécessaire de maximum local).** Si  $f$  est définie au voisinage de  $a$  et deux fois différentiable en  $a$ , une condition nécessaire de maximum local est que  $df(a) = 0$  et que la forme quadratique  $d^2 f(a)$  soit négative, i.e. que  $d^2 f(a).(h, h) \leq 0$ .

Dans les deux cas, la démonstration peut se ramener au cas d'une variable. On pose  $\phi(t) = f(a+th)$  et on note que  $\phi'(t) = df(a+th).h$ , d'où  $\phi''(t) = (d^2 f(a+th).h).h = d^2 f(a+th).(h)^2$ ; ainsi  $\phi'(0) = df(a).h$  et  $\phi''(0) = d^2 f(a).(h)^2$ .

**Théorème (condition suffisante de maximum local).** Si  $f$  est définie au voisinage de  $a$  et deux fois différentiable en  $a$ , une condition suffisante de maximum local est que  $df(a) = 0$  et que la forme quadratique  $d^2 f(a)$  soit définie négative, i.e. que l'on ait

$$d^2 f(a).(h, h) \leq -c\|h\|^2$$

pour une constante  $c > 0$ .

**Théorème.** En dimension finie, pour qu'une forme quadratique soit définie négative il suffit qu'elle soit négative et non nulle en tout  $h \neq 0$ .

Ici la démonstration repose sur le théorème de Taylor-Young et le fait que  $-c\|h\|^2 + o(\|h\|^2) \leq 0$  pour  $\|h\|$  assez petit. En fait on a même un maximum local strict, à savoir  $f(a+h) \leq f(a)$  pour  $h \neq 0$  assez petit.

Reprenons l'exemple d'une fonction de deux variables. On avait obtenu

$$d^2 f(x_1, x_2).(h)^2 = h_1^2 \frac{\partial^2 f}{\partial x_1^2} + 2h_1 h_2 \frac{\partial^2 f}{\partial x_1 \partial x_2} + h_2^2 \frac{\partial^2 f}{\partial x_2^2}.$$

Cette forme sera négative si

$$\left(\frac{\partial^2 f}{\partial x_1 \partial x_2}\right)^2 - \frac{\partial^2 f}{\partial x_1^2} \frac{\partial^2 f}{\partial x_2^2} \leq 0$$

et, par exemple

$$\frac{\partial^2 f}{\partial x_1^2} \leq 0 .$$

Elle sera définie négative avec une première inégalité qui est stricte.

## 10. Suites et intégrales de fonctions différentiables.

**Théorème.** Soit  $f_n$  une suite de fonctions différentiables sur une boule ouverte bornée  $B$  d'un espace normé  $E$  et à valeurs dans un espace normé complet  $F$ . On suppose

(a) que la suite  $f_n$  converge en un point  $a$  de  $B$ .

(b) que la suite  $df_n$  converge *uniformément* dans  $B$  vers une forme  $g$ .

Alors la suite  $f_n$  converge uniformément dans  $B$  vers une fonction  $f$ , laquelle est différentiable et vérifie  $df = g$ .

La démonstration est plus simple pour des  $f_n$  de classe  $\mathcal{C}^1$ ; nous le supposons.

Si  $x = a + h$  est un point de  $B$ , on écrit la formule intégrale des accroissements finis

$$f_n(a + h) = f_n(a) + \int_0^1 df_n(a + th).h dt ,$$

où le second membre converge uniformément vers

$$f(a) + \int_0^1 g(a + th).h dt .$$

Notons  $f(x)$  la limite. On tire de la dernière expression le reste des conclusions.

**Théorème.** Soit  $U$  une partie ouverte d'un espace normé  $E$  et  $f = f(t, x)$  une application de  $U \times [a, b]$  dans un espace normé complet  $F$ .

On suppose que  $f$  est continue et admet une différentielle partielle  $d_x f$  en  $x$  continue dans  $[a, b] \times U$ .

Alors

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

est continûment différentiable dans  $U$ , de différentielle

$$dF(x) = \int_a^b d_x f(t, x) dt .$$

C'est une application directe de la formule intégrale des accroissements finis. En effet; si  $x = a + h$ , l'expression

$$F(a + h) - F(a) - \int_a^b d_x f(t, a).h dt = \int_a^b \left( \int_0^1 (d_x f(t, a + uh).h - d_x f(t, a).h) du \right) dt$$

se majore en un  $o(\|h\|)$ .

## Chapitre 2. Inversion locale et applications

### 1. Inversion locale.

**Théorème.** Soit  $f$  une application est de classe  $\mathcal{C}^1$  au voisinage du point  $a$  de  $E$  et à valeurs dans  $F$ , où  $E, F$  sont des espaces normés complets; on pose  $b = f(a)$ .

Si la différentielle  $df(a)$  est inversible, comme application linéaire continue entre espaces normés, alors on peut trouver des voisinages ouverts  $U$  de  $a$  et  $b$  de  $V$  tels que  $f$  induise une bijection de  $U$  sur  $V$  dont l'inverse soit de classe  $\mathcal{C}^1$  (on dit que  $f$  induit un  $\mathcal{C}^1$ -difféomorphisme de  $U$  sur  $V$ ).

Voici l'idée de la démonstration. Par translation on se ramène au cas où  $a = b = 0$ , puis en appliquant  $df(a)^{-1}$  au cas où  $df(a)$  est l'application identique  $1_E$  de  $E$ .

On construit d'abord une **bijection**. Par la continuité de  $df$ , il vient

$$\|1_E - df(x)\| \leq \frac{1}{2}$$

pour  $x$  dans une boule fermée  $B' = \{\|x\| \leq r\}$  où  $r > 0$ . Par le théorème des accroissements finis, il en résulte que l'application  $g_y$  définie par

$$g_y(x) = x + y - f(x)$$

est contractante dans le rapport  $1/2$  dans  $B'$  :

$$\|g_y(x) - g_y(x')\| \leq \frac{1}{2}\|x - x'\| ;$$

de plus, pour  $\|y\| \leq r/2$ , on vérifie que cette application envoie  $B'$  dans elle-même.

Le théorème du point fixe assure que l'on peut, pour  $\{\|y\| \leq r/2\}$ , résoudre  $g_y(x) = x$  et donc  $y = f(x)$  en  $x$  dans la boule  $B'$  de façon unique. On en déduit que l'on peut aussi, pour  $y$  dans la boule ouverte  $V = \{\|y\| < r/2\}$ , résoudre  $y = f(x)$  en  $x$  dans la boule ouverte  $U_1 = \{\|x\| < r\}$  de façon unique. Ainsi  $f$  induit-elle une bijection de  $U = U_1 \cap f^{-1}(V)$  sur  $V$ .

Maintenant on montre que c'est un **homéomorphisme**. la continuité de l'inverse résulte de l'étude de  $g_0$  :

$$\|f(x) - f(x')\| = \|x - g_0(x) - x' + g_0(x')\| \geq \|x - x'\| - \|g_0(x) - g_0(x')\| \geq \|x - x'\|/2 .$$

Enfin c'est un  **$\mathcal{C}^1$ -difféomorphisme**. On applique le théorème d'inversion des différentielles, sachant que  $df(x)$  est inversible dans  $B'$  et en particulier dans  $U$ . En effet  $u = 1_E - df(x)$  vérifie  $\|u\| \leq 1/2$  et  $\|u^n\| \leq \|u\|^n \leq 1/2^n$ ; on inverse  $df(x) = 1_E - u$  par la somme de la série

$$1_E + u + u^2 + \cdots + u^n + \cdots$$

qui converge absolument en norme d'opérateur.

Si  $f$  est, en plus, de classe  $\mathcal{C}^k$ , alors son inverse local est de la même classe. C'est facile à vérifier en dimension finie à l'aide des dérivées partielles. Pour le cas général on renvoie au livre d'Henri Cartan.

Dans la suite on se placera en *dimension finie* et toutes les applications différentiables seront supposées de classe  $\mathcal{C}^k$  pour un certain  $k$  compris entre 1 et  $+\infty$ , sans que cela soit rappelé.

## 2. Coordonnées locales.

Etant donné un point  $a$  d'un espace vectoriel de dimension  $n$ , un système de coordonnées locales en  $a$  est simplement un difféomorphisme local à valeurs dans  $\mathbf{R}^n$  en  $a$  et 0; les composantes constituent les coordonnées locales.

Cependant, de manière un peu abusive, on parlera aussi de coordonnées locales lorsque l'espace d'arrivée est un espace vectoriel de dimension finie; on laisse au lecteur le soin d'y considérer une base et de composer avec les coordonnées suivant cette base.

Dans les deux sections qui suivent, on considèrera deux espaces vectoriels  $E, F$  de dimensions respectives  $m$  et  $n$ , une application différentiable  $f$  définie au voisinage d'un point  $a$  de  $E$  et à valeurs dans  $F$ , pour laquelle on pose  $b = f(a)$ . On désignera par  $x, y$  des systèmes de coordonnées linéaires sur  $E, F$  les identifiant à  $\mathbf{R}^m, \mathbf{R}^n$ .

## 3. Réduction locale des submersions.

**Théorème.** Les propriétés suivantes, qui imposent  $m \geq n$ , sont équivalentes; elles caractérisent les *submersions* locales en  $a$  et  $b$ .

- (i) L'application linéaire tangente  $df(a)$  est surjective; autrement dit elle est de rang  $n$ .
- (ii) On peut trouver localement une factorisation

$$\begin{array}{ccc} E & \xrightarrow{f} & F \\ \Phi \downarrow & \nearrow p & \\ G \times F & & \end{array}$$

où  $G$  est un espace vectoriel de dimension  $m - n$ , où  $\Phi$  est un difféomorphisme local en  $a$  et  $(0, b)$  et où  $p$  est la projection naturelle  $G \times F \rightarrow F$ .

- (ii') Dans (ii), on peut prendre une application linéaire surjective  $q : E \rightarrow G$  donnée de noyau supplémentaire de celui de  $df(a)$  et définir  $\Phi$  par  $\Phi(x) = (q(x), f(x))$ .
- (iii) On peut trouver un système  $(u, v)$  de coordonnées locales sur  $E$  tel que  $f$  soit décrite localement par l'équation

$$y = v .$$

Le passage de (ii) à (i) est évident. Celui de (i) à (ii'), en particulier à (ii), repose simplement sur le théorème d'inversion locale. Enfin (iii) n'est qu'une réécriture de (ii); en effet le système  $(u, v)$  de coordonnées locales est essentiellement  $\Phi$ ; en particulier  $v = f$ .

**Exercice.** Expliciter les cas  $\mathbf{R}^2 \rightarrow \mathbf{R}, \mathbf{R}^3 \rightarrow \mathbf{R}, \mathbf{R}^3 \rightarrow \mathbf{R}^2$ . On se placera à l'origine, on écrira  $f$  sous la forme d'équation(s) scalaire(s) et on répondra aux questions suivantes.

- Que signifie (i)?
- On se donne un cas où (i) est valide. Que convient-il d'ajouter pour avoir un difféomorphisme local en 0?
- Que suffit-il de choisir pour cela?
- Expliciter des coordonnées locales et des équations dans ces coordonnées.

Une situation remarquable de submersion locale correspond à ce qui suit.

**Théorème des fonctions implicites.** Soit  $f$  une application différentiable définie au voisinage de  $(a, b)$  dans  $E \times F$  et à valeurs dans  $G$ , vérifiant  $f(a, b) = 0$ . On suppose que la différentielle partielle  $\partial f / \partial y(a, b)$  par rapport au second facteur est inversible.

Alors localement en  $a$  et  $b$ , on peut trouver une application différentiable  $g$  de  $E$  dans  $F$  telle que l'équation  $f(x, y) = 0$  soit équivalente à  $y = g(x)$ .

Autrement dit on exprimer localement  $y$  comme une fonction de  $x$ .

C'est un cas particulier de la situation (ii') ci-dessus. On prend pour  $q$  la projection  $E \times F \rightarrow E$ , sachant que le noyau de  $df(a)$  ne rencontre pas  $0 \times F$ . La factorisation s'écrit

$$\begin{array}{ccc} E \times F & \xrightarrow{f} & G \\ \Phi \downarrow & \nearrow p & \\ E \times G & & \end{array}$$

et pour obtenir  $g(x)$  il suffit de prendre la seconde projection de  $\Phi^{-1}(x, 0)$ .

#### 4. Réduction locale des immersions.

**Théorème.** Les propriétés suivantes, qui imposent  $m \leq n$ , sont équivalentes; elles caractérisent les *immersions* locales en  $a$  et  $b$ .

- (i) L'application linéaire tangente  $df(a)$  est injective; autrement dit elle est de rang  $m$ .
- (ii) On peut trouver localement une factorisation

$$\begin{array}{ccc} E & \xrightarrow{f} & F \\ i \searrow & & \downarrow \Phi \\ & & E \times G \end{array}$$

où  $G$  est un espace vectoriel de dimension  $n - m$ , où  $\Phi$  est un difféomorphisme local en  $b$  et  $(a, 0)$  et où  $i$  est l'injection naturelle  $E \rightarrow E \times G$ .

- (ii') Dans (ii), on peut prendre une application linéaire injective  $j : G \rightarrow F$  donnée dont l'image est supplémentaire de celle de  $df(a)$  et définir  $\Phi$  par  $\Phi^{-1}(x, z) = f(x) + j(z)$ .
- (iii) On peut trouver un système  $(u, v)$  de coordonnées locales sur  $F$  tel que  $f$  soit décrite localement par les équations

$$\begin{cases} u = x \\ v = 0 \end{cases} .$$

Le passage de (ii) à (i). Celui de (i) à (ii'), et en particulier à (ii), repose simplement sur le théorème d'inversion locale. Enfin la formulation (iii) est une réécriture de (ii) : le système de coordonnées locales est essentiellement  $\Phi$ ; en particulier  $u \circ f = x$  sur l'image de  $E$ .

Précisément, dans la construction (ii'), les coordonnées locales  $u, v$  prennent au point  $f(x) + j(z)$  les valeurs respectives  $x$  et  $z$ .

**Exercice.** Expliciter les cas  $\mathbf{R} \rightarrow \mathbf{R}^2$ ,  $\mathbf{R} \rightarrow \mathbf{R}^3$ ,  $\mathbf{R}^2 \rightarrow \mathbf{R}^3$ . On se placera à l'origine, on écrira  $f$  sous la forme d'équation(s) scalaire(s) et on répondra aux questions suivantes.

- Que signifie (i)?
- On se donne un cas où (i) est valide. Que convient-il d'ajouter pour avoir un difféomorphisme local en 0?
- Que suffit-il de choisir pour cela?
- Expliciter des coordonnées locales et des équations dans ces coordonnées.

**N.B.** On trouverait un exposé bien plus satisfaisant, où le cadre est celui des variétés différentielles, dans l'ouvrage *Lie Algebras and Lie Groups* de Jean-Pierre Serre. D'autres caractérisations sont données : on peut se ramener à une surjection (resp. injection) linéaire par des difféomorphismes locaux au départ et à l'arrivée, ou bien l'inversibilité locale à droite (resp. à gauche).

## 5. Sous-variétés.

**Définition.** Etant donné un espace vectoriel  $E$  de dimension  $n$ , on dit qu'une partie  $V$  de  $E$  est une *sous-variété* de dimension  $m$  de  $E$  si, par un choix convenable de coordonnées locales, on peut la ramener en chacun de ses points au sous-espace  $\mathbf{R}^m \times 0$  de  $\mathbf{R}^n$ .

Autrement dit, on peut trouver en chacun de ses points un système de coordonnées locales  $(x, y)$ , où  $x = (x_1, \dots, x_m)$  et  $y = (x_{m+1}, \dots, x_n)$  tel que  $V$  soit définie localement par  $y = 0$ .

La définition s'étend au cas où l'on remplace  $E$  par une partie ouverte  $U$  de ce dernier.

**Exemple 1.** Une application différentiable  $f$  de la partie ouverte  $U$  de  $E$  dans  $F$  est dite une *submersion* si c'en est une en chaque point de  $U$ . La partie définie par  $f(x) = 0$  est alors une sous-variété fermée de  $U$  de dimension  $\dim E - \dim F$ . On dit que  $f$  en est un système d'équations cartésiennes ou implicites.

C'est simplement la réduction locale des submersions qui le donne : on se ramène localement à une projection  $(u, v) \mapsto v$ .

**Exemple 2.** Une application différentiable injective propre de la partie ouverte  $V$  de  $F$  dans la partie ouverte  $U$  de  $E$  est dite un *plongement* si c'est une immersion en chaque point. Son image, qui est fermée dans  $U$ , est alors une sous-variété de dimension  $\dim F$ . On dit que  $f$  en est un système d'équations paramétriques.

Cette fois-ci c'est la réduction locale des immersions qui le donne : on se ramène localement à une injection  $u \mapsto (u, 0)$ .

## 6. Espace tangent.

Soient  $V$  une sous-variété de dimension  $m$  de  $U$  et  $a$  un point de  $V$ . Les vecteurs de la forme

$$\left. \frac{d}{dt} \gamma(t) \right|_{t=0},$$

où  $\gamma$  est une application dérivable d'un intervalle ouvert contenant 0 dans  $V$  telle que  $\gamma(0) = a$ , sont dits *vecteurs tangents* à  $V$  au point  $a$ .

**Théorème.** Les vecteurs tangents à  $V$  en  $a$  constituent un espace vectoriel de dimension  $\dim V$ , appelé *espace tangent* en  $a$ . Si  $x : V \rightarrow \mathbf{R}^n$  est un système de coordonnées locales en  $a$  envoyant  $V$  sur  $\mathbf{R}^m \times 0$ , c'est l'espace  $dx^{-1}(0)(\mathbf{R}^m \times 0)$ .

On fait la démonstration pour la sous-variété  $\mathbf{R}^m \times 0$  de  $\mathbf{R}^n$  : l'espace tangent est alors  $\mathbf{R}^m \times 0$ . Le résultat s'étend par le principe de composition des différentielles. Ce qui suit est également évident.

**Exemple 1.** Dans le cas d'une variété définie par une équation implicite  $f$ , le plan tangent en  $a$  est le noyau de  $df(a)$ .

**Exemple 2.** Dans le cas d'une variété définie par une équation paramétrique  $f$ , le plan tangent en  $a$  est l'image de  $df(a)$ .

## 7. Extrema liés.

On se donne une fonction différentiable  $f$  à valeurs réelles sur une variété  $V$  définie par des équations implicites  $g_1 = \dots = g_p = 0$ .

**Théorème.** Une condition nécessaire pour que le point  $a$  de  $V$  soit un extremum local de  $f$  sur  $V$  est qu'il existe des scalaires  $\lambda_1, \dots, \lambda_p$  (dits *multiplicateurs de Lagrange*) vérifiant

$$df(a) + \lambda_1 dg_1(a) + \dots + \lambda_p dg_p(a) = 0 .$$

Pour la démonstration, on se ramène simplement au cas où chaque fonction  $g_i$  est une coordonnée  $x_i$ .

**N.B.** On vérifiera bien sûr pour  $(g_1, \dots, g_p)$  la condition exigée d'une submersion, à savoir que son rang en chaque point est  $p$ .

## Chapitre 3. Equations différentielles ordinaires

### 1. Le cadre.

On se donne un espace normé  $E$  complet (par exemple de dimension finie), une partie  $U$  de  $\mathbf{R} \times E$  ouverte (sauf mention contraire) et une application  $f$  de  $U$  dans  $E$  continue (avec d'autres propriétés s'il le faut).

Une **solution sur l'intervalle  $I$  de l'équation différentielle**

$$\frac{dx}{dt} = f(t, x)$$

est une fonction  $x$  de  $I$  dans  $E$  qui est dérivable, pour laquelle  $(t, x(t))$  reste dans  $U$  et qui vérifie sur  $I$  la propriété indiquée. On parle aussi de *trajectoire* ou de *courbe intégrale*.

Précisément, étant donné un intervalle  $I$ , on désigne par  $V_I$  la partie de l'espace normé  $\mathcal{C}(I, E)$ , équipé de la norme uniforme, qui est composée des fonctions continues  $x$  sur  $I$  à valeurs dans  $E$  telles que  $\delta_x(t) = (t, x(t))$  reste dans  $U$  pour  $t$  dans  $I$ .

Une solution sur  $I$  est alors une fonction  $x$  de  $V_I$  qui est dérivable et qui vérifie sur  $I$  la propriété indiquée.

Si  $I$  est un segment, on notera que  $V_I$  est une partie *ouverte* de  $\mathcal{C}(I, E)$ ; en effet, pour  $x$  dans  $V_I$ , l'image de  $I$  par  $\delta_x$  est compacte dans  $U$  et sa distance au bord de  $U$  est minorée par un nombre  $c > 0$ . Si l'application  $y$  de  $\mathcal{C}(I, E)$  vérifie alors  $\|x - y\|_\infty < c$ , il apparaît que  $\|\delta_x(t) - \delta_y(t)\| < \epsilon$  pour tout  $t$  dans  $I$ , ce qui implique que  $\delta_y(t)$  soit dans  $U$  pour tout  $t$  dans  $I$ , et donc que  $y$  est dans  $V_I$ .

On retiendra surtout que  $f$ ,  $t_0$  et  $x_0$  sont les données et que  $x$  est l'inconnue, qui est ici une fonction.

### 2. Condition initiale.

On se donne encore un point  $(t_0, x_0)$  de  $U$  et on s'intéresse aux solutions sur un intervalle  $I$  contenant  $t_0$  telles que  $x(t_0) = x_0$ .

Pour qu'une fonction  $x$  de  $V_I$  soit une telle solution, il faut et il suffit que  $x$  vérifie **l'équation intégrale**

$$x(t) = x_0 + \int_{t_0}^t f(u, x(u)) du$$

sur  $U$ .

La vérification est immédiate : si  $x$  vérifie l'équation intégrale, alors  $x(t_0) = x_0$  et on peut dériver l'intégrale par rapport à sa borne supérieure pour obtenir  $f(t, x(t))$ ; inversement si  $x$  vérifie l'équation différentielle avec la condition initiale, alors la dérivée  $dx/dt$  est continue, puisque valant  $f(t, x(t))$ , et on peut l'intégrer entre  $t_0$  et  $t$  pour obtenir la relation intégrale considérée.

### 3. Diverses réductions.

a) L'équation différentielle est dite **autonome** si le second membre (i.e. la fonction  $f$  ci-dessus) est indépendant du temps. C'est une situation, la plus courante en physique ou en mécanique, à laquelle il est toujours possible de se ramener en changeant la fonction inconnue  $x$ .

Si dans

$$\begin{cases} \frac{dx}{dt} = f(t, x) \\ x(t_0) = x_0 \end{cases}$$

on pose

$$(t, x) = y$$

on obtient

$$\begin{cases} \frac{dy}{dt} = (1, f(y)) \\ y(t_0) = (t_0, x_0) \end{cases}$$

tout simplement.

b) Une équation dépendant d'un paramètre  $\alpha$  peut être ramenée à une équation sans paramètres. Contentons-nous du cas autonome. Si dans

$$\begin{cases} \frac{dx}{dt} = f(x, \alpha) \\ x(t_0) = x_0(\alpha) \end{cases}$$

on pose

$$(x, \alpha) = y$$

on obtient

$$\begin{cases} \frac{dy}{dt} = (f(y), 0) \\ y(t_0) = (x_0(\alpha), \alpha) \end{cases}$$

maintenant.

c) Une équation d'ordre supérieur peut être ramenée à une équation du premier ordre. Si dans

$$\begin{cases} \frac{dx}{dt} = f(t, x, \frac{dx}{dt}) \\ x(t_0) = x_0 \\ \frac{dx}{dt}(t_0) = v_0 \end{cases}$$

on pose

$$\begin{pmatrix} x \\ \frac{dx}{dt} \end{pmatrix} = u$$

on obtient

$$\begin{cases} \frac{dy}{dt} = \begin{pmatrix} p_2(y) \\ f(t, y) \end{pmatrix} \\ y(t_0) = (t_0, v_0) \end{cases}$$

où  $p_2$  est la projection sur le second facteur.

Dans la suite nous nous limiterons au premier ordre, en l'absence de paramètres, mais nous ne supposerons pas l'équation autonome car la simplification n'est pas significative.

#### 4. Cylindres de sécurité.

Un **cylindre de sécurité** est le produit  $I \times B'(x_0; r)$  d'un segment  $I$  contenant  $t_0$  et contenu dans  $[t_0 - \tau, t_0 + \tau]$  par une boule fermée  $B'(x_0; r)$ , produit inclus dans  $U$  et sur lequel la norme  $\|f(t, x)\|$  est majorée par  $M$ , où

$$M\tau \leq r .$$

L'intérêt d'un cylindre de sécurité est qu'une solution ne peut en sortir que par les extrémités de  $I$ .

Une fonction  $x$  dérivable sur  $I$  à valeur dans  $E$ , qui vérifie  $x(t_0) = x_0$  ainsi que l'équation en tout point  $t$  tel que  $(t, x(t))$  soit dans  $U$ , est alors nécessairement dans  $V_I$ , de sorte que c'est une solution. Montrons, par exemple, que l'on sort à droite par l'extrémité de  $I$ . La borne supérieure  $\theta$  des  $t$  de  $I$  pour lequel l'image par  $x$  de  $[t_0, t]$  est dans la boule fermée  $B'(x_0; r)$  est atteinte. Supposons par l'absurde que  $\theta$  n'est pas une extrémité de  $I$ . On aurait alors  $\|x(\theta) - x(t_0)\| = r$ . Mais il viendrait  $\|x(\theta) - x(t_0)\| < M\tau$  par l'inégalité des accroissements finis.

On peut toujours trouver un cylindre de sécurité  $[t_0 - \tau, t_0 + \tau] \times B'(x_0; r)$  centré en  $(t_0, x_0)$  où  $\tau$  est  $> 0$ . On commence par choisir un voisinage de  $(t_0, x_0)$  dans  $U$  de la forme  $[t_0 - \sigma, t_0 + \sigma] \times B'(x_0; r)$  sur lequel  $\|f(t, x)\|$  est majorée par  $M = \|f(t_0, x_0)\| + 1$ . Ensuite on prend  $\tau = \min(\sigma, r/M)$ .

**Solutions approchées.** C'est surtout pour la construction de solutions approchées que l'on a recours à un cylindre de sécurité.

Considérons un tel cylindre, de la forme

$$[t_0 - \tau, t_0 + \tau] \times B'(x_0; r) .$$

Nous allons voir que nous pouvons y développer le **schéma d'Euler**. Donnons-nous un pas  $h > 0$ . Nous allons construire une solution approchée  $x$  telle que  $x(t_0) = x_0$  à droite; on ferait de même à gauche.

On définit à partir de  $t_0$  et  $x_0$ , pour  $1 \leq k \leq \tau/h$ , la suite  $t_k = t_0 + kh$  et, par récurrence, la suite  $x_k$  et la solution approchée  $x$  sur  $[t_k, t_k + h]$  en posant

$$x(t) = x_k + (t - t_k)f(t_k, x_k)$$

puis  $x_{k+1} = x(t_{k+1})$ .

Le point important est que le procédé permet de définir  $x$  sur  $[t_0 - \tau, t_0 + \tau]$ . C'est le raisonnement que nous avons fait précédemment avec la dérivée à droite ou à gauche au lieu de la dérivée.

Les solutions approchées ainsi construites sont dérivables à gauche et à droite et vérifient l'équation différentielle d'un côté, ici à droite, aux points  $t_k$ . Elles y sont tangentes de ce côté-là à une courbe intégrale exacte (quand il y en a).

Il ne faudrait pas croire qu'une solution approchée est une approximation affine par morceaux d'une courbe intégrale. Une telle solution saute de courbe intégrale en courbe intégrale, prenant la tangente au sens propre aux points  $t_k$ .

Maintenant on peut montrer, sous les hypothèses du théorème de Cauchy-Lipschitz que nous donnerons, qu'en choisissant une suite de pas tendant vers 0 comme  $h_n = \tau/n$  la suite des solutions approchées converge localement vers une solution exacte.

## 5. Existence et unicité locales.

Nous fixons quelques points de terminologie.

On dit qu'il y a **existence locale** en  $(t_0, x_0)$  si l'on peut trouver un intervalle  $I$  contenant  $t_0$  en son intérieur et une solution  $x$  sur  $I$ .

On dit qu'il y a **existence et unicité locales** en  $(t_0, x_0)$  si l'on peut trouver un intervalle  $I$  contenant  $t_0$  en son intérieur tel que l'équation admette une solution  $x$  et une seule sur  $I$ .

Une solution sur  $I$  est dite **maximale** s'il n'est pas possible de la prolonger à un intervalle strictement plus grand sur lequel elle soit encore une solution.

De façon pédante, c'est un élément maximal dans l'ensemble des couples  $(I, x)$  où  $x$  est une solution sur  $I$  pour la relation d'ordre définie entre  $(I, x)$  et  $(J, y)$  par

$$I \subset J$$

et

$$y|_I = x$$

où  $y|_I$  est la restriction de  $x$  à  $I$ .

**Théorème.** Toute solution peut être prolongée en une solution maximale.

C'est une conséquence du lemme de Zorn. Cependant on peut être plus précis dans certains cas.

Commençons par quelques remarques sur le **comportement des solutions**.

Pour qu'une solution  $x$  définie sur un intervalle  $I = (a, b[$  ouvert en l'extrémité (finie)  $b$  puisse être prolongée sur  $(a, b]$ , il faut il suffit qu'elle admette une limite  $l$  en  $b$  telle que  $(b, l)$  soit dans  $U$ . C'est évidemment nécessaire. Maintenant c'est suffisant car la dérivée  $dx/dt = f(t, x(t))$  admettra la limite  $(b, l)$ . Dans ce cas la fonction  $x$  admet comme dérivée en  $b$  cette limite et l'équation est vérifiée sur  $(a, b]$ .

Si, pour une solution  $x$  sur  $I = (a, b[$ , le point  $(t, x(t))$  reste dans une partie compacte de  $U$  quand  $t$  est au voisinage de  $b$ , alors la solution admet une limite  $l$  en  $b$  et  $(b, l)$  est alors dans  $U$ . En effet la dérivée  $dx/dt$  sera bornée en norme par le maximum de  $\|f(t, x)\|$  sur cette partie compacte. Etant donnée une suite  $b_n \rightarrow b$  de  $I$ , on montre, par l'inégalité des accroissements finis, que la suite  $x(b_n)$  a une limite.

Revenons aux solutions maximales.

Dans le cas où  $U$  est ouvert et où il y a existence locale, une solution  $x$  définie sur un intervalle  $(a, b]$  fermé en  $b$  peut toujours être prolongée au-delà de  $b$  : on choisit, sur un petit intervalle à droite, une solution  $y$  telle que  $y(b) = x(b)$ .

**Principe.** En dimension finie, cas où nous verrons qu'il y a toujours existence locale, une solution maximale  $x$  est nécessairement définie sur un intervalle ouvert et la fonction  $(t, x(t))$  ne peut rester, au voisinage d'aucune des extrémités, dans une partie compacte de  $U$ ; elle doit nécessairement avoir des valeurs d'adhérence au bord de  $U$  ou à l'infini.

Pour ce dernier point, on note en effet que si l'on avait

$$d((t, x(t)), U^c) \geq c \quad \text{et} \quad 1/\|(t, x(t))\| \geq c$$

pour  $c > 0$ , alors  $(t, x(t))$  resterait dans une partie compacte de  $U$ .

**N.B.** Il s'agit des valeurs de la fonction  $(t, x(t))$  et non de la fonction  $x(t)$ , laquelle peut rester dans une partie compacte au voisinage de  $b = +\infty$  par exemple.

## 6. Le théorème de Cauchy-Lipschitz

Il s'énonce ainsi : *si la fonction continue  $f$  est de classe  $\mathcal{C}^1$ , alors il y a existence et unicité locales en tout point de  $U$ .*

Une forme plus générale s'énonce encore comme ceci : *si, au voisinage de  $(t_0, x_0)$ , la fonction continue  $f$  est lipschitzienne par rapport à la seconde variable, i.e. s'il existe un voisinage  $V$  de ce point et une constante  $k \geq 0$  pour lesquels*

$$\|f(t, x) - f(t, y)\| \leq k\|x - y\|$$

*pour  $(t, x)$  et  $(t, y)$  dans  $V$ , alors il y a existence et unicité locale en ce point.*

L'idée de la démonstration est la suivante. D'abord on se ramène à l'équation intégrale donnée en **2**, que l'on écrit

$$x = F(x)$$

où

$$F(x)(t) = x_0 + \int_{t_0}^t f(u, x(u)) du .$$

On va travailler dans l'espace des fonctions définies sur un segment  $I = [t_0 - \tau, t_0 + \tau]$  et à valeurs dans  $E$ . On en fait un espace normé en y prenant la norme uniforme

$$\|x\|_\infty = \sup_{t \in I} \|x(t)\| .$$

On majore alors  $\|F(x) - F(y)\|$  en commençant par tirer

$$\|f(u, x(u)) - f(u, y(u))\| \leq M\|x(u) - y(u)\| \leq M\|x - y\|_\infty$$

de l'inégalité des accroissements finis, si  $M$  majore la seconde différentielle partielle de  $f$ . Il vient ensuite

$$\|F(x)(t) - F(y)(t)\| \leq \tau M\|x - y\|_\infty$$

en intégrant. Enfin

$$\|F(x) - F(y)\|_\infty \leq \tau M\|x - y\|_\infty$$

en passant à la borne supérieure en  $t$ .

Si on s'arrange pour que  $\tau M < 1$ , on aura une application contractante et on pourra appliquer le théorème du point fixe.

Pour être plus précis, on démontre en fait ceci :

**Lemme.** Soient  $I = [t_0 - \tau', t_0 + \tau'']$  et  $\tau = \max(\tau', \tau'')$ . On suppose  $f$  continue sur  $I \times B'(x_0; r)$  et vérifiant

$$\|f(t, x) - f(t, y)\| \leq k\|x - y\|$$

dans ce cylindre. Si l'on a

$$\tau k < 1 \quad \text{et} \quad r \geq \frac{\tau M}{1 - \tau k}$$

où

$$M = \sup_{t \in I} \|f(t, x_0)\|$$

alors il y a existence et unicité locales sur  $I$ .

Il est bien évident qu'on pourra réaliser les hypothèses en choisissant d'abord  $r > 0$  assez petit, puis  $\tau' = \tau'' = \tau$  assez petit. Pour démontrer le lemme on utilise la forme intégrale.

Si  $\underline{x}_0$  désigne la fonction constante valant  $x_0$ , on considère l'application  $F$  de la boule fermée  $B'(\underline{x}_0; r)$  de l'espace  $\mathcal{C}(I, E)$  à valeurs dans l'espace  $\mathcal{C}(I, E)$  lui-même, muni de la norme uniforme, qui est définie par

$$F(x)(t) = x_0 + \int_{t_0}^t f(u, x(u)) du .$$

Démontrons le lemme. Les solutions vérifiant  $x(t_0) = x_0$  sont exactement les points fixes de  $F$ . Cette application vérifie

$$\|F(x)(t) - F(y)(t)\| \leq \left| \int_{t_0}^t \|f(u, x(u)) - f(u, y(u))\| du \right| \leq \tau k \|x - y\|_\infty .$$

Elle est contractante dès que  $\tau k < 1$ . Par ailleurs elle applique la boule dans elle-même. On vérifie que

$$\|F(\underline{x}_0)(t) - \underline{x}_0(t)\| \leq \left| \int_{t_0}^t \|f(u, x_0) - x_0\| du \right| \leq \tau M$$

d'où l'on déduit

$$\|F(\underline{x}_0) - \underline{x}_0\|_\infty \leq (1 - \tau k)r$$

sous les hypothèses du lemme. Alors, si  $x$  est dans la boule, on a

$$\|F(x) - \underline{x}_0\|_\infty \leq \|F(x) - F(\underline{x}_0)\|_\infty + \|F(\underline{x}_0) - \underline{x}_0\|_\infty \leq \tau k r + (1 - \tau k)r = r$$

et il en est de même de  $F(x)$ .

Pour finir, le fait que la boule  $B'(\underline{x}_0; r)$  soit un espace métrique complet permet d'appliquer le théorème du point fixe de Picard.

## 7. Equations linéaires.

Si  $I$  est un intervalle quelconque, si  $A$  est une application continue de  $I$  dans l'espace  $\mathcal{L}(E)$  des endomorphismes continus de  $E$  et  $B$  une application continue de  $I$  dans  $E$ , l'équation

$$\frac{dx}{dt} = Ax + B$$

est dite *linéaire*.

On notera que  $U$  est ici le produit  $I \times E$ , qui n'est ouvert que si  $I$  l'est. Cependant on peut se ramener à ce cas en prolongeant continûment  $A$  et  $B$ .

Le théorème de Cauchy-Lipschitz s'applique au cas linéaire. Cependant on a davantage : il y a *existence globale*. Pour le voir on peut se ramener au cas où  $I$  est un segment. Une solution maximale sera alors définie partout.

Dans le lemme que nous avons établi, on prendrait  $r = +\infty$ , ce qui fait que seule reste la condition  $\tau k < 1$ . Il faut s'en débarrasser et on considère pour cela une itérée  $F^n$  de  $F$  où  $n$  est assez grand pour qu'elle soit contractante.

En effet on montre par récurrence que

$$\|F^n(x)(t) - F^n(y)(t)\| \leq k^n \frac{|t - t_0|^n}{n!} \|x - y\|_\infty .$$

C'est clair pour  $n = 0$  et pour passer du rang au rang  $n + 1$ , on note que

$$\|F^{n+1}(x)(t) - F^{n+1}(y)(t)\| \leq \left| \int_{t_0}^t k \cdot k^n \frac{|u - u_0|^n}{n!} \|x - y\|_\infty \right| = k^{n+1} \frac{|t - t_0|^{n+1}}{(n+1)!} \|x - y\|_\infty .$$

Alors

$$\|F^n(x) - F^n(y)\|_\infty \leq \frac{(\tau k)^n}{n!} \|x - y\|_\infty$$

où le facteur  $(\tau k)^n/n!$  tend vers 0 quand  $n \rightarrow \infty$  comme terme général d'une série exponentielle.

**Remarque.** Pour une équation non linéaire, le fait que  $f$  soit définie sur  $I \times E$  n'implique pas l'existence globale de solutions. Par exemple les solution de l'équation

$$\frac{dx}{dt} = x^2$$

sont en

$$x(t) = -\frac{1}{t - a} .$$

Une telle solution explose pour  $t = a$ .

## 8. Le théorème de Cauchy-Arzelà.

Il s'énonce ainsi : *si  $E$  est de dimension finie et  $f$  continue, il y a existence locale.*

Pour le démontrer on se donne un cylindre de sécurité

$$I \times B'(x_0; r)$$

où  $I = [t_0 - \tau, t_0 + \tau]$  en  $x_0$ ; on  $\tau M \leq r$  et  $\|f(t, x)\| \leq M$  sur le cylindre.

Soit  $x_n$  la solution approchée dans  $I$  issue du schéma d'Euler pour le pas  $\tau/n$ . Les valeurs des  $x_n$  en  $t_0$  sont les mêmes et la majoration uniforme des dérivées en fait une suite équicontinue. Par le théorème d'Ascoli, cette suite est incluse dans une partie compacte de  $\mathcal{C}(I, E)$  et on peut en extraire une sous-suite qui converge uniformément sur  $I$  vers une fonction continue  $x$ .

Nous allons voir que

$$\left\| \frac{dx_n}{dt}(t) - f_n(t, x_n(t)) \right\| \leq \epsilon_n$$

où la suite  $\epsilon_n$  tend vers 0. En effet, étant donné  $\epsilon > 0$ , par la continuité uniforme de  $f$  sur le cylindre, on peut choisir  $\eta > 0$  tel que  $|t - u| \leq \eta$  et  $\|x - y\| \leq \eta$  impliquent

$$\|f(t, x) - f(u, y)\| \leq \epsilon .$$

Soit alors  $N$  tel que  $\tau/N < \eta$  et  $\tau M/N < \eta$ . Pour  $n \geq N$  on a

$$\frac{dx_n}{dt}(t) = f(u, x_n(u))$$

où  $|t - u| \leq \tau/n$  et  $\|x_n(t) - x_n(u)\| \leq \tau M/n$ .

Il ne reste plus qu'à intégrer en  $t_0$  et  $t$  cette inégalité, puis passer à la limite uniforme dans l'expression intégrale obtenue.

**Remarque.** Il peut ne pas y avoir unicité locale. Prenons l'exemple du ballon qui glisse sur l'arête d'un toit. Soit  $h$  l'altitude, vers le bas, à partir de l'arête. Par le théorème de l'énergie cinétique, on a

$$\frac{1}{2} \left( \frac{dh}{dt} \right)^2 = gh .$$

Or une équation du type

$$\frac{dh}{dt} = \sqrt{h}$$

s'intègre, dans le demi-plan  $h \geq 0$ , en

$$h(t) = \frac{(t - t_0)^2}{2}$$

pour  $t \geq t_0$ , ou bien en

$$h(t) = 0 .$$

Les courbes intégrales maximales suivent l'axe des  $t$  jusqu'en  $t_0$  où elles empruntent une parabole. En  $(t_0, 0)$  on n'a pas unicité locale; il y a une infinité de possibilités pour les  $t$  voisins.

## 9. Différentiabilité locale du flot.

En situation d'existence et d'unicité locales, pour une valeur  $t_0$  donnée, le **flot** désigne la solution de l'équation différentielle comme fonction de la variable  $t$  et de la condition initiale  $\xi$ .

**Théorème.** Si  $f$  est de classe  $\mathcal{C}^k$ , le flot est aussi de classe  $\mathcal{C}^k$ .

On commence par en établir une version locale. Pour cela on reprend ce qui a été fait en **2**. Une solution  $x$  sur  $I$  vérifiant  $x(t_0) = \xi$  est une fonction continue telle que

$$x(t) - \xi - \int_{t_0}^t f(u, x(u)) du = 0$$

pour  $t$  dans  $I$ . Etant donnée une solution prenant la valeur  $x_0$  en  $t_0$ , si l'on peut, au voisinage, tirer  $x$  comme fonction implicite de  $\xi$ , on aura bien le résultat cherché.

Considérons un segment  $I$  et l'application  $\Psi$  qui à la fonction  $x$  de  $V_I$  associe la fonction de  $F = \mathcal{C}(I, E)$  définie par

$$(\Psi(x))(t) = \int_{t_0}^t f(u, x(u)) du$$

sur  $I$ . Considérons encore la fonction  $\Phi$  qui au point  $\xi$  de  $E$  et à la fonction  $x$  de  $V_I$  associe  $\Phi(\xi, x) = x - \xi - \Psi(x)$ . Il s'agit de voir que  $\Phi$  est différentiable et que sa différentielle partielle en  $x$  est inversible.

**Lemme.** Si  $f$  est de classe  $\mathcal{C}^1$  dans  $U$ , alors  $\Psi$  est différentiable dans  $V_I$  et sa différentielle donnée par

$$(d\Psi(x).h)(t) = \int_{t_0}^t \frac{\partial f}{\partial x}(u, x(u)) h(u) du .$$

Il est facile de voir que c'est une application linéaire continue sur  $\mathcal{C}(I, E)$ . On majore en effet l'intégrale par

$$\tau k \|h\|_\infty$$

si  $I$  est inclus dans  $[t_0 - \tau, t_0 + \tau]$  et si  $\|\partial f / \partial x\| \leq k$  sur l'image par  $\delta_x$  de  $I$ .

Maintenant on vérifie que la fonction  $\Psi(x+h) - \Psi(x) - d\Psi.h$  qui prend en  $t$  la valeur

$$\int_{t_0}^t \left[ f(u, x(u) + h(u)) - f(u, x(u)) - \frac{\partial f}{\partial x}(u, x(u)) h(u) \right] du$$

est tangente à zéro en  $x$ . Etant donné  $\epsilon > 0$ , on peut trouver  $\eta > 0$  tel que

$$\left\| \frac{\partial f}{\partial x}(t, y+z) - \frac{\partial f}{\partial x}(t, y) \right\| \leq \epsilon$$

pour  $\|z\| \leq \eta$  et  $(t, y)$  dans  $\delta_x(I)$  : ici la compacité de  $I$  a été utile pour obtenir l'uniformité en  $t$ . Il en résulte

$$\left\| f(t, y+z) - f(t, y) - z \frac{\partial f}{\partial x}(t, y) \right\| \leq \epsilon \|z\|$$

pour  $\|z\| \leq \eta$  et  $(t, y)$  dans  $\delta_x(I)$ , par la formule des accroissements finis. Il n'y a plus qu'à porter cela dans l'intégrale.

Le lemme étant montré, la dérivée partielle en  $x$  de  $\Phi$  est donnée par

$$\frac{\partial \Phi}{\partial x} = 1_F - d\Psi .$$

Elle est inversible, comme application linéaire continue, dès que l'on peut prendre  $\tau k \leq 1/2$ , donc dès que  $I$  est un segment assez petit.

Maintenant, si l'on veut appliquer la stratégie précédente pour obtenir en même temps l'existence locale des solutions, il faut travailler un peu plus. Dans ce qui suit, on se limite au cas où  $E$  est de dimension finie.

Pour simplifier encore, prenons désormais  $t_0 = 0$  et  $x_0 = 0$  et essayons de résoudre, avec la condition initiale correspondante, l'équation autonome

$$\frac{dx}{dt} = f(x) .$$

Sous les hypothèses faites plus haut, l'application  $\Phi_0 : x \mapsto \Phi(0, x) = x - \Psi(x)$  a une différentielle inversible. C'est donc, par le théorème d'inversion locale, une application ouverte. Il reste à voir que 0 est dans l'image.

Si  $I = [-\tau, \tau]$ , prenons  $\tau > 0$  et  $r > 0$  assez petits pour que le cylindre  $I \times B'_E(0; r)$  soit inclus dans  $U$  et que  $\tau M < r/2$  où  $M$  majore  $\|f\|$  sur le cylindre. Soient, dans  $\mathcal{C}(I, E)$ , les boules ouvertes  $U = B(0; r)$  et  $V = B(0; r/2)$ .

Alors  $\Phi_0(U) \cap V$  est une partie fermée de  $V$  : si une suite  $\Xi(x_n)$  converge vers  $y$  dans  $V$ , on peut établir une propriété de compacité permettant de se ramener, par extraction de sous-suite, au cas où  $z_n = x_n - \Phi_0(x_n)$  converge vers  $z$  et  $x_n$  convergera vers  $x = y + z$ . De  $\|z_n\| \leq r/2$  il résulte que  $\|x\| < r$ ; d'où la propriété.

Enfin  $\Phi_0(U) \cap V$  est une partie ouverte de  $V$ . Comme  $V$  est connexe, ce ne peut être que la partie vide ou  $V$ . Mais elle contient déjà  $\Phi_0(0)$ . Ainsi s'achève la démonstration.

## 10. Vers la différentiabilité globale du flot.

Etant donnée une solution définie sur un segment  $I = [a, b]$  et y prenant en  $t_0$  une valeur initiale  $x_0$ , peut-on, pour  $\xi$  voisin de  $x_0$ , trouver une solution sur  $I$  prenant la valeur initiale  $\xi$  en  $t_0$  qui dépend différentiablement de  $\xi$ ? C'est toujours un problème de fonction implicite, mais, pour traiter le cas d'un intervalle  $[a, b]$  qu'on ne peut pas réduire à volonté, il faut améliorer la stratégie développée en **2**.

On va commencer par simplifier les notations en se ramenant au cas où  $t_0 = 0$ ,  $x_0 = 0$ ,  $I = [-\tau', \tau'']$  et où la solution est elle-même 0; on pose  $\tau = \max(\tau', \tau'')$ . On considère toujours l'équation

$$\frac{dx}{dt} = f(t, x)$$

où  $f$  est de classe  $\mathcal{C}^1$ . Une possibilité est de remplacer  $\Phi$  par une intégrée de la forme

$$(\xi, x) \mapsto x - \Theta(\xi, \Theta(\xi, x))$$

où  $\Theta(\xi, x) = \xi + \Psi(x)$  et où le nombre d'itérations, ici égal à 2, est adapté. La différentielle partielle en  $x$  s'obtiendra à l'aide du théorème de composition; ce sera ainsi

$$h \mapsto \int_0^t d(f(u, \Psi(0)))(u) \cdot \left( \int_0^u df(u, 0) \cdot h \, dv \right) du$$

pour  $\Theta(\xi, \Theta(\xi, x))$  en  $(0, 0)$ . On majore cette dernière en norme par

$$k^2 \int_0^t \left( \int_0^u dv \right) dt \leq \frac{(\tau k)^2}{2}$$

si  $k$  majore  $\|f'\|$  sur un cylindre convenable contenant  $\Theta(0, 0)(I)$ .

Avec  $n$  itérations, on obtiendrait une majoration en  $(\tau k)^n/n!$  qui sera toujours  $< 1$  pour  $n$  assez grand.

## Chapitre 4. Exponentielle de matrice

### 1. Equations différentielles linéaires.

Comme nous l'avons vu, étant donné un espace normé complet  $E$ , une équation différentielle **linéaire** est du type

$$\frac{dx}{dt} = A(t).x + B(t)$$

où  $A, B$  sont respectivement des applications continues d'un intervalle  $I$  dans  $L(E)$  et dans  $E$ ; ici on a fait apparaître la variable  $t$  pour montrer la dépendance. Une telle équation admet, pour toute donnée initiale  $(t_0, x_0)$  dans  $I \times E$ , une solution globale, définie sur  $I$  tout entier, qui est unique.

On s'intéresse à la valeur en  $t$  de la solution de l'équation homogène

$$\frac{dx}{dt} = A(t).x$$

qui vaut  $x_0$  en  $t_0$ . Pour  $t_0$  fixé, cette dernière est donnée par

$$R(t).x_0$$

où l'**endomorphisme fondamental**  $R$  est la solution de l'**équation résolvente**

$$\frac{dR}{dt} = A(t) \circ R$$

qui vérifie la condition

$$R(t_0) = 1_E ;$$

ici  $1_E$  est, comme d'habitude, l'application identique ou endomorphisme unité de  $E$ .

Cette équation, qui est également linéaire, porte sur un endomorphisme et non plus sur un vecteur. La vérification de ce qui est annoncé est évidente : la dérivée de  $x(t) = R(t).x_0$  est

$$\frac{dR}{dt}.x_0 = A(t).R(t).x_0 = A(t).x(t)$$

et par ailleurs  $R(t_0).x_0 = 1_E.x_0 = x_0$ .

En dimension finie, le déterminant  $W$  de  $R$ , appelé **déterminant wronskien**, s'obtient par la formule

$$W(t) = W(t_0) \exp\left(\int_{t_0}^t \operatorname{tr} A(u) du\right)$$

de Liouville. La démonstration repose sur le fait que la différentielle du déterminant en l'identité est la trace.

L'intérêt de la résolvente est la résolution de l'équation inhomogène par la méthode de la variation des constantes. Nous expliquerons la méthode, qui est générale, dans le cas particulier des équations dites à coefficients constants, pour lesquelles  $A$  ne dépend pas de  $t$ .

D'ailleurs nous commençons par nous intéresser à l'équation résolvente dans cet exemple, en notant désormais  $u$  l'endomorphisme constant qui apparaît.

## 2. Définition de l'exponentielle.

Dans toute la suite le corps de base est le corps  $\mathbf{C}$  des nombres complexes. On sait qu'une matrice carrée  $M$  d'ordre  $n$  peut être identifiée à un endomorphisme de  $\mathbf{C}^n$ . On lui associe précisément celui qui transforme les vecteurs de la base canonique en les vecteurs colonne de la matrice. Par conséquent, plutôt qu'une matrice carrée, on se donnera plus généralement un endomorphisme  $u$  d'un espace vectoriel  $E$  de dimension finie sur  $\mathbf{C}$ .

Nous pouvons énoncer. On note  $1_E$  l'endomorphisme identique de  $E$ .

**Théorème et définition.** Il existe une unique application dérivable  $x$  de  $\mathbf{R}$  dans l'espace  $L(E)$  des endomorphismes de  $E$ , dont les valeurs sont des polynômes en  $u$  et qui vérifie

$$\frac{dx}{dt} = u \circ x$$

ainsi que  $x(0) = 1_E$ . Sa valeur pour  $t = 1$  est, par définition, l'exponentielle  $\exp(u)$  de  $u$ .

**N.B.** Dans le cas où l'on considère des matrices plutôt que des endomorphismes, on écrit  $u.x$  pour le produit des matrices  $u$  et  $x$ . Nous ferons cette convention d'écriture désormais.

Si nous connaissons le théorème de Cauchy-Lipschitz nous pouvons omettre la condition sur les valeurs de  $x$  pour l'unicité. Le fait que l'équation soit linéaire assure en plus l'existence d'une solution globale.

La forme indiquée permet de démontrer l'unicité de façon élémentaire. Si  $x$  et  $y$  sont deux solutions de l'équation complétée par la condition initiale, il vient

$$\frac{d}{dt}(x(t)y(-t)) = u.x.y - x.u.y = 0$$

puisque  $y$  (entre autres) commute avec  $u$ . Par suite  $x(t).y(-t)$  est constant, égale à sa valeur en 0 qui est  $1_E$ .

On obtient notamment  $y(t).y(-t) = 1_E$  en remplaçant  $x$  par  $y$ . Par suite les valeurs des solutions sont nécessairement dans le groupe  $GL(E)$  des automorphismes de  $E$ . Enfin  $x(t) = y(t)$ , d'où l'unicité.

De plus

$$x(t+u) = x(t)x(u)$$

parce que  $x(u)^{-1}x(t+u)$  vérifie l'équation et la condition initiale. On a donc un homomorphisme du groupe additif  $\mathbf{R}$  dans le groupe  $GL(E)$ .

Admettons que l'on ait construit une solution  $x$ . On obtient

$$\frac{d}{dt}x(\tau t) = \tau u.x(\tau t)$$

de sorte que  $x(\tau) = \exp(\tau u)$ . La solution est donc  $x(t) = \exp(tu)$ .

Si maintenant  $u$  et  $v$  sont des endomorphismes qui commutent, on vérifie que

$$\exp(u+v) = \exp(u).\exp(v) .$$

En effet  $\exp(v)^{-1}.\exp(tu+v)$  vérifie l'équation et la condition initiale. Pour cette raison on emploie aussi la notation  $e^u$  pour  $\exp u$ .

Maintenant la fonction  $e^{tu}$  est de classe  $\mathcal{C}^\infty$  et le reste de Taylor donné par

$$\frac{u^k}{k!} \int_0^1 (1-t)^{k-1} e^{tu} dt$$

est majoré par

$$\frac{\|u\|^k}{k!} \int_0^1 \|e^{tu}\| dt$$

si l'on a pris une norme sur  $E$  et la norme d'opérateur correspondante sur  $L(E)$ . Ce reste tend vers 0 quand  $n \rightarrow \infty$  de sorte que

$$e^u = 1 + \frac{u}{1!} + \dots + \frac{u^k}{k!} + \dots$$

dans  $L(E)$ .

### 3. Applications linéaires à coefficients constants.

A l'aide de ce qui précède, on peut résoudre l'équation homogène

$$\frac{dx}{dt} = u.x$$

dans laquelle  $u$  est un endomorphisme de  $E$  donné et où  $x$  est une fonction inconnue à valeurs dans  $E$ .

**Théorème.** La solution de l'équation homogène valant  $x_0$  en  $t_0$  est

$$x(t) = e^{(t-t_0)u} . x_0 .$$

En effet  $x(t_0) = x_0$  et l'on obtient

$$x'(t) = u . e^{(t-t_0)u} . x_0$$

en dérivant.

Considérons l'équation inhomogène

$$\frac{dx}{dt} = u.x + f$$

dans laquelle  $f$  est une fonction continue quelconque. Pour la résoudre on applique la méthode de Lagrange, encore dite de variation des constantes, qui consiste à chercher une solution du type

$$x = e^{tu} . c$$

où  $c$  est une fonction inconnue. En dérivant il vient

$$u . e^{tu} . c + e^{tu} . c'$$

de sorte que l'équation se ramène à

$$e^{tu} . c' = f \quad \text{ou} \quad c' = e^{-tu} . f ,$$

à savoir à une primitivation.

On notera ceci. Pour une équation d'ordre supérieur, une équation scalaire d'ordre 2 comme

$$x'' = ax' + bx + f$$

par exemple, on se ramène au système

$$\frac{d}{dt} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ b & a \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0 \\ f \end{pmatrix}.$$

en posant  $y = dx/dt$ . Il s'agit alors de prendre l'exponentielle de la matrice  $A$  carrée d'ordre 2 ci-dessus. La méthode de Lagrange conduit à

$$e^{tA} \cdot \begin{pmatrix} c' \\ d' \end{pmatrix} = \begin{pmatrix} 0 \\ f \end{pmatrix}.$$

Cela revient à ajouter, pour la première ligne de la solution générale

$$e^{tA} \cdot \begin{pmatrix} c \\ d \end{pmatrix}$$

de l'équation homogène, la relation obtenue en égalant à 0 l'expression où l'on ne dérive que les constantes et à reporter dans l'équation inhomogène.

C'est la construction de l'exponentielle qui va maintenant nous occuper.

#### 4. $\mathbf{C}$ -algèbres.

Une algèbre possède une structure d'anneau avec une multiplication externe qui en fait un espace vectoriel. Nous ne considérerons que des algèbres unitaires, i.e. possédant un élément unité, lequel sera noté 1. Nous identifierons le scalaire  $s$  et l'élément  $s.1$ , ce qui revient à considérer le corps de base comme inclus dans l'algèbre. Une algèbre unitaire est ainsi simplement un anneau unitaire qui contient  $\mathbf{C}$ .

Nous considérons la sous-algèbre  $A$  engendrée par  $u$  dans l'algèbre  $L(E)$  des endomorphismes de  $E$ . Elle est constituée des endomorphismes  $p(u)$ , où  $p$  parcourt les polynômes de  $\mathbf{C}[X]$ .

C'est dans cette algèbre que nous allons travailler. Elle est évidemment commutative et de dimension finie comme espace vectoriel. Le morphisme surjectif d'algèbres

$$\mathbf{C}[X] \rightarrow A$$

qui a  $X$  associe  $u$  a donc un noyau non nul, qui est un idéal de  $\mathbf{C}[X]$ , lequel est engendré (par la division euclidienne) par un polynôme unitaire non constant  $D$ . Ce dernier s'appelle le *polynôme minimal* de  $u$ .

L'homomorphisme

$$\mathbf{C}[X]/D \rightarrow A$$

obtenu par passage au quotient est un isomorphisme d'algèbres. Par conséquent nous pouvons commencer par oublier  $E$  et l'endomorphisme  $u$  et travailler dans l'algèbre  $\mathbf{C}[X]/D$  des classes modulo  $D$  de polynômes.

**Le Spectre.** Pour un élément  $a$  d'une algèbre, le spectre est l'ensemble des scalaires  $s$  tels que  $a - s$  ne soit pas inversible. En particulier le spectre d'un opérateur  $u$  de  $E$  est son spectre dans l'algèbre  $L(E)$ .

Revenons à l'algèbre  $\mathbf{C}[X]/D$ . Par le théorème fondamental de l'algèbre, on écrit

$$D = \prod_{s \in S} (X - s)^{m_s}$$

où  $S$  est l'ensemble des racines de  $D$  et où  $m_s$  est la multiplicité de la racine  $s$ .

L'ensemble  $S$  des racines est exactement le spectre de  $X$  dans  $\mathbf{C}[X]/D$ . D'abord, si  $s$  est dans  $S$ , le facteur  $X - s$  apparaît comme un diviseur de 0 modulo  $D$  et il n'est donc pas inversible. Ensuite, si  $s$  n'est pas dans  $S$ , alors  $D(s) \neq 0$ ; comme  $D - D(s) = (X - s)Q$  où  $Q$  est un polynôme, il vient  $1 = (X - s) \cdot \frac{Q}{-D(s)}$  modulo  $D$  et  $X - s$  est inversible.

Notre souci, à partir d'ici, est de disposer d'une propriété de décomposition qui permette d'isoler le contribution de chaque élément du spectre.

**Le théorème chinois.** Ceux qui n'aiment pas spécialement les chinoiseries et les isomorphismes canoniques peuvent sauter directement au paragraphe suivant.

On a une application naturelle

$$\mathbf{C}[X]/D \rightarrow \prod_{s \in S} \mathbf{C}[X]/(X - s)^{m_s}$$

obtenue, à partir de l'application diagonale de  $\mathbf{C}[X]$  dans  $\mathbf{C}[X]^S$  qui à  $P$  associe la famille constante égale à  $P$ , en composant à droite avec les projections qui envoient le facteur  $\mathbf{C}[X]$  de rang  $s$  dans  $\mathbf{C}[X]/(X - s)^{m_s}$ , puis en passant au quotient par  $D$ . C'est un isomorphisme parce que les polynômes  $(X - s)^{m_s}$  sont premiers entre eux.

Ainsi a-t-on décomposé l'algèbre donnée en un produit d'algèbres. Dans ce produit apparaissent des éléments idempotents : on notera  $I_s$  celui dont la projection de rang  $s$  est 1 et dont les autres sont nulles. Clairement

$$1 = \sum_{s \in S} I_s .$$

Pour expliciter un représentant de l'élément  $I_s$  de l'algèbre  $\mathbf{C}[X]/D$ , nous cherchons  $P_s$  ayant les propriétés suivantes : il est congru à 1 modulo  $(X - s)^{m_s}$  et à 0 modulo les autres facteurs  $(X - r)^{m_r}$ . La seconde condition signifie qu'il est divisible par le produit

$$D_s = \frac{D}{(X - s)^{m_s}} = \prod_{r \neq s} (X - r)^{m_r}$$

de ces autres facteurs, donc de la forme  $P_s = Q_s D_s$ . La première veut que  $Q_s D_s = 1$  modulo  $(X - s)^{m_s}$ , autrement dit que  $Q_s$  inverse  $D_s$  modulo  $(X - s)^{m_s}$ . Or travailler modulo  $(X - s)^{m_s}$  veut dire qu'on prend un développement limité en  $s$  d'ordre  $m_s - 1$  (au moins). La réponse est donc

$$Q_s = \text{dl}_s\left(\frac{1}{D_s}\right)$$

où il est entendu que le développement limité  $\text{dl}_s$  en  $s$  est pris à l'ordre  $m_s - 1$ .

Finalement

$$P_s = \text{dl}_s\left(\frac{1}{D_s}\right) D_s \quad \text{et} \quad 1 = \sum_{s \in S} \text{dl}_s\left(\frac{1}{D_s}\right) D_s .$$

## 5. La décomposition des fractions rationnelles en éléments simples.

Ce paragraphe est une digression. On peut aussi en sauter la lecture.

Le problème dont nous avons parlé se rencontre déjà à l'occasion de la décomposition en éléments simples d'une fraction comme  $1/D$  — ou  $N/D$  plus généralement. Si l'on cherche à faire apparaître la partie singulière en  $s$ , on écrira

$$\frac{1}{D} = \frac{1}{(X-s)^{m_s}} \frac{1}{D_s}$$

où  $D_s$  a été défini comme plus haut et on fera un développement limité de  $1/D_s$  en  $s$  à l'ordre  $m_s - 1$  au moins.

En fait on peut écrire de façon unique

$$\frac{1}{D} = \frac{Q}{(X-s)^{m_s}} + \frac{R}{D_s}$$

où  $Q, R$  sont des polynômes et où le degré de  $Q$  est strictement inférieur à  $m_s$ .

En multipliant par  $D$  la première égalité, on obtient

$$1 = Q \cdot D_s + (X-s)^{m_s} R$$

et la recherche de  $Q$  est celle d'un inverse de  $D_s$  modulo  $(X-s)^{m_s}$ , que l'on trouve bien sûr avec le développement limité  $dl_s(\frac{1}{D_s})$  à l'ordre  $m_s - 1$  de  $1/D_s$ .

## 6. La formule de Taylor-Gauss.

Des calculs du théorème chinois ou des motivations provenant de la décomposition en éléments simples qui précèdent, nous ne retenons que la définition de  $D_s$  et l'idée de considérer le développement limité

$$dl_s\left(\frac{1}{D_s}\right)$$

de  $\frac{1}{D_s}$  à l'ordre  $m_s - 1$  en  $s$ . Nous avons

$$(1) \quad 1 = dl_s\left(\frac{1}{D_s}\right) D_s \quad \text{mod} \quad (X-s)^{m_s}$$

puis

$$(2) \quad 1 = \sum_{s \in S} dl_s\left(\frac{1}{D_s}\right) D_s \quad \text{dans} \quad \mathbf{C}[X]/D$$

que l'on transforme en

$$(1') \quad f = dl_s(f) dl_s\left(\frac{1}{D_s}\right) D_s \quad \text{mod} \quad (X-s)^{m_s}$$

et

$$(2') \quad f = \sum_{s \in S} dl_s(f) dl_s\left(\frac{1}{D_s}\right) D_s \quad \text{dans} \quad \mathbf{C}[X]/D$$

pour tout polynôme  $f$  plus généralement.

La relation (1) est claire. La relation (2) est vraie modulo tous les  $(X - s)^{m_s}$  puisque  $D_r$  est divisible par ce facteur pour  $r \neq s$ , donc modulo  $D$ . Les relations (1') et (2') s'obtiennent en multipliant par  $f$  les précédentes, lesquelles apparaissent comme des cas particuliers. Les développements limités en  $s$  sont à l'ordre  $m_s - 1$ ; dans le produit les termes d'ordre supérieur sont ignorés.

On remarquera d'ailleurs que l'ordre auquel il faut calculer le développement limité de  $1/D_s$  est évident sur les formules (2) et (2') elles-mêmes : on néglige  $(X - s)^{m_s}$  parce que  $(X - s)^{m_s} D_s = D$  est nul dans  $\mathbf{C}[X]/D$ . Mais rien n'interdit d'aller au delà.

Le second membre de (1) définit pour chaque  $s$  un élément  $P_s$  et on vérifie que  $P_s P_r$  est divisible par  $D$ , donc nul dans  $\mathbf{C}[X]/D$  si  $r \neq s$ . Par suite  $P_s = P_s \cdot 1 = P_s^2$  de sorte que (2) permet de retrouver une décomposition de 1 en éléments idempotents orthogonaux.

## 7. Le cas d'un endomorphisme.

La même relation (2) fournit une décomposition de l'identité en projecteurs orthogonaux  $p_s = P_s(u)$ , appelés projecteurs spectraux, quand on l'applique à un endomorphisme  $u$ . Ainsi

$$1 = \sum_{s \in S} p_s .$$

Maintenant la relation (2'), avec  $f = X$ , appliquée encore à  $u$  donne la décomposition de  $u$  en parties semi-simple et nilpotente

$$u = \sum_{s \in S} s \cdot p_s + n_s$$

où la partie nilpotente  $n_s = N_s(u)$  s'obtient par

$$N_s = (X - s) dl_x \left( \frac{1}{D_s} \right) D_s$$

où le développement est pris à l'ordre  $m_s - 2$ ; sa puissance  $n_s^k = N_s^k(u)$  s'obtient par

$$N_s^k = (X - s)^k dl_s \left( \frac{1}{D_s} \right) D_s$$

plus généralement, où le développement est pris à l'ordre  $m_s - k - 1$ , avec  $n_s^{m_s} = 0$  évidemment.

Dans ces relations, tous les éléments commutent : ce sont des polynômes en  $u$ . De plus les éléments en des points  $s$  et  $t$  différents sont orthogonaux, ce qui signifie que leur produit est nul : dans l'algèbre produit du **4**, les composants en  $t$  du premier et en  $s$  du second sont nulles.

Chaque  $p_s$  est idempotent car ils provient de l'unité de  $\mathcal{C}[X]/(X - s)^{m_s}$ ; il vérifie  $p_s^2 = p_s$  : c'est un projecteur; il projette sur son image, le sous-espace spectral en  $s$ , parallèlement à son noyau. Leur somme étant l'identité, l'espace est décomposé en somme directe des sous-espaces spectraux.

Chaque  $n_s$  est nilpotent car il provient de  $X - s$ , lequel est évidemment nilpotent dans  $\mathcal{C}[X]/(X - s)^{m_s}$ ; précisément  $n_s^{m_s} = 0$ . De plus  $n_s$  est subordonné à  $p_s$  : on a  $n_s p_s = n_s$ ; c'est aussi une conséquence du reste.

## 8. Le calcul fonctionnel.

Dans la formule

$$(2') \quad f = \sum_{s \in S} dl_s(f) dl_s\left(\frac{1}{D_s}\right) D_s \quad \text{dans } \mathbf{C}[X]/D$$

précédente, il suffit d'être capable de donner un sens à  $dl_s(f)$  en tout point  $s$  du spectre, pour donner un sens au second membre. Précisément il suffit de se donner en chaque  $s$  un développement limité abstrait d'ordre  $m_s - 1$ ; c'est ce qu'on appelle un *jet* d'ordre  $m_s - 1$ .

Quand on applique la formule à un endomorphisme  $u$ , on définit l'endomorphisme que nous noterons  $f[u]$ .

**Retour sur l'exponentielle d'un endomorphisme.** C'est ici que le calcul fonctionnel précédent trouve sa place. On peut définir  $e^{tu}$  comme  $f[u]$  en prenant pour  $f$ , en chaque point  $s$  du spectre de  $u$ , le développement limité à l'ordre  $m_s - 1$  en  $s$  tiré du développement en série formelle et obtenu en la développant en  $X - s$  selon

$$\exp(tX) = \exp(ts + t(X - s)) = e^{ts} \left[ 1 + \frac{t(X - s)}{1!} + \dots + \frac{t^k(X - s)^k}{k!} + \dots \right]$$

Il reste à voir que  $e^{tu}$  est solution de l'équation différentielle linéaire

$$\frac{dx}{dt} = u \circ x$$

avec la condition initiale

$$x(0) = 1_E .$$

On l'obtient aussi en dérivant en  $t$  le développement limité en  $s$  à l'ordre  $m_s - 1$ , ce qui revient à dériver une somme partielle de la série formelle  $\exp(tX)$ . Or la dérivée en  $t$  de cette dernière série est

$$se^{ts} \exp(tX) + e^{ts}(X - s). \exp(tX) = X. \exp(tX)$$

de sorte que la dérivée en  $t$  de  $e^{tu}$  est  $u.e^{tu}$ .

## 9. En pratique.

On va considérer le cas d'une matrice carrée  $A$  d'ordre 3 ayant une valeur propre simple  $\lambda$  et une valeur propre double  $\mu$ , autrement dit dont le polynôme caractéristique est

$$D = (X - \lambda)(X - \mu)^2$$

où  $\lambda \neq \mu$ . Plus généralement on va supposer que  $u$  est un endomorphisme annulé par  $D$ ; le cas le plus intéressant est celui où  $D$  est le polynôme minimal, mais il n'est pas nécessaire de s'y limiter.

Une première stratégie consiste à décomposer le spectre en séparant les facteurs  $(X - \mu)^2$  et  $X - \lambda$ . On part de

$$1 = \frac{X - \mu}{\lambda - \mu} + \frac{X - \lambda}{\mu - \lambda} .$$

Si on multiplie le premier terme de droite par l'expression de 1 ainsi obtenue, il vient

$$1 = P_\lambda + P_\mu$$

où

$$P_\lambda = \frac{(X - \mu)^2}{(\lambda - \mu)^2} \quad , \quad P_\mu = \frac{X - \lambda}{\mu - \lambda} \left[ 1 + \frac{X - \mu}{\lambda - \mu} \right] .$$

Il vient encore

$$\begin{aligned} X &= (\lambda + X - \lambda)P_\lambda + (\mu + X - \mu)P_\mu \\ &= \lambda P_\lambda + \mu P_\mu + \frac{(X - \lambda)(X - \mu)}{\mu - \lambda} \quad \text{modulo } D \end{aligned}$$

En remplaçant  $X$  par l'endomorphisme  $u$  de polynôme minimal  $D$ , on obtient aussitôt sa décomposition spectrale.

Le même calcul avec  $\exp(tX)$  donne

$$\exp(tX) = e^{\lambda t} P_\lambda + e^{\mu t} (1 + t(X - \mu)) P_\mu \quad \text{modulo } D$$

soit

$$e^{\lambda t} P_\lambda + e^{\mu t} \left[ P_\mu + t \frac{(X - \lambda)(X - \mu)}{\mu - \lambda} \right] \quad \text{modulo } D$$

et il suffit de remplacer  $X$  par  $u$  pour obtenir  $e^{tu}$ .

Cependant calculer l'exponentielle de  $u$  revient aussi à calculer sa décomposition spectrale puisqu'on aura

$$e^{tu} = \sum_{s \in S} e^{ts} \left[ p_s + tn_s + \cdots + \frac{t^k n_s^k}{k!} + \cdots \right]$$

où chaque somme est arrêtée à  $k = m_s - 1$ , en remplaçant  $X$  par  $u$  dans la formule du 8. Cela étant, il convient de faire le calcul de la décomposition en termes de polynômes en  $u$ .

Dans le cas considéré en exemple, on écrira a priori

$$A = \lambda p + \mu q + n$$

où les projecteurs spectraux  $p, q$  vérifient  $p^2 = p, q^2 = q, pq = 0$  et où la partie nilpotente  $n$  vérifie  $np = 0, nq = n, n^2 = 0$ .

On en déduira

$$e^{tu} = e^{\lambda t} p \cdot e^{\mu t} [q + tn] .$$

On calcule  $p, q, n$  à l'aide de  $I, A$  et  $A^2$  ou plutôt de  $I, A - \mu I$  et  $(A - \mu I)^2$ .

## Chapitre 5. Retour sur la topologie générale

Dans l'étymologie du mot "topologie" — ou "analysis situs" — on trouve la notion d'espace, du grec *topos* ou du latin *situs*. On parlera donc d'**espaces**. On se placera couramment "dans un espace  $X$ ". On considèrera une application "d'un espace  $X$  dans un espace  $Y$ ". Dans de tels cas rien ne pourra avoir de sens en dehors.

Précisément la donnée d'un espace sera celle d'un ensemble et d'une distance sur celui-ci. Cependant les notions considérées n'utiliseront qu'une partie de l'information apportée par la distance, laquelle sera rapidement reléguée au second plan.

### 1. Topologie.

**Dans un espace  $X$** , on considèra des voisinages, des parties ouvertes ou fermées. Ce sont des notions *relatives*, précisément relatives à un espace donné.

Pour voir qu'une partie  $V$  de l'espace  $X$  est un *voisinage* d'un point  $a$  de  $X$ , on doit trouver un nombre  $r > 0$  tel que  $V$  contienne la boule (ouverte ou fermée) de centre  $a$  et de rayon  $r$ .

Pour voir qu'une partie  $U$  de l'espace  $X$  est *ouverte*, on se donne un point  $a$  arbitraire dans  $U$  et on doit montrer que c'est un voisinage de  $a$ , autrement dit trouver un nombre  $r > 0$  tel que  $U$  contienne la boule (ouverte ou fermée) de centre  $a$  et de rayon  $r$ .

Pour voir qu'une partie  $F$  de l'espace  $X$  est *fermée*, on peut montrer que son complémentaire dans  $X$  est ouvert. Dans le cas métrique que nous considérons, on peut aussi se donner une suite arbitraire  $(x_n)$  de  $F$  qui converge vers un point  $a$  de  $X$ , et montrer que nécessairement  $a$  est dans  $F$ .

On notera que la notion de voisinage, comme les qualificatifs ouvert et fermé, s'appliquent à des *parties* de  $X$ . Pour éviter toute confusion on parlera de voisinage de  $a$  **dans**  $X$ , de partie ouverte ou fermée **dans**  $X$  ou **de**  $X$ .

### 2. Applications continues.

Pour voir qu'une application continue d'un espace  $X$  dans un espace  $Y$  est continue en un point  $a$  de  $X$ , on se donne un voisinage  $V$  de  $b = f(a)$  arbitraire et on montre que son image réciproque  $f^{-1}(V)$ , ensemble des  $x$  tels que  $f(x) \in V$ , est un voisinage de  $a$ .

Cela revient à trouver un voisinage  $U$  de  $A$  tel que  $f(U) \subset V$ .

Dans le cas métrique que nous considérons, cela revient à se donner  $\epsilon > 0$  arbitraire et à trouver  $\eta > 0$  tel que l'image de la boule (ouverte ou fermée) de centre  $a$  et de rayon  $\eta$  soit incluse dans celle de centre  $b = f(a)$  et de rayon  $\epsilon$ .

On peut aussi se donner une suite  $x_n \rightarrow a$  et montrer que  $f(x_n) \rightarrow f(a)$ .

Cela se décline en termes de limites. Si  $a$  adhère à  $A$ , sans être dans  $A$ , et si  $l$  est dans  $Y$ , pour voir que  $f(x) \rightarrow l$  quand  $x \rightarrow a$  en restant dans  $A$ , on se donne un voisinage  $V$  de  $l$  arbitraire et on doit trouver un voisinage  $U$  de  $a$  tel que  $f(U \cap A) \subset V$ .

Cela vaut en particulier pour les suites : la convergence de  $x_n$  vers  $l$  est celle en  $\infty$  de la fonction sur  $\mathbf{N} \cup \infty$  définie par la suite et  $l$ . Ici les boules centrées à l'infini sont les intervalles  $[N, \infty]$ ; leurs traces sur  $\mathbf{N}$  sont les intervalles  $[N, \infty[$ .

Pour voir qu'une application continue d'un espace  $X$  dans un espace  $Y$  est continue, on montre qu'elle l'est en tout point de  $X$ . On peut aussi se donner une partie ouverte  $V$  de  $Y$  arbitraire et montrer que son image réciproque  $f^{-1}(V)$ , ensemble des  $x$  tels que  $f(x) \in V$ , est ouverte dans  $X$ .

### 3. Sous-espaces, produits.

On fait d'une partie  $Y$  d'un espace  $X$  un **sous-espace** en y considérant la distance induite par celle de  $d$ . Sa topologie est appelée *topologie induite*.

Une application à valeurs dans  $Y$  est continue si et seulement si on prend  $X$  comme espace d'arrivée; dans ce sens il n'y a pas de souci. Dans l'autre sens, une application définie sur  $X$  peut avoir une restriction à  $Y$  continue sans être pour autant continue en tout point de  $Y$ .

On retiendra que les parties ouvertes ou fermées du sous-espace  $Y$  sont les traces sur  $Y$  des parties ouvertes ou fermées de  $X$ .

La notion de sous-espace est très importante, mais piègeuse. Des lors qu'on se place dans le sous-espace  $Y$ , les autres points de  $X$  disparaissent du décor, ce qui le clarifie. Mais les parties ouvertes (resp. fermées) de  $Y$  sont rarement ouvertes ou fermées dans  $X$  : c'est cependant le cas lorsque  $Y$  est lui-même ouvert (resp. fermé) dans  $X$ .

L'**espace produit** de  $X, Y$  est défini par  $d((x, y), (x', y')) = \max(d(x, x'), d(y, y'))$ .

Une application à valeurs dans un produit est continue si et seulement si ses composantes les sont. Dans l'autre sens c'est plus délicat.

### 4. Compacité.

La notion de compacité est *absolue* ou *intrinsèque* : elle concerne un espace  $X$ . Elle ne s'applique que secondairement aux parties d'un espace : une partie est dite compacte si elle constitue un sous-espace compact. On dit donc "compact" tout court.

Dans le cas métrique, la compacité se définit à partir de l'une ou l'autre des propriétés équivalentes — comme résultat d'une démonstration sérieuse — qui suivent.

(Bolzano-Weierstrass) De toute suite on peut extraire une suite convergente.

(Heine-Borel-Lebesgue) De tout recouvrement ouvert on peut extraire un recouvrement fini.

Pour montrer qu'un espace est compact, on peut retenir ceci.

L'image continue d'un espace compact est compacte.

Un produit d'espaces compacts est compact.

Une partie fermée d'un espace compact est compacte.

On peut surtout noter cela.

**Théorème.** Dans un espace vectoriel de dimension finie  $E$ , par exemple dans  $\mathbf{R}^n$ , une partie est compacte si et seulement si elle est fermée et bornée.

Cependant, en calcul différentiel on se place rarement dans l'espace  $E$  tout entier. On se place souvent dans le sous-espace  $U$  défini par une partie ouverte. Pour montrer qu'une partie ouverte de  $U$  est compacte, on doit montrer qu'elle est fermée dans  $U$  — ce qui n'implique pas fermée dans  $E$  jusqu'ici — et qu'elle est bornée et à distance  $> 0$  du bord.

### 5. Utilisation de la compacité.

Deux grands théorèmes sont à retenir.

**Théorème de la borne supérieure.** Si  $K$  est un espace compact non vide, toute application continue de  $K$  dans  $\mathbf{R}$  est bornée et atteint ses bornes.

**Théorème de Heine ou de l'uniforme continuité.** Si  $f$  est une application continue d'un espace métrique  $K$  dans un espace métrique  $M$ , alors  $f$  est uniformément continue.

Citons encore une propriété utile : soit  $(F_n)$  est une suite décroissante de parties fermées de l'espace compact  $K$ ; si l'intersection est vide, l'une des parties  $F_n$  est vide.

## 6. Connexité.

Comme la compacité, la notion de connexité est *absolue* ou *intrinsèque* : elle concerne un espace  $X$ . Elle ne s'applique que secondairement aux parties d'un espace : une partie est dite connexe si elle constitue un sous-espace connexe. On dit donc "connexe" tout court.

La connexité se définit à partir de l'une autre des propriétés — directement — équivalentes qui suivent.

On ne peut pas partager l'espace en deux parties ouvertes non vides.

On ne peut pas partager l'espace en deux parties fermées non vides.

Il n'existe pas de partie ouverte et fermée autre que la partie vide et la partie pleine.

Pour montrer qu'un espace est connexe, on peut retenir ceci.

L'image continue d'un espace connexe est connexe.

Un produit d'espaces connexes est connexe.

L'adhérence d'une partie connexe est connexe.

On peut surtout noter cela.

**Théorème.** Les intervalles de  $\mathbf{R}$  sont connexes.

Ce sont — accessoirement — les seules parties connexes.

Ce théorème est le théorème des valeurs intermédiaires. Il montre en effet que l'image continue d'un intervalle  $[a, b]$  est connexe; si c'est une partie de  $\mathbf{R}$ , elle contiendra les points entre  $f(a)$  et  $f(b)$ .

Il fournit surtout un **critère de connexité**. Pour voir que  $X$  est connexe, il suffit de se donner deux points arbitraires  $a, b$  de  $X$  et de montrer qu'on peut les relier dans  $X$  par un chemin continu, i.e. par une application continue  $\phi : [\alpha, \beta] \rightarrow X$  telle que  $\phi(\alpha) = a, \phi(\beta) = b$ .

## 7. Utilisation de la connexité.

Pour montrer qu'une partie  $A$  de l'espace  $X$  est pleine, on montre qu'elle est ouverte, fermée et non vide.

## 8. Une précision importante.

On n'utilisera jamais le symbole  $\lim$  a priori, avant de connaître l'existence d'une limite, donc a fortiori pour tenter de l'établir.

En fait on ne raisonne jamais sur des limites avec ce symbole. On écrit une relation, égalité ou inégalité **large**, et on y passe à la limite.

On fera notamment très attention à l'écriture correcte des inégalités. En l'absence d'intention particulière, c'est toujours l'inégalité large qu'il faut prendre. L'usage de l'inégalité stricte doit être assorti d'un but précis; il doit être valide sans restriction; et on ne peut pas y passer à la limite — sauf en prenant l'inégalité large.

## Chapitre 6. Retour sur l'intégration

On considère l'intégrale de Lebesgue de fonctions définies sur  $\mathbf{R}$  ou  $\mathbf{R}^n$ ; l'intégrale d'une fonction définie sur une partie  $A$  s'obtiendra en prolongeant la fonction en lui donnant la valeur 0 hors de  $A$ .

A la notion d'intégrale, on relie celle de mesure d'une partie  $A$ ; on pose

$$\text{mes } A = \int 1_A$$

où  $1_A$  est la fonction indicatrice de  $A$ , qui vaut 1 sur  $A$  et 0 ailleurs.

### 1. Sommations infinies.

Si  $x_n$  est une série à termes positifs ou nuls, on peut toujours définir sa somme

$$\sum x_n$$

en lui attribuant éventuellement la valeur  $+\infty$ ; si la valeur est finie, la série est dite **convergente**; sinon elle est dite **divergente**. Dans tous les cas la somme est la borne supérieure des sommes finies.

Si  $x_n$  est une série réelle ou complexe, on n'utilisera jamais le symbole de sommation a priori. Mais on peut définir la somme  $\sum x_n$  lorsque

$$\sum |x_n| < +\infty$$

auquel cas la série est dite **absolument convergente**.

### 2. Intégrales.

Si  $f$  est une fonction à valeurs positives ou nulles, on peut toujours définir son intégrale

$$\int f$$

en lui attribuant éventuellement la valeur  $+\infty$ . Dans le cas où la valeur est cependant finie, on dit que  $f$  est **intégrable**.

Si  $f$  est une fonction à valeurs réelles ou complexes, on n'utilisera jamais le symbole d'intégration a priori. Cependant on définit l'intégrale  $\int f$  lorsque

$$\int |f| < +\infty$$

auquel cas on dit aussi que  $f$  est **intégrable**.

**N.B.** Ce n'est pas tout à fait correct. Il faudrait ajouter une hypothèse de mesurabilité. Cependant, en Analyse et contrairement à ce qui vaut en Probabilités, on s'est arrangé pour que la mesurabilité soit quasi-automatique. En effet considère au départ la tribu de Lebesgue, la plus grande possible, et à l'arrivée la tribu borélienne, la plus petite possible.

### 3. Critères d'intégrabilité.

Pour une fonction positive tous les calculs sont licites, notamment le passage par une primitive. Pour une fonction réelle ou complexe, on place une valeur absolue ou un module autour de la fonction, on majore et on calcule. Si le résultat est fini, la fonction est intégrable.

Si  $f$  est continue sur  $[a, +\infty[$ , elle est intégrable sur cet intervalle dès qu'elle est un  $O(1/x^\alpha)$  où  $\alpha > 1$  quand  $x \rightarrow +\infty$ .

En effet

$$\int_a^{+\infty} \frac{dx}{x^\alpha} = \frac{x^{\alpha-1}}{\alpha-1} \Big|_{x=a}^{x=+\infty} \quad \text{si } \alpha \neq 1 \quad \text{ou} \quad \log x \Big|_{x=a}^{x=+\infty} \quad \text{si } \alpha = 1$$

donne un résultat fini si et seulement si  $\alpha > 1$ .

Si  $f$  est continue sur  $]0, a]$ , elle est intégrable sur cet intervalle dès qu'elle est un  $O(1/x^\alpha)$  où  $\alpha < 1$  quand  $x \rightarrow 0$ .

En effet

$$\int_0^a \frac{dx}{x^\alpha} = \frac{x^{\alpha-1}}{\alpha-1} \Big|_{x=0}^{x=a} \quad \text{si } \alpha \neq 1 \quad \text{ou} \quad \log x \Big|_{x=0}^{x=a} \quad \text{si } \alpha = 1$$

donne un résultat fini si et seulement si  $\alpha < 1$ .

**Attention.** On parle d'intégrabilité *sur un intervalle*, jamais en un point.

### 4. Négligeabilité.

L'intégrale de Lebesgue dispose d'une notion à la fois puissante et troublante, celle de négligeabilité. Certaines parties auront une *mesure nulle*. On dira qu'elles sont négligeables parce qu'on peut y modifier à volonté la valeur d'une fonction sans modifier son intégrabilité et son intégrale.

On emploiera la locution "presque partout" pour dire "en dehors d'une partie négligeable". Les fonctions à intégrer n'auront besoin d'être définies que presque partout.

Voici quelques exemples de parties négligeables :

dans  $\mathbf{R}$  les points, l'ensemble des points d'une suite;

dans  $\mathbf{R}^2$  une courbe de classe  $\mathcal{C}^1$ ;

dans  $\mathbf{R}^3$  une surface de classe  $\mathcal{C}^1$ .

### 5. Un grand théorème.

On va donner deux énoncés, qui sont en fait deux variantes du même dans une situation un peu plus générale.

**Théorème d'intégration des séries.** Si les fonctions  $u_n$  sont à valeurs positives, ou si soit  $\sum \int |u_n| < +\infty$  soit  $\int \sum |u_n| < +\infty$ , alors

$$\int \sum u_n = \sum \int u_n$$

les deux membres ayant simultanément un sens.

**Théorème de Fubini.** Si la fonction de deux variables  $f(x, y)$  est à valeurs positives, ou intégrable, alors

$$\int \int f(x, y) dx dy = \int \left( \int f(x, y) dx \right) dy = \int \left( \int f(x, y) dy \right) dx$$

les trois expressions ayant simultanément un sens.

Du premier énoncé découlent les théorèmes dits de la convergence monotone et de la convergence dominée; on ne les citera pas.

## 6. Intégrales à paramètres.

On considère une fonction

$$F(x) = \int f(t, x) dt$$

définie par une intégrale à paramètre.

**Théorème.** Si  $f$  est continue en  $x$  et vérifie la condition de domination

$$|f(t, x)| \leq g(t)$$

où  $g$  est intégrable, alors  $F$  est continue.

**Théorème.** Si, de plus,  $f$  est dérivable en  $x$  et vérifie la condition de domination

$$\left| \frac{\partial f}{\partial x}(t, x) \right| \leq g(t)$$

où  $g$  est intégrable, alors  $F$  est dérivable et

$$F'(x) = \int \frac{\partial f}{\partial x}(t, x) dt .$$

On a donné le second énoncé parce que c'est l'usage. Cependant l'utilisation du théorème de Fubini pour dériver sous l'intégrale est bien souvent préférable.

## 7. Un cas important.

Ce qui précède comprend le cas élémentaire d'une fonction continue sur un segment ou sur une partie compacte de  $\mathbf{R}^n$ . Par exemple, si  $f$  est continue sur  $[a, b]$ , elle y est bornée par une constante  $M \geq 0$  et

$$\int |f| \leq (b - a)M < +\infty .$$

Le cas d'une intégrale à paramètres portant sur une partie compacte et pour laquelle  $f$ , comme  $\partial f / \partial x$  éventuellement, est continue en  $(t, x)$  entre aussi dans ce cadre. Par exemple si  $f$  est continue sur  $[a, b] \times [c, d]$ , on a

$$|f(t, x)| \cdot 1_{[a, b]}(t) \leq M \cdot 1_{[a, b]}(t)$$

où la fonction de droite est intégrable.

## Chapitre 7. Changement de variables

### 1. En dimension 1.

On commence, classiquement, par **intégrer des fonctions**. Etant donné un segment  $I$  de  $\mathbf{R}$  et une fonction continue  $f$  sur  $I$ , on définit

$$\int_I f(x)dx = \int \tilde{f}(x)dx$$

où  $\tilde{f}$  s'obtient en prolongeant  $f$  par 0 en dehors de  $I$ .

Lorsque  $I = [a, b]$ , où  $a \leq b$ , on note aussi cette intégrale

$$\int_a^b f(x)dx .$$

Cependant ce n'est pas exactement le concept **usuel**. On verra plus loin à quoi il faut rattacher ce dernier. Le plus souvent on se donne un intervalle  $I$  de  $\mathbf{R}$ , une fonction  $f$  sur  $I$  et deux points  $a, b$ ; on définit alors

$$\int_a^b f(x)dx$$

comme

$$\int_{[a,b]} f(x)dx$$

lorsque  $a \leq b$  et

$$- \int_{[b,a]} f(x)dx$$

dans le cas contraire. Il ne s'agit plus d'intégrer une fonction sur un ensemble.

C'est dans ce cadre qu'on énonce le théorème du **changement de variable**. Etant donnée une application  $\phi$  de classe  $\mathcal{C}^1$  d'un intervalle  $J$  dans  $I$  et des points  $\alpha, \beta$  tels que  $a = \phi(\alpha)$ ,  $b = \phi(\beta)$ , on a

$$\int_a^b f(x)dx = \int_\alpha^\beta f(\phi(u))\phi'(u)du .$$

La démonstration se fait en considérant une primitive  $F$  de  $f$ . Alors  $F \circ \phi$  est une primitive de  $(f \circ \phi) \cdot \phi'$  et le résultat s'en déduit.

On notera que l'existence d'une primitive est fournie par l'intégrale dépendant de la borne supérieure. Cependant on a aussi besoin de l'unicité, à une constante additive près. C'est le fait qu'une fonction de dérivée nulle sur un intervalle est constante; ici on a besoin du théorème des accroissements finis, au moins sous cette forme affaiblie.

Revenons à l'intégrale d'une **fonction**. Comment y traduire le changement de variable. On voudrait avoir une relation du genre

$$\int_I f(x)dx = \int_J f(\phi(u))\phi'(u)du .$$

lorsque  $I, J$  sont des segments. Il faudra déjà que  $\phi(J) = I$ . Mais il faudra encore que les bornes correspondent par  $\phi$ . Si c'est dans le bon ordre on aura bien l'égalité ci-dessus. Si c'est en changeant l'ordre on aura

$$\int_I f(x)dx = - \int_J f(\phi(u))\phi'(u)du .$$

Si  $\phi'$  garde un signe constant, en s'annulant peut-être, autrement dit si  $\phi$  est une surjection monotone, on aura

$$\int_I f(x)dx = \int_J f(\phi(u))|\phi'(u)|du$$

dans les deux cas. Voilà la formule du changement de variable pour l'intégrale d'une fonction.

Peut-on chercher une démonstration directe de cette égalité. On va supposer  $\phi$  croissante. Le second membre peut être approché par une somme de Riemann dy type

$$\sum_k f(\phi(\alpha_k))\phi'(\alpha_k)(\alpha_{k+1} - \alpha_k)$$

alors que le premier sera approché par la somme de Riemann

$$\sum_k f(\phi(\alpha_k))(\phi(\alpha_{k+1}) - \phi(\alpha_k)) .$$

Pour obtenir l'égalité il faudra rendre petite la différence

$$\phi(\alpha_{k+1}) - \phi(\alpha_k) - \phi'(\alpha_k)(\alpha_{k+1} - \alpha_k) .$$

Pour  $\alpha_{k+1}$  suffisamment proche de  $\alpha_k$ , on peut bien rendre cette différence inférieure à  $\epsilon(\alpha_{k+1} - \alpha_k)$ , par la définition même de la dérivée. Il reste cependant une question : peut-on joindre  $\alpha$  à  $\beta$  par une chaîne de  $\alpha_k$  ayant cette propriété? C'est la compacité de  $[\alpha, \beta]$  qui va le permettre.

Au passage on remarquera que le théorème des accroissements finis, au moins pour une fonction continûment dérivable, résultera de cette étude.

Cela vaut si  $\phi'$  ne s'annule pas, autrement dit lorsque  $\phi$  est essentiellement un  $\mathcal{C}^1$  difféomorphisme de  $J$  sur  $I$ . C'est sous cette forme qu'on généralise le théorème en dimension supérieure, mais il faut se placer sur des parties ouvertes de  $\mathbf{R}^n$ . Pour cette raison on considère l'intégrabilité au sens de Lebesgue.

## 2. En dimension quelconque.

**Théorème.** Soient  $U, V$  des parties ouvertes de  $\mathbf{R}^n$  et  $\phi$  un  $\mathcal{C}^1$ -difféomorphisme de  $U$  sur  $V$ ; pour une fonction  $f$  sur  $V$  on a

$$\int_V f(y)dy = \int_U f(\phi(x))|\det d\phi(x)|dx$$

chaque fois que l'un des deux membres a un sens, auquel cas l'autre en a également un.

Le résultat est une conséquence immédiate de l'énoncé suivant.

**Lemme.** Soit  $U$  une partie ouverte de  $\mathbf{R}^n$  et  $\phi$  une application de classe  $\mathcal{C}^1$  de  $U$  dans  $\mathbf{R}^n$ ; pour une fonction (mesurable) positive  $f$  sur  $\mathbf{R}^n$  on a

$$\int_{\phi(U)} f(y)dy \leq \int_U f(\phi(x))|\det d\phi(x)|dx .$$

### Démonstration.

On commence par considérer une fonction  $f$  continue positive à support compact dans  $\mathbf{R}^n$ , une partie compacte  $K$  de  $U$  et une fonction  $g$  continue à support compact dans  $U$ , à valeurs dans  $[0, 1]$  et égale à 1 dans un voisinage  $L$  de  $K$ . Il s'agit de montrer que

$$\int_{\phi(K)} f(y)dy \leq \int_U g(x)f(\phi(x))|\det d\phi(x)|dx .$$

Cette dernière inégalité exprime en effet la majoration de la mesure de Lebesgue sur  $\phi(K)$  par l'image de la mesure de densité  $g(x)|\det d\phi(x)|$ . Elle s'étend alors au cas où  $f$  est (mesurable) positive : on remplace  $g$  par 1 et  $\phi(K)$  par  $\phi(U)$  en passant à la limite monotone .

On va se limiter à  $n = 2$  et munir  $\mathbf{R}^2$  de la norme produit pour simplifier l'exposé. On suppose le support de  $g(x)f(\phi(x))$  inclus dans le carré  $[-M, M]^2$ .

On se donne  $\epsilon \in ]0, 1]$  et on découpe le carré précédent en carrés  $C_a = a + C$  de demi-côté  $\eta = M/k$  et de centre  $a$ , où  $\eta > 0$  est assez petit pour que tout  $C_a$  rencontrant  $K$  soit inclus dans  $L$ .

**Principe.** Une majoration approchée du membre de gauche de la relation à établir est donnée par une somme de termes du type

$$f(\phi(a))\text{mes}(\phi(C_a)) .$$

- On aura pris les carrés assez petits, pour que leurs images soient elles-mêmes assez petites et pour que la fonction  $f$  puisse y être approchée par sa valeur en  $\phi(a)$ .

- On aura noté que les  $\phi(C_a)$  recouvrent  $\phi(K)$ , mais peuvent se chevaucher; c'est pourquoi on parle de majoration approchée.

Une approximation du membre de droite est donnée par une somme (de Riemann) de termes du type

$$f(\phi(a))|\det d\phi(a)|\text{mes}(C_a) .$$

Il va falloir comparer les deux expressions obtenues, en majorant, à petite erreur près, la première par la seconde, c'est-à-dire  $\text{mes}(\phi(C_a))$  par  $|\det d\phi(a)|\text{mes}(C_a)$ . Si les carrés sont assez petits, l'application  $\phi$  est proche sur  $C_a$  de l'application affine donnée par  $\phi(a) + d\phi(x - a)$ . Cette dernière transforme le carré en parallélogramme d'aire  $|\det d\phi(a)|\text{mes}(C_a)$ . En fait il faut agrandir le parallélogramme d'une bordure de petite largeur pour obtenir l'inclusion exacte. Cependant l'erreur alors commise est petite par rapport à l'aire de  $C_a$ .

Cela étant il faut prendre quelques précautions.

- Il faut s'assurer qu'on a un contrôle uniforme de l'erreur. C'est là qu'intervient la compacité, par le théorème de Heine pour les valeurs de  $f$  et par le théorème de la borne supérieure pour la norme de  $d\phi(x)$ .

- Il faut voir qu'on ajoute chaque fois  $k \times k$  termes d'erreur. On doit donc s'assurer que ce sont des erreurs en  $\epsilon/k^2$ .

**Mise en forme.** Pour  $x$  dans un carré  $C_a$  inclus dans  $U$ , on écrit

$$\phi(x) = \phi(a) + df(a).(x - a) + \int_0^1 [d\phi(a + t(x - a)) - d\phi(a)].(x - a)dt$$

ce qui donne d'abord

$$\|\phi(x) - \phi(a)\| \leq K\eta$$

si  $K$  majore  $\|d\phi(x)\|$ , puis

$$\|\phi(x) - \phi(a) - df(a).(x - a)\| \leq \epsilon\eta$$

si  $\eta > 0$  est choisi assez petit pour que le crochet soit inférieur à  $\epsilon$  en norme, et ce pour  $\eta > 0$  assez petit, comme le permet la continuité uniforme de  $d\phi$ .

Maintenant, si on remplace  $f(y)$  par  $f(\phi(a))$  sur  $\phi(C_a)$ , on commet une erreur  $\leq \epsilon$ , dès que  $\eta > 0$  est assez petit, par continuité uniforme de  $f$ .

De même, si on remplace  $g(x)|\det d\phi(x)|f(x)$  par  $g(a)|\det d\phi(a)|f(a)$  sur  $C_a$ , on commet une erreur  $\leq \epsilon$ , dès que  $\eta > 0$  est assez petit, par continuité uniforme de  $g(x)|\det d\phi(x)|f(x)$ .

En additionnant et majorant les erreurs ainsi commises, on obtient d'une part  $k^2.\epsilon\text{mes}(C_a) = 4\epsilon M^2$ , et d'autre part  $k^2.\epsilon\text{mes}(\phi(C_a)) \leq 4\epsilon K^2 M^2$ .

Il reste à comparer les mesures (les surfaces ici) de  $\phi(C_a)$  et de  $d\phi(a)(C)$ , soit  $|\det d\phi(a)|.\text{mes}(C)$ . Evaluant la mesure d'une allée bordant le parallélogramme  $d\phi_a(C)$  tracée avec un carré centré sur le bord et de demi-côté  $\epsilon\eta$ , on obtient facilement une majoration par  $4K'\epsilon\eta^2$ , d'où  $16\epsilon K' M^2$  en additionnant. Le reste suit.

### 3. Le théorème de Sard.

Comme conséquence du lemme, on a l'énoncé suivant : l'image de l'ensemble critique d'une application  $\phi$  de classe  $C^1$  de  $\mathbf{R}^n$  dans lui-même est de mesure nulle.

Un point est dit critique si la différentielle n'y est pas de rang maximum; ici cela signifie qu'elle n'est pas inversible. Si  $C$  est l'ensemble critique, on a simplement

$$\int 1_{\phi(C)}(y)dy \leq \int 1_C(x).0.dx = 0 .$$

#### 4. Partitions de l'unité.

**Proposition.** Etant donné un recouvrement ouvert fini  $(U_i)$  d'un espace métrique compact  $X$ , on peut trouver une famille  $(\phi_i)$ , où chaque  $\phi_i$  est une fonction positive ou nulle à support compact inclus dans  $U_i$ , telle que

$$\sum_i \phi_i = 1$$

sur  $X$ . On dit que la famille  $(\phi_i)$  est une partition continue de l'unité subordonnée au recouvrement  $(U_i)$ .

**Démonstration.** On prend d'abord

$$\psi_i(x) = d(x, U_i^c) .$$

Puisque  $\psi_i(x) > 0$  pour  $x$  dans  $U_i$ , on a

$$\sup_i \psi_i > 0$$

partout. Par le théorème de la borne inférieure, on peut trouver  $m > 0$  tel que

$$\sup_i \psi_i \geq m .$$

Soit alors

$$\theta_i = (\psi_i - m/2)^+ = \max(0, \psi_i - m/2) .$$

Clairement le support de  $\theta_i$  est inclus dans l'ensemble  $\{\psi_i \geq m/2\}$  donc dans  $U_i$ . De plus en chaque  $x$  l'un des  $\theta_i(x)$  est non nul : en effet l'un des  $\psi_i(x)$  est  $> m/2$ . Alors

$$\phi_i(x) = \frac{\theta_i}{\sum_j \theta_j}$$

donne la partition cherchée.

**Remarque.** En l'absence de compacité, on sera moins exigeant sur les supports.

Etant donné un recouvrement ouvert fini  $(U_i)$  d'un espace métrique  $X$ , on peut trouver une famille  $(\phi_i)$ , où chaque  $\phi_i$  est une fonction positive ou nulle et nulle en dehors de  $U_i$ , telle que

$$\sum_i \phi_i = 1$$

sur  $X$ .

La démonstration est la même : on passe directement des  $\psi_i$  aux  $\phi_i$ .

**Remarque.** Lorsqu'on remplace  $X$  par une partie ouverte  $U$  de  $\mathbf{R}^n$ , on peut imposer aux  $\phi_i$  d'être de classe  $\mathcal{C}^\infty$ . Mais c'est plus compliqué.

#### 5. Cas d'un $\mathcal{C}^1$ -difféomorphisme de $U$ sur $V$ en dimension 2.

On suppose que  $\Phi : U \rightarrow V$  est le  $\mathcal{C}^1$ -difféomorphisme donné par

$$\begin{cases} x = \phi(u, v) \\ y = \psi(u, v) \end{cases} .$$

Soit  $U_1$  la partie ouverte de  $U$  où  $\frac{\partial\phi}{\partial u} \neq 0$  et  $U_2$  celle où  $\frac{\partial\phi}{\partial v} \neq 0$ . Clairement  $U$  est la réunion de  $U_1$  et de  $U_2$ . On peut trouver des fonctions  $g$  et  $h$  continues dans  $V$  à valeurs dans  $[0, 1]$  telles que  $g + h = 1$ , que  $g$  soit nulle en dehors de l'image  $V_1$  de  $U_1$  et  $h$  en dehors de l'image  $V_2$  de  $U_2$ . Soit maintenant  $f$  une fonction à support compact dans  $V$ . En écrivant  $f$  sous la forme  $fg + fh$ , on se ramène au cas où  $f$  est à support compact dans  $V_1$  ou  $V_2$ .

Plaçons-nous dans le premier cas. Pour  $v$  fixé, l'application qui à  $u$  associe  $\phi(u, v)$  est inversible, son inverse associant à  $x$  la valeur  $\theta(x, v)$ , où  $\theta$  est de classe  $\mathcal{C}^1$ .

On a ainsi décomposé le difféomorphisme donné en deux, donnés par

$$\begin{cases} x = \phi(u, v) \\ v = v \end{cases}, \quad \begin{cases} x = x \\ y = \psi(\theta(x, v), v) \end{cases}$$

où, chaque fois, une seule des variables est changée. On se ramène à faire la démonstration dans un tel cas, par exemple dans le premier.

Calculons

$$\int \int h(x, v) dx dv = \int dv \int h(x, v) dx$$

en posant  $x = \phi(u, v)$  à  $v$  fixé. Il vient

$$dx = \frac{\partial\phi}{\partial u} du$$

et on obtient

$$\int dv \int h(\phi(u, v), v) \left| \frac{\partial\phi}{\partial u}(u, v) \right| du.$$

Il n'y a plus qu'à remarquer que  $\frac{\partial\phi}{\partial u}$  est le déterminant jacobien de la transformation.

## 6. Images directes.

Soient  $A$  une partie borélienne de  $\mathbf{R}^n$  et  $\phi$  une application borélienne de  $A$  dans  $\mathbf{R}^p$ . Si  $\mu$  est une mesure bornée sur  $\mathbf{R}^n$  portée par  $A$ , on lui associe la mesure  $\nu$  sur  $\mathbf{R}^p$  définie par

$$\nu(B) = \mu(\phi^{-1}(B))$$

pour toute partie borélienne  $B$ , ou par

$$\int f(y) d\nu(y) = \int f(\phi(x)) d\mu(x)$$

pour toute fonction  $f$  continue à support compact sur  $\mathbf{R}^p$ . Cette mesure  $\nu$  est dite l'image (directe) de  $\mu$  par  $\phi$ .

On a alors

$$\int f(y) d\nu(y) = \int f(\phi(x)) d\mu(x)$$

pour toute fonction  $f$  mesurable, chaque fois que l'un des membres a un sens, auquel l'autre membre en a également un. Par exemple pour  $f$  mesurable positive sur  $\mathbf{R}^p$ .

Le théorème de ce chapitre peut s'énoncer en disant que la mesure  $dy$  est l'image de la mesure  $|\det d\phi(x)| dx$ .

## Chapitre 8. Formes différentielles

On désigne dans toute la suite par  $E$  un espace vectoriel réel de dimension  $n$ . Y choisir une base  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$  revient à se donner un isomorphisme  $\mathbf{R}^n \rightarrow E$ . Dans ce cas on fera comme si  $E$  était l'espace  $\mathbf{R}^n$  lui-même, avec sa base canonique  $\mathbf{e}_1 = (1, 0, \dots, 0), \dots$

On désigne par  $E^*$  le dual de  $E$ . Une base étant choisie, il lui correspond la base duale définie par  $e_i^*(\mathbf{e}_j) = \delta_{ij}$ ;  $E$  et  $E^*$  sont alors identifiés à  $\mathbf{R}^n$  et on a  $e_i^*(\mathbf{h}) = h_i$ .

Une base étant choisie, on ne notera pas  $e_i^*$  les éléments de la base duale, mais plutôt  $x_i$  ou  $dx_i$ , sachant que

$$dx_i(\mathbf{h}) = h_i .$$

Réglons encore un problème d'indices. Si des  $h_j$  sont des vecteurs de  $\mathbf{R}^n$ , on notera  $h_{ij}$  la composante de  $h_j$  suivant  $\mathbf{e}_i$ . Ainsi aura-t-on

$$h_j = \sum_i \mathbf{e}_i h_{ij} .$$

Maintenant, pour plus de clarté, on écrit les vecteurs en gras.

### 1. Formes multilinéaires alternées.

Soit  $p$  un nombre entier positif ou nul.

**Définition.** Une forme  $p$ -linéaire  $S$  sur  $E$  est dite *alternée* si

$$S(\mathbf{h}_1, \dots, \mathbf{h}_p) = 0$$

dès que  $\mathbf{h}_i = \mathbf{h}_j$  pour un couple  $(i, j)$  tel que  $i \neq j$ .

Cela implique

$$S(\mathbf{h}_1, \dots, \mathbf{h}_i, \dots, \mathbf{h}_j, \dots, \mathbf{h}_p) = -S(\mathbf{h}_1, \dots, \mathbf{h}_j, \dots, \mathbf{h}_i, \dots, \mathbf{h}_p)$$

et

$$S(\mathbf{h}_{\sigma(1)}, \mathbf{h}_{\sigma(2)}, \dots, \mathbf{h}_{\sigma(p)}) = \epsilon(\sigma) S(\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_p)$$

où  $\epsilon(\sigma)$  est la signature de la permutation  $\sigma$  des nombres entiers  $1, 2, \dots, p$ .

On note  $\bigwedge^p E^*$  l'espace vectoriel des formes  $p$ -linéaires alternées sur  $E$ .

#### Exemples.

Pour  $p = 0$ , les formes alternées sont les fonctions n'ayant aucun argument, donc les éléments du corps de base  $\mathbf{R}$ .

Pour  $p = 1$ , les formes linéaires sont toutes alternées.

Pour  $p = 2$ , la condition se réduit à

$$S(\mathbf{h}, \mathbf{h}) = 0$$

ou

$$S\mathbf{h}, \mathbf{k} = -S(\mathbf{k}, \mathbf{h}) .$$

## 2. Antisymétrisation.

Voici un procédé pour construire des formes alternées. Soient  $f_1, f_2, \dots, f_p$  des éléments de  $E^*$ . On définit la forme  $p$ -linéaire alternée, dite décomposable

$$f_1 \wedge f_2 \wedge \dots \wedge f_p$$

en antisymétrisant l'expression

$$f_1(\mathbf{h}_1) \cdot f_2(\mathbf{h}_2) \cdot \dots \cdot f_p(\mathbf{h}_p)$$

pour obtenir

$$\sum_{\sigma} \epsilon(\sigma) f_1(\mathbf{h}_{\sigma(1)}) \cdot f_2(\mathbf{h}_{\sigma(2)}) \cdot \dots \cdot f_p(\mathbf{h}_{\sigma(p)})$$

sur les vecteurs  $\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_p$ , où  $\sigma$  parcourt l'ensemble des permutations des nombres entiers  $1, 2, \dots, p$ .

La propriété pour  $S$  d'être alternée est facile à vérifier.

### Exemples.

Pour  $p = 2$ , une base étant choisie, on a

$$(dx_1 \wedge dx_2)(\mathbf{h}_1, \mathbf{h}_2) = h_{11}h_{22} - h_{12}h_{21}$$

notamment.

Pour  $p = n$ , une base étant choisie, la forme

$$dx_1 \wedge dx_2 \wedge \dots \wedge dx_n$$

est le déterminant dans la base, encore appelé forme volume. Elle vérifie

$$(dx_1 \wedge dx_2 \wedge \dots \wedge dx_n)(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n) = 1$$

sur la base; en particulier elle est non nulle.

**Proposition.** Une base étant choisie, les formes  $p$ -linéaires alternées du type  $dx_{\phi(1)} \wedge dx_{\phi(2)} \wedge \dots \wedge dx_{\phi(p)}$ , où  $\phi : [1, p] \rightarrow [1, n]$  est une application *strictement croissante* entre intervalles de  $\mathbf{N}$ , forment une base de  $\bigwedge^p E^*$ .

Faisons la démonstration pour  $n = 3$  et  $p = 2$  à titre d'exemple. On a

$$\begin{aligned} S(h_1\mathbf{e}_1 + h_2\mathbf{e}_2 + h_3\mathbf{e}_3, k_1\mathbf{e}_1 + k_2\mathbf{e}_2 + k_3\mathbf{e}_3) \\ = (h_1k_2 - h_2k_1)S(\mathbf{e}_1, \mathbf{e}_2) + (h_2k_3 - h_3k_2)S(\mathbf{e}_2, \mathbf{e}_3) + (h_3k_1 - h_1k_3)S(\mathbf{e}_3, \mathbf{e}_1) \end{aligned}$$

d'où

$$S = \lambda_1 dx_1 \wedge dx_2 + \lambda_2 dx_2 \wedge dx_3 - \lambda_3 dx_1 \wedge dx_3 .$$

Ainsi a-t-on un système générateur. Il est également libre. Donnons-nous une relation

$$0 = \lambda_1 dx_1 \wedge dx_2 + \lambda_2 dx_2 \wedge dx_3 - \lambda_3 dx_1 \wedge dx_3 .$$

En se limitant aux vecteurs dont la troisième composante est nulle, identifiables à des vecteurs de  $\mathbf{R}^2$ , on obtient

$$0 = \lambda_1 dx_1 \wedge dx_2 .$$

On reconnaît la forme déterminant dans la base canonique, laquelle n'est pas nulle, comme on le vérifie sur des vecteurs de la base. Ainsi  $\lambda_1 = 0$ . Il en est de même pour les autres coefficients.

### Exemples.

Pour  $p > n$  l'espace  $\bigwedge^p E^*$  est nul.

L'espace  $\bigwedge^n E^*$  est de dimension 1; une base étant choisie, il est engendré par le déterminant dans la base.

### 3. Effet d'une application linéaire.

Une application linéaire  $u$  de  $F$  dans  $E$  définit naturellement une application linéaire de  $\bigwedge^p E^*$  dans  $\bigwedge^p F^*$  : à la forme  $p$ -linéaire alternée  $S$  sur  $E$ , cette dernière associe la forme  $p$ -linéaire alternée sur  $F$  valant

$$S(u(\mathbf{h}_1), u(\mathbf{h}_2), \dots, u(\mathbf{h}_n))$$

en  $\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_n$ .

En particulier, un endomorphisme  $u$  de  $E$  définit un endomorphisme de  $\bigwedge^n E^*$ ; puisque cet espace est de dimension 1, c'est la multiplication par un scalaire  $\lambda$ .

**Définition.** Le coefficient  $\lambda$  ci-dessus est le déterminant  $\det u$  de l'endomorphisme  $u$ .

On notera que cette définition est indépendante du choix d'une base. Mais, une base étant choisie, en appliquant la définition à  $S = dx_1 \wedge dx_2 \wedge \dots \wedge dx_n$ , on constate que le déterminant de  $u$  est encore le déterminant des vecteurs colonne dans la base.

### 4. Produit extérieur.

**Proposition.** Il existe une application bilinéaire et une seule de  $\bigwedge^p E^* \times \bigwedge^q E^*$  dans  $\bigwedge^{p+q} E^*$ , qui à  $S$  et  $T$  associe le produit extérieur  $S \wedge T$  de  $S$  et  $T$ , vérifiant

$$(f_1 \wedge \dots \wedge f_p) \wedge (g_1 \wedge \dots \wedge g_q) = f_1 \wedge \dots \wedge f_p \wedge g_1 \wedge \dots \wedge g_q$$

si  $f_1, \dots, f_p, g_1, \dots, g_q$  sont des formes linéaires.

On a alors

$$T \wedge S = (-1)^{pq} S \wedge T .$$

L'unicité se voit facilement en prenant une base. L'existence se montre à partir des formules obtenues. On peut de cette façon interpréter les formes alternées décomposables comme des produits extérieurs de formes linéaires.

### 5. Orientation.

Rappelons que l'espace  $\bigwedge^n E^*$  est de dimension 1, donc une droite. En revanche il ne possède pas de vecteur de base canonique.

Une orientation d'un espace vectoriel  $E$  sur  $\mathbf{R}$  de dimension 1 est le choix d'une demi-droite vectorielle parmi les deux possibles. Les bases positives sont les vecteurs non nuls de la demi-droite choisie.

Bien sûr le corps de base  $\mathbf{R}$  est canoniquement orienté par la demi-droite positive.

On constate aussitôt qu'orienter  $E$  et orienter  $E^*$  reviennent au même : la compatibilité entre leurs orientations est donnée par la relation  $f^*(e) > 0$  entre une base positive  $f^*$  de  $E^*$  et une base positive  $e$  de  $E$ .

Une orientation d'un espace vectoriel réel de dimension  $n$  peut être définie comme une orientation de l'espace  $\bigwedge^n E^*$ . Une base  $e_1, \dots, e_n$  est positive si  $f_1^*(e_1) \cdots f_n^*(e_n) > 0$  dès lors que  $f_1^* \wedge \dots \wedge f_n^* > 0$ .

Choisir une base oriente l'espace, car  $\bigwedge^n \mathbf{R}^*$  est orienté par  $dx_1 \wedge \dots \wedge dx_n$ .

Enfin un endomorphisme préserve l'orientation quand son déterminant est positif.

## 6. Formes différentielles.

Soit maintenant  $U$  une partie ouverte de  $E$ .

**Définition.** Une  $p$ -forme différentielle sur  $U$  est une application de  $U$  dans  $\bigwedge^p E^*$ .

L'espace vectoriel des  $p$ -formes sur  $U$  est noté  $\Lambda^p(U)$ .

### Exemples.

Une 0-forme est simplement une fonction  $f$ .

Une base étant choisie, une 1-forme s'écrit

$$f_1 dx_1 + f_2 dx_2 + \cdots + f_n dx_n$$

où les  $f_i$  sont des fonctions. La différentielle d'une fonction est une telle forme; on verra qu'il y en a d'autres.

Une base étant choisie, une  $n$ -forme s'écrit

$$f dx_1 \wedge dx_2 \wedge \cdots \wedge dx_n$$

où  $f$  est une fonction.

Une  $p$ -forme peut s'écrire

$$\sum f_\phi dx_{\phi(1)} \wedge dx_{\phi(2)} \wedge \cdots \wedge dx_{\phi(p)}$$

où  $\phi$  parcourt les applications strictement croissantes de  $[1, p]$  dans  $[1, n]$  et où chaque  $f_\phi$  est une fonction.

En dimension 3, une base étant choisie, en notant  $x, y, z$  les fonctions coordonnées, une 2-forme peut aussi s'écrire

$$f dx \wedge dy + g dy \wedge dz + h dz \wedge dx$$

où  $f, g, h$  sont des fonctions.

Le **produit extérieur** de deux formes différentielles se définit, en chaque point, en prenant le produit extérieur des formes alternées.

## 7. Différentielle extérieure.

Désormais, les  $p$ -formes que nous considérerons seront toujours supposées suffisamment régulières, les coefficients étant, s'il le faut, des fonctions indéfiniment dérivables.

On se place sur l'espace vectoriel  $\Lambda(U)$  somme des  $\Lambda^p(U)$ .

**Définition.** Il existe une application linéaire et une seule de  $\Lambda(U)$  dans lui-même, notée  $d$  et appelée différentielle extérieure, ayant les propriétés suivantes :

- (i) Si  $f$  est une fonction, alors  $df$  est sa différentielle.
- (ii)  $d(\omega \wedge \omega') = d\omega \wedge \omega' + (-1)^{\deg \omega} \omega \wedge d\omega'$ .
- (iii)  $d(d\omega) = 0$ .

L'unicité est facile; on montre en même temps que  $d$  envoie  $\Lambda^p(U)$  dans  $\Lambda^{p+1}(U)$ . En effet, une base étant choisie, on a

$$d(f dx_{\phi(1)} \wedge \cdots \wedge dx_{\phi(p)}) = df \wedge dx_{\phi(1)} \wedge \cdots \wedge dx_{\phi(p)}$$

d'après (ii) et (iii), où  $df$  a été précisé par (i).

L'existence se montre à partir de la formule ci-dessus. On commence par établir les propriétés suivantes.

$$(ii)_a \quad d(f\omega) = df \wedge \omega + f d\omega.$$

$$(iii)_a \quad d(df_1 \wedge \cdots \wedge df_p) = 0 \text{ si } f_1, \dots, f_p \text{ sont des fonctions.}$$

La propriété (iii)<sub>a</sub> résulte par exemple de  $d(df) = 0$ , ce qui provient du théorème de Schwarz.

**Remarque.** Les propriétés (i), (ii)<sub>a</sub> et (iii)<sub>a</sub> suffisent à caractériser  $d$ .

La différentielle extérieure définit une suite

$$\Lambda^0(U) \rightarrow \Lambda^1(U) \rightarrow \cdots \rightarrow \Lambda^n(U) \rightarrow 0$$

dans laquelle le produit de deux flèches consécutives est nul.

On dit qu'une forme  $\omega$  est exacte si elle s'écrit  $d\eta$ . Une condition nécessaire pour cela est qu'elle soit *fermée*, i.e. que  $d\omega = 0$ .

Cette condition n'est pas suffisante. En revanche on a l'énoncé suivant.

**Théorème.** si  $U$  est étoilé par rapport à un de ses points, alors toute forme fermée sur  $U$  est exacte.

On parlera de la démonstration plus tard.

## 8. Cas euclidien, champ de vecteurs, gradient.

Si  $E$  est muni d'une structure euclidienne, son dual  $E^*$  s'identifie à lui-même; on peut en effet associer canoniquement vecteurs et formes linéaires par l'application qui à  $\mathbf{v}$  associe la forme linéaire  $\langle \mathbf{v}, \cdot \rangle$  qui en  $\mathbf{h}$  vaut  $\langle \mathbf{v}, \mathbf{h} \rangle$ .

C'est déjà le cas, si l'on s'est donné une base. En effet l'espace  $\mathbf{R}^n$  dispose d'une structure euclidienne canonique définie par

$$\langle \mathbf{h}, \mathbf{k} \rangle = \sum_i h_i k_i .$$

On retrouve ainsi l'identification déjà considérée.

Ainsi, dans un espace euclidien, une 1-forme différentielle  $\omega = \langle \mathbf{v}, \cdot \rangle$  sur  $U$  est-elle associée à une fonction  $\mathbf{v}$  sur  $U$  à valeurs dans  $E$ , appelée *champ de vecteurs*, par

$$\omega(\mathbf{h}) = \langle \mathbf{v}, \mathbf{h} \rangle .$$

En particulier la différentielle  $df$  de  $f$  est associée au gradient  $\mathbf{grad} f$  par

$$df(\mathbf{h}) = \langle \mathbf{grad} f, \mathbf{h} \rangle .$$

## 9. Cas euclidien orienté, rotationnel, divergence.

Désormais nous nous limitons au cas de la dimension 3. La forme volume

$$e^* = dx_1 \wedge dx_2 \wedge dx_3$$

est définie sans ambiguïté dès lors qu'on s'est donné à la fois une structure euclidienne et une orientation.

Elle est en effet préservée au signe près par un endomorphisme orthogonal. Le signe lui-même est préservé par l'orientation.

La forme  $e^*$  fait correspondre à un scalaire  $f$  l'élément  $fe^*$  de  $\bigwedge^3 E^*$ .  
Ainsi fait-elle correspondre à une fonction  $f$  la 3-forme  $\phi(f) = fe^*$ .

Elle fait encore correspondre à un vecteur  $\mathbf{v}$  de  $E$  un élément  $i_{\mathbf{v}}e^*$  de  $\bigwedge^2 E^*$ , produit intérieur gauche de  $\mathbf{v}$  et  $e^*$ , par la formule

$$i_{\mathbf{v}}e^*(\mathbf{h}, \mathbf{k}) = e^*(\mathbf{v}, \mathbf{h}, \mathbf{k}) \quad (1)$$

où  $\mathbf{h}, \mathbf{k}$  sont des vecteurs arbitraires.

Ainsi fait-elle correspondre à un champ de vecteurs  $\mathbf{v}$  une 2-forme  $\phi(\mathbf{v}) = i_{\mathbf{v}}e^*$ .

Maintenant si  $\mathbf{v}$  est un champ de vecteurs, la différentielle extérieure de la 1-forme associée, correspond à un nouveau champ de vecteurs, le champ *rotationnel*  $\mathbf{rot} \mathbf{v}$  défini par  $d\langle \mathbf{v}, \cdot \rangle = \phi(\mathbf{rot} \mathbf{v})$ .

Par ailleurs, si  $\mathbf{v}$  est un champ de vecteurs, la différentielle extérieure de la 2-forme correspondante elle-même correspond à une fonction scalaire, la fonction *divergence*  $\text{div} \mathbf{v}$  définie par  $d\phi(\mathbf{v}) = \phi(\text{div} \mathbf{v})$ .

## 10. Formes vectorielles.

Ce qu'on vient de faire s'étend aussitôt lorsque les coefficients d'une forme sont à valeurs vectorielles. Désormais, à la manière des physiciens, on utilisera le point  $\cdot$  pour le produit scalaire plutôt que les pointes  $\langle \cdot, \cdot \rangle$ ,  $\mathbf{e}_x$  plutôt que  $\mathbf{e}_1$ ,  $dv$  plutôt que  $e^*$  etc.

Par exemple, en dimension 3, on considèrera l'élément vectoriel de longueur qui s'écrit

$$d\mathbf{M} = \mathbf{e}_x dx + \mathbf{e}_y dy + \mathbf{e}_z dz$$

dans une base, et, dans le cas euclidien orienté, l'élément vectoriel de surface qui s'écrit

$$d\mathbf{S} = \mathbf{e}_x dy \wedge dz + \mathbf{e}_y dz \wedge dx + \mathbf{e}_z dx \wedge dy$$

dans une base orthonormée positive. Ce dernier permet d'écrire

$$\phi(\mathbf{v}) = \mathbf{v} \cdot d\mathbf{S} \quad (2)$$

ce qui le caractérise indépendamment du choix de la base.

On notera alors que

$$df = \mathbf{grad} f \cdot d\mathbf{M}$$

$$d(\mathbf{v} \cdot d\mathbf{M}) = \mathbf{rot} \mathbf{v} \cdot d\mathbf{S}$$

$$d(\mathbf{v} \cdot d\mathbf{S}) = \text{div} \mathbf{v} \, dv .$$

**Exercice.** Relier les propriétés (1), (2) et vérifier que  $d\mathbf{S}(\mathbf{h}, \mathbf{k})$  définit le produit vectoriel  $\mathbf{h} \wedge \mathbf{k}$ .

## Chapitre 9. Intégration des formes différentielles

### 1. Image inverse.

Soient  $V, U$  des parties ouvertes des espaces vectoriels réels  $F, E$  de dimensions respectives  $m$  et  $n$  et  $\phi$  une application de classe  $\mathcal{C}^1$  de  $V$  dans  $U$ .

En chaque point  $y$  de  $V$  la différentielle  $d\phi(y)$  est une application linéaire de  $F$  dans  $E$ ; elle définit donc une application linéaire de  $\bigwedge^p E^*$  dans  $\bigwedge^p F^*$  pour tout  $p$  et on obtient ainsi une application linéaire de  $\Lambda(U)$  dans  $\Lambda(V)$ , appelée *image inverse* par  $\phi$  et notée  $\phi^*$ .

**Proposition.** L'image inverse  $\phi^*$  est l'unique application linéaire de  $\Lambda(U)$  dans  $\Lambda(V)$  vérifiant

(i)

$$\phi^* f = f \circ \phi .$$

(ii)

$$\phi^*(\omega \wedge \omega') = \phi^*\omega \wedge \phi^*\omega' .$$

(iii)

$$\phi^*(d\omega) = d(\phi^*\omega)$$

**N.B.** On peut remplacer (iii) par

(iii)<sub>a</sub>  $\phi^*(df) = d(\phi^*f)$ .

Les vérifications sont faciles. On retiendra la règle suivante : si  $x = \phi(u)$ , pour calculer  $\phi^*\omega$ , on exprime, dans les coefficients, les coordonnées  $x$  en fonction des coordonnées  $u$  et on remplace les éléments différentiels  $dx_i$  par leur différentielle en  $u$ . Formellement rien ne change.

**Proposition.**

$$\psi^*(\phi^*\omega) = (\phi \circ \psi)^*\omega .$$

C'est la théorème de composition des différentielles appliqué à  $x = \phi(u)$  et  $u = \psi(v)$ .

### 2. Sous-variétés orientées.

Soit  $V$  une sous-variété de dimension  $p$  de  $\mathbf{R}^n$ . Une *orientation* de  $V$  est le choix en chaque  $x$  d'une orientation de l'espace tangent  $T_x$  en  $x$ , de façon que l'orientation soit localement constante sur les cartes.

Une carte en  $a$  est un difféomorphisme  $\Phi$  d'un voisinage  $U$  de  $a$  sur un voisinage  $V$  de 0 dans  $\mathbf{R}^p$ . L'espace tangent en un point  $x$  voisin de  $a$  est envoyé par  $d\Phi$  sur l'espace tangent à  $\mathbf{R}^p$ , qui est encore  $\mathbf{R}^p$ ; on demande que l'orientation image dans  $\mathbf{R}^p$  soit la même qu'en  $a$  aux points  $x$  voisins.

Une façon d'orienter une sous-variété est de se donner en chaque point une carte locale de façon que, là où l'on dispose de deux cartes  $\phi$  et  $\psi$  sur une même partie ouverte, l'application  $\phi \circ \psi^{-1}$  préserve l'orientation, autrement dit ait un déterminant jacobien positif.

### 3. Intégration d'une $p$ -forme sur une sous-variété de dimension $p$ .

Soient une partie ouverte  $U$  de l'espace vectoriel  $E$  de dimension  $n$ , une sous-variété orientée  $V$  de dimension  $p$  de  $U$  et une  $p$ -forme  $\omega$  dans  $U$ .

**1er cas.** On suppose que  $E = \mathbf{R}^n$ , que  $p = n$  et que  $V = U$  avec l'orientation canonique de  $\mathbf{R}^n$ .

Alors

$$\omega = f dx_1 \wedge \cdots \wedge dx_n .$$

On pose simplement

$$\int_U \omega = \int_U f dx_1 \cdots dx_n$$

sous réserve que l'intégrale ait un sens.

**Proposition.** Si  $\phi : U' \rightarrow U$  est un  $\mathcal{C}^1$ -difféomorphisme préservant l'orientation, alors

$$\int_U \omega = \int_{U'} \phi^* \omega .$$

En effet si  $\phi = (\phi_1, \dots, \phi_n)$  alors

$$\phi^* \omega = f \circ \phi d\phi_1 \wedge \cdots \wedge d\phi_n = f \circ \phi \det d\phi dx_1 \wedge \cdots \wedge dx_n .$$

L'égalité résulte du théorème de changement de variables.

**2ème cas.** On suppose que  $p \leq n$  et que  $V$  est l'image dans  $U$  d'une partie ouverte  $W$  de  $\mathbf{R}^p$  par une immersion  $\phi$  injective préservant les orientations de  $\mathbf{R}^p$  et de  $V$ . On pose alors

$$\int_V \omega = \int_W \phi^* \omega$$

lorsque le membre de droite a un sens. On est en effet ramené au cas précédent, avec  $p$  au lieu de  $n$ .

**3ème cas.** On suppose toujours  $p \leq n$ , mais que  $V$  puisse être recouverte par un nombre fini de parties ouvertes  $V_i$  sur lesquelles on dispose d'une paramétrisation locale  $\phi_i$  préservant l'orientation de  $V$ . Lorsque cela a un sens, on pose alors

$$\int_V \omega = \sum_i \int_{V_i} f_i \omega$$

où chaque fonction  $f_i$  est continue, positive, nulle en dehors de  $V_i$  et où  $\sum f_i = 1$ . Sur chaque  $V_i$  on est en effet ramené au cas précédent.

On se ramène au troisième cas lorsque  $V$  est fermée dans  $U$  et que  $\omega$  est à support compact dans  $U$ , en prenant les paramétrisations locales  $\phi_i$  fournies par des cartes locales.

**Proposition.** L'intégrale

$$\int_V \omega$$

est indépendante du choix de paramétrisations locales préservant l'orientation.

**Démonstration.** On se donne aussi un recouvrement par des parties ouvertes  $V'_j$  et des fonctions  $g_j$  associées. On note  $\phi_i : W_i \rightarrow V_i$  et  $\psi_j : W'_j \rightarrow V'_j$  des paramétrisations locales. Il s'agit de comparer

$$\sum_i \int_{V_i} f_i \omega \quad \text{et} \quad \sum_j \int_{V'_j} g_j \omega .$$

Posant  $V_{ij} = V_i \cap V'_j$ , on se ramène en fait à comparer deux expressions de

$$\sum_{i,j} \int_{V_{ij}} f_i g_j \omega .$$

Si  $W = \phi_i^{-1}(V_{ij})$  et  $W' = \psi_i^{-1}(V_{ij})$ , on est amené à comparer

$$\int_W f_i g_j \phi_i^* \omega \quad \text{et} \quad \int_{W'} f_i g_j \psi_j^* \omega .$$

Soit  $\theta : W \rightarrow W'$  le difféomorphisme induit par  $\psi_j^{-1} \circ \phi_i$ . Le second membre vaut encore

$$\int_W f_i g_j \theta^* \psi_j^* \omega = \int_W f_i g_j \phi_i^* \omega .$$

**Exercice.** Montrer que si  $U$  est étoilé, par exemple en 0, toute 1-forme  $\omega$  vérifiant  $d\omega = 0$  (i.e. fermée) est une différentielle  $df$  (i.e. exacte). On prendra pour  $f(x)$  l'intégrale de  $\omega$  sur le segment  $[0, x]$  orienté de 0 à  $x$ .

#### 4. Sous-variétés à bord.

On définit les sous-variétés à bord de la même façon que les variétés, mais en prenant comme modèle soit  $\mathbf{R}^p \times 0$  (pour un point intérieur), soit  $\mathbf{R}_+ \times \mathbf{R}^{p-1} \times 0$  (pour un point du bord). L'intérieur et le bord sont de vraies sous-variétés de dimensions  $p$  et  $p - 1$ .

Si la variété est orientée, son bord l'est également comme suit : une base  $e_1, \dots, e_{p-1}$  du plan tangent au bord est positive si la base  $e, e_1, \dots, e_{p-1}$  du plan tangent à la variété l'est quand on prend pour  $e$  un vecteur sortant.

#### 5. La formule de Stokes.

Une première version de ce théorème est facile à énoncer.

**Théorème.** Soient une partie ouverte  $U$  de  $E$ , une sous-variété orientée fermée  $V$  de dimension  $p$  de  $U$  et une forme  $\omega$  de degré  $p - 1$  et de classe  $\mathcal{C}^1$  à support compact dans  $U$ . On a

$$\int_V d\omega = 0 .$$

Une version un peu plus délicate est la suivante.

**Théorème.** Soient une sous-variété à bord orientée compacte  $V$  de dimension  $p$ , de bord orienté  $\partial V$ , et une forme  $\omega$  de degré  $p - 1$  et de classe  $\mathcal{C}^1$  au voisinage de  $V$  dans  $\mathbf{R}^n$ . On a

$$\int_V d\omega = \int_{\partial V} \omega .$$

**Démonstrations.** Etablissons d'abord le premier énoncé. L'hypothèse de compacité permet d'écrire

$$\int_V d\omega = \int_V d\left(\sum_i f_i \omega\right) = \sum_i \int_V d(f_i \omega)$$

où les fonctions  $f_i$  sont continues, positives, de somme 1 et où il existe une carte locale au voisinage du support de chaque  $f_i$ .

Or chaque terme de la somme est nul. On est en ramené, en prenant l'image inverse par une paramétrisation locale, au cas où  $V = \mathbf{R}^p$  et où  $\omega$  est une forme de degré  $p - 1$  à support compact dans  $\mathbf{R}^p$ .

Dans ce cas, il s'agit d'évaluer des intégrales du type

$$\int_{-A}^A \dots \int_{-A}^A (-1)^{i+1} \frac{\partial f}{\partial x_i} dx_1 \dots dx_p$$

pour  $A$  assez grand. Il s'agit en effet des intégrales des différentielles extérieures de formes du type

$$f dx_1 \wedge \dots \wedge \hat{dx}_i \wedge \dots \wedge dx_p$$

où le chapeau indique que  $dx_i$  a été retiré. On sait que  $\omega$  en est une somme.

Or, dans l'intégrale multiple ci-dessus, l'intégration en  $x_i$  donne zéro.

Le second énoncé s'établit de façon analogue. On se ramènera cette fois au cas où  $V = ]-\infty, 0] \times \mathbf{R}^{p-1}$ . Il s'agira alors d'évaluer une intégrale du type

$$\int_{-A}^0 \dots \int_{-A}^A (-1)^{i+1} \frac{\partial f}{\partial x_i} dx_1 \dots dx_p .$$

Pour  $i \neq 1$  on obtient encore 0, alors que pour  $i = 1$  il vient

$$- \int_{-A}^A \dots \int_{-A}^A f(0, x_2, \dots, x_p) dx_2 \dots dx_p .$$

Les intégrales de ce type constituent le second membre de la formule.

On peut aussi utiliser le premier énoncé.

On commence par prendre une partition de l'unité pour se ramener au cas où le support de  $\omega$  est inclus dans une partie ouverte sur laquelle on dispose d'une carte locale convenable.

Par image inverse de la paramétrisation locale, on se ramène finalement au cas où la variété  $V$  est définie au voisinage de 0 dans  $\mathbf{R}^p$  par  $x_1 \leq 0$ . Etant donné  $\epsilon > 0$ , on introduit une fonction régulière décroissante  $\rho_\epsilon(x_1)$  où  $\rho_\epsilon(x_1) = 1$  pour  $x_1 \geq -\epsilon$  et  $\rho_\epsilon(x_1) = 0$  pour  $x_1 \geq 0$ . Alors

$$0 = \int_V d(\rho_\epsilon \omega) = \int_V \rho_\epsilon d\omega + \int_V d\rho_\epsilon \wedge \omega \rightarrow \int_V d\omega - \int_{\partial V} \omega$$

quand  $\epsilon \rightarrow 0$ .

En effet

$$\int_V d\rho_\epsilon \wedge \omega = \int \dots \int_{-\epsilon < x_1 < 0} \rho'_\epsilon(x_1) f_1 dx_1 . dx_2 \dots dx_p$$

si  $\omega = f_1 dx_2 \wedge \dots \wedge dx_p + \dots$ . A la limite quand  $\epsilon \rightarrow 0$ , il vient

$$- \int \dots \int f_1 . dx_2 \dots dx_p$$

sachant que

$$\int_{-\epsilon < x_1 < 0} \rho'_\epsilon(x_1) dx_1 = -1 .$$

QED.

**Remarque.** Le théorème vaut encore si  $\omega$  est une forme continue jusqu'au bord et de classe  $\mathcal{C}^1$  à l'intérieur. Cependant, pour cela, il faut définir la différentielle extérieure d'une forme différentielle sur une sous-variété, ce qu'on peut faire en prenant des cartes. Alors, pour donner aux formes différentielles un caractère intrinsèque, leur valeur en un point  $x$  doit être prise  $\wedge^p T_x^*$  où  $T_x$  est le plan tangent en  $x$ .

## Chapitre 10. Longueurs, surfaces, etc.

Ce chapitre est essentiellement culturel. On fait un peu le tour de questions souvent négligées dans l'enseignement universitaire. Sauf indication contraire, on se place dans un espace euclidien  $E$  de dimension au plus 3.

### 1. En dimension maximum.

Il s'agit de la mesure des longueurs sur la droite, de la mesure des surfaces dans le plan, de la mesure des volumes dans l'espace à trois dimensions etc.

Avec ce problème aucune chausse-trape n'est à craindre. En dimension maximum la mesure possède en effet des propriétés conformes à l'intuition, comme l'additivité simple; en particulier la mesure du tout est plus grande que celle de la partie etc.

Partant de là, prenant la dimension 2 comme exemple, on peut chercher à mesurer une surface plane en traçant un premier quadrillage, avec une maille carrée de côté 1 dont la surface sera prise comme unité; on encadrera la partie à mesurer entre la réunion des carrés qu'elle contient et celle des carrés qui la rencontre. Ensuite on diminuera le côté de la maille pour lui donner la valeur  $1/2$ , puis  $1/4$  et ainsi de suite, améliorant l'encadrement à chaque étape. Si les deux suites ainsi mises en évidence, l'une croissante et l'autre décroissante, ont une limite commune, ce sera la surface cherchée. On dira que la partie concernée est *quarrable* au sens de Jordan.

Cela permet de mesurer carrés, triangles, disques et autres domaines bornés assez réguliers. Avec une propriété géométrique remarquable : la mesure est invariante par les isométries, en particulier les déplacements; cela peut se voir dès que l'on a su mesurer — théoriquement — la surface du disque.

**Exercice.** Montrer, plus généralement, qu'une partie convexe bornée est quarrable.

Malheureusement on ne peut pas atteindre, par ce type d'encadrement, la mesure d'une partie quelconque. Toutes les parties bornées du plan ne sont pas quarrables au sens de Jordan.

Il y aurait une façon de s'en tirer. On peut en effet établir, en dimension 2, l'existence d'une mesure additive et invariante pour toutes les parties; c'est un théorème de Banach. Mais le résultat est faux dès la dimension 3.

Surtout cette approche est fondamentalement incorrecte. Ce que l'intuition suggère est non seulement une propriété d'existence, mais également une propriété d'unicité. Pour ce faire, on imposera

- l'invariance par isométrie, donc par l'action d'un groupe, qui fera de la mesure cherchée ce qu'on appelle une mesure de Haar;
- ainsi que la petite normalisation consistant à fixer la mesure d'un corps convexe donné; c'est le choix d'une unité.

Pour obtenir cette unicité, il convient d'ajouter à l'additivité simple le passage à la limite monotone, ce qui revient à supposer l'additivité dénombrable. C'est la raison pour laquelle on réfère à la théorie de Lebesgue, qui demande des constructions un peu plus savantes.

Savantes, mais intuitives quand même. Prenant toujours la dimension 2 en exemple, voici comment mesurer une partie  $U$  ouverte du plan. On prend, comme précédemment, un quadrillage de côté 1 et on retient la somme  $A_0$  des surfaces des carrés inclus dans  $U$ .

Recommençant avec un quadrillage de côté  $1/2$  on y ajoute la somme  $A_n$  des surfaces des carrés inclus dans  $U$  qui ne sont pas déjà dans les carrés précédents. Et ainsi de suite, avec une surface  $A_n$  avec un quadrillage de côté  $2^{-n}$ . La surface cherchée sera

$$\sum_n A_n$$

par additivité dénombrable, du moins lorsqu'on aura montré qu'on épuise ainsi  $U$ . Or si  $a$  est dans  $U$ , on peut trouver  $\epsilon > 0$  tel que le carré parallèle au quadrillage de centre  $a$  et de côté  $2\epsilon$  soit inclus dans  $U$ ; alors, pour  $n$  tel que  $2^{-n} \leq \epsilon$ , le point  $a$  est dans un des carrés du quadrillage de côté  $2^{-n}$  et ce dernier est inclus dans  $U$ , donc retenu.

Par passage au complémentaire, on peut mesurer les parties compactes. Ensuite on encadrerait une partie entre des parties compactes plus petites et des parties ouvertes plus grandes. C'est là qu'il y a des démonstrations à faire.

Si l'on entre dans les détails, le point difficile est la construction de la mesure de Lebesgue sur  $[0, 1]$ . Le passage à  $[-N, N]$  est alors immédiat, celui à  $\mathbf{R}$  se fait par une limite croissante :

$$\text{mes } A = \lim_{N \rightarrow \infty} \text{mes}(A \cap [-N, N]) .$$

Ensuite le passage à  $\mathbf{R}^n$  relève des mesures produit, ou du théorème de Fubini.

Maintenant, plutôt que de construire ce qui est encore la loi uniforme sur  $[0, 1]$ , il peut être plus pratique de raisonner en termes probabilistes. Comment choisir un nombre entre 0 et 1 en respectant l'uniformité, i.e. de façon que la probabilité de tomber dans l'intervalle  $(a, b)$  soit  $b - a$ ? On peut toujours écrire  $x$  sous la forme d'un développement dyadique

$$0, \alpha_1 \alpha_2 \dots \alpha_n \dots$$

où les  $\alpha_n$  valent 0 ou 1. Le premier chiffre  $\alpha_1$  vaut 0 entre 0 et  $1/2$  et 1 entre  $1/2$  et 1. Il a autant de chances de valoir 0 ou 1. C'est, bien sûr, la même chose pour les autres.

Ainsi le tirage d'un nombre réel entre 0 et 1 revient-il très précisément au jeu non faussé de pile ou face. Mais il s'agit du jeu infini, pour lequel on ne s'arrête jamais. Construire la mesure de Lebesgue revient à construire un modèle de ce jeu. L'ensemble des épreuves est ici  $\{0, 1\}^{\mathbf{N}}$ ; il reste à construire une tribu et une mesure. On donne rarement ce modèle, qu'on trouvera néanmoins en exercice dans le livre de Foata et Fuchs.

De façon plus savante, on obtiendra la loi uniforme cherchée en prenant l'image de la mesure de probabilité du jeu précédent par l'application qui à la suite  $\alpha_1, \dots, \alpha_n, \dots$  associe le nombre réel

$$\sum_{i>0} 2^{-i} \alpha_i .$$

Ce n'est pas exactement une bijection; à cause de l'ambiguïté du développement dyadique, certaines valeurs sont obtenues deux fois. Mais ces anomalies constituent une partie dénombrable, donc négligeable.

Maintenant, avec la théorie de Lebesgue on peut attribuer à pratiquement toute partie une mesure, éventuellement infinie. On peut même, grâce aux travaux du logicien Solovay, supposer que c'est possible pour toute dès lors qu'on ne fait pas appel à l'axiome du choix général et à ce qui en découle : lemme de Zorn, récurrence transfinie. On ne risque pas de contradiction autre que celles pouvant déjà faire de la théorie des ensembles.

## 2. Longueurs; lignes brisées.

Nous sommes bien sûr ici en dimension  $> 1$ . Le problème de la longueur d'une courbe remonte à Archimède, avec son calcul de la longueur du cercle, calcul pour lequel il utilisait, entre autres, des polygones réguliers inscrits.

Si  $\gamma : [a, b] \rightarrow E$  est une courbe paramétrée, on peut considérer les lignes brisées inscrites sur la courbe, composées de segments  $[\gamma(\alpha_i), \gamma(\alpha_{i+1})]$  où

$$a = \alpha_0 < \alpha_1 < \dots < \alpha_n = b$$

est une subdivision de  $[a, b]$ . Chaque ligne brisée inscrite a une longueur et on convient de définir la longueur  $L(\gamma)$  de la courbe comme la borne supérieure, éventuellement infinie, des longueurs des lignes brisées inscrites.

On notera qu'il faut imposer un sens de parcours; si on prend les  $\alpha_i$  dans  $[a, b]$  n'importe comment, on trouvera presque toujours une borne supérieure infinie. Pour la longueur du cercle, on peut prendre des polygones réguliers inscrits, mais pas des polygones étoilés.

La longueur ainsi définie s'applique à une fonction  $\gamma$  quelconque. Cependant pour donner un sens géométrique à la longueur ainsi définie, mieux vaut supposer la fonction continue; sinon on mesurera aussi les sauts. Et pour s'abstraire du paramétrage, mieux vaut supposer  $\gamma$  injective autant qu'il est possible, sachant que ce ne l'est pas tout à fait pour le cercle; sinon on pourrait, par exemple, parcourir deux fois la courbe et trouver le double pour la longueur.

Cela étant, dans le cas le plus général de nouveau, une courbe est dite *rectifiable* si sa longueur est finie.

**Exercice.** Donner quelques exemples de courbes continues non rectifiables.

**Indication.** Essayer  $x \sin(1/x^2)$ , voire  $x \sin(1/x)$  sur  $[0, a]$ , avec la valeur 0 en 0; on calculera la variation totale en décomposant  $]0, a]$  en intervalles sur lesquels la fonction est monotone.

Un cas particulier important est celui du graphe  $\Gamma_f$  dans l'espace  $\mathbf{R}^2$  d'une fonction  $f : [a, b] \rightarrow \mathbf{R}$ . On vérifie facilement que ce graphe est rectifiable dès lors que la *variation totale*  $V(f)$  de  $f$  est finie, où  $V(f)$  est définie comme la borne supérieure des sommes

$$|f(\alpha_{i+1}) - f(\alpha_i)|$$

où la borne est prise sur l'ensemble des subdivisions de  $[a, b]$ . En effet, dans le cas euclidien par exemple, pour calculer la longueur, on prendra

$$\sqrt{(\alpha_{i+1} - \alpha_i)^2 + (f(\alpha_{i+1}) - f(\alpha_i))^2}$$

et comme ce terme est encadré entre

$$|f(\alpha_{i+1}) - f(\alpha_i)| \quad \text{et} \quad |\alpha_{i+1} - \alpha_i| + |f(\alpha_{i+1}) - f(\alpha_i)|$$

il vient

$$V(f) \leq L(\Gamma_f) \leq V(f) + b - a .$$

Les fonctions dont la variation totale est finie sont dites à *variation bornée*. Elles constituent un espace vectoriel. Les fonctions réelles monotones en sont; leur variation totale est, en valeur absolue, la différence des valeurs extrêmes; dans le cas général elle est toujours plus grande.

Inversement une fonction réelle à variation bornée est la différence de deux fonctions croissantes. On écrit

$$f = g - (g - f)$$

où  $g(x)$  est la variation totale sur  $[a, x]$ . Clairement  $g$  est croissante et  $g \geq f$ ; le fait que  $g - f$  soit croissante vient de ce que la variation totale vérifie la relation de Chasles

$$V_{[a,b]}(f) = V_{[a,c]}(f) + V_{[c,b]}(f)$$

pour  $a \leq c \leq b$ . Alors si  $x \leq y$ , on a

$$g(y) - f(y) = g(x) - f(x) + V_{[x,y]}(f) - (f(y) - f(x)) .$$

Revenons aux longueurs. Voici un moyen de les calculer.

**Théorème.** Si  $E$  est un espace normé complet, une courbe paramétrée  $\gamma : [a, b] \rightarrow E$  de classe  $\mathcal{C}^1$  est rectifiable et sa longueur donnée par

$$L(\gamma) = \int_a^b \|\gamma'(t)\| dt .$$

La démonstration en sera donnée à la section suivante.

### 3. Élément de longueur.

Maintenant on se place dans un plan ou un espace euclidien. Une idée possible, pour des courbes régulières, est de prendre en quelque sorte la formule intégrale ci-dessus comme définition. Cependant on voudrait le faire d'une façon qui soit plus intrinsèque, moins dépendante du paramétrage de la courbe.

Précisément on va considérer une courbe  $\gamma : [a, b] \rightarrow E$  injective de classe  $\mathcal{C}^1$  dont la dérivée est partout non nulle; c'est un arc inclus dans une sous-variété de dimension 1 définie par une immersion.

La dérivabilité étant calquée sur le modèle linéaire, on va définir la longueur "sur la tangente".

La petite question à résoudre est le calcul de l'élément de longueur sur une droite. On constate aisément que c'est

$$\|d\vec{M}\| = |\vec{\tau}.d\vec{M}|$$

si  $\vec{\tau}$  est un vecteur directeur unitaire de la droite. En effet la quantité dans la norme ou la valeur absolue est linéaire dans les vecteurs de la droite et sur une base orthonormée — composée du vecteur  $\vec{\tau}$  par exemple — elle vaut 1 en norme ou en module.

C'est simplement une façon compliquée d'exprimer  $\|\vec{h}\|$  ou  $\vec{h} = \lambda\vec{\tau}$ , sachant que  $d\vec{M}(\vec{h}) = \vec{h}$  par définition et que  $|\vec{\tau}.\lambda\vec{\tau}| = |\lambda| = \|\lambda\vec{\tau}\|$ . On peut sincèrement se demander s'il est vraiment utile d'introduire ce formalisme creux.

En fait la réponse est la même que pour le  $dx$  qui figure dans l'intégrale

$$\int_a^b f(x)dx$$

alors que

$$\int_{x=a}^{x=b} f(x)$$

pourrait suffire. Pourquoi ce  $dx$  qui fait écrire  $h$  sous la forme plus compliquée  $dx(h)$ ? Pour l'intégrale, la réponse est qu'on a un formalisme qui rend automatique le changement de variable. La réponse est la même pour l'élément de longueur. On parle pour  $\|d\vec{M}\| = |\vec{\tau} \cdot d\vec{M}|$  de *densité différentielle* pour signifier que les règles de l'image inverse s'appliquent.

Si  $\vec{M} = \gamma(t)$ , par image inverse l'élément de longueur donne

$$\gamma^* \|d\vec{M}\| = \|\gamma'(t)\| dt$$

une fois convenu que  $dt$  n'est plus ici une forme différentielle mais une mesure, donc à considérer formellement comme un nombre positif.

Maintenant on peut rattacher cela aux lignes brisées. Au point de paramètre  $\alpha$  un vecteur tangent unitaire est

$$\vec{\tau} = \frac{\gamma'(\alpha)}{\|\gamma'(\alpha)\|} .$$

Projetée sur la tangente, la longueur du segment  $[\gamma(t), \gamma(t')]$  est

$$|\vec{\tau} \cdot (\gamma(t') - \gamma(t))|$$

dont la partie principale quand  $t, t'$  tendent vers  $\alpha$  est

$$|\vec{\tau} \cdot ((t' - t)\gamma'(\alpha))| = |t' - t| \|\gamma'(\alpha)\| .$$

La longueur elle-même est donnée par

$$\sqrt{(\gamma(t') - \gamma(t)) \cdot (\gamma(t') - \gamma(t))}$$

dont la partie principale quand  $t, t'$  tendent vers  $\alpha$  est

$$\sqrt{(t' - t)^2 \gamma'(\alpha) \cdot \gamma'(\alpha)} = |t' - t| \|\gamma'(\alpha)\|$$

aussi. Voilà qui renforce le choix de l'élément de longueur choisie et la formule considérée à la section précédente comme définition .

**Exercice.** Donner une démonstration en forme du théorème de la section 2.

**Indications.** Pour établir l'égalité de deux expressions qui sont, comme ici, définies par des opérations assimilables à des "limites", qui plus est de nature très différente, on ne peut pas effectuer de calcul formel à partir de l'une ou de l'autre. On se donne donc une quantité  $\epsilon > 0$  arbitraire et on montre que la différence est majorée, en valeur absolue, par cet  $\epsilon$ .

Le membre de gauche est défini par

$$L(\gamma) = \sup \sum_{i=0}^{n-1} \|\gamma(\alpha_{i+1}) - \gamma(\alpha_i)\|$$

où la borne est prise sur l'ensemble des subdivisions  $\mathcal{S}$  de  $[a, b]$ , de la forme

$$a = \alpha_0 < \alpha_1 < \dots < \alpha_n = b$$

où le nombre  $n$  dépend lui même de la subdivision  $\mathcal{S}$ , nombre qu'on devrait plutôt noter  $n_{\mathcal{S}}$  par conséquent.

La première chose à faire est de se débarrasser de cette borne supérieure, en revenant à la définition de celle-ci : on peut trouver une subdivision  $\mathcal{S}$  telle que l'on ait

$$L(\gamma) - \epsilon \leq \sum_{i=0}^{n-1} \|\gamma(\alpha_{i+1}) - \gamma(\alpha_i)\| \leq \epsilon .$$

Regardons maintenant du côté du membre de gauche. L'intégrale d'une fonction continue sur un segment est une "limite" de somme de Riemann. Précisément, il existe un nombre  $\delta > 0$  tel que la somme de Riemann

$$\sum_{i=0}^{n-1} (\alpha_{i+1} - \alpha_i) \|\gamma'(\alpha_i)\|$$

diffère de l'intégrale considérée de moins de  $\epsilon$  pour toute subdivision  $\mathcal{S}$  dont le pas est inférieur à  $\delta$ , i.e. telle que  $\alpha_{i+1} - \alpha_i \leq \delta$  pour tout  $i$ .

On se ramène ainsi à comparer les expressions

$$\|\gamma(\alpha_{i+1}) - \gamma(\alpha_i)\| \quad \text{et} \quad (\alpha_{i+1} - \alpha_i) \|\gamma'(\alpha_i)\|$$

ou à évaluer leur différence, elle-même majorée par la norme de

$$\gamma(\alpha_{i+1}) - \gamma(\alpha_i) - (\alpha_{i+1} - \alpha_i) \gamma'(\alpha_i)$$

grâce à l'inégalité triangulaire.

Dans cette dernière on peut voir un  $\gamma(\alpha_i + h) - \gamma(\alpha_i) - h\gamma'(\alpha_i)$  qui fait penser à la définition d'une différentielle ou d'une dérivée. C'est un  $o(h)$ , c'est-à-dire un  $o(\alpha_{i+1} - \alpha_i)$ ; c'est négligeable devant  $\alpha_{i+1} - \alpha_i$  dès que ce dernier est assez petit. Cependant ce raisonnement exige que l'on ait fixé  $\alpha_i$  d'abord, pour prendre  $\alpha_{i+1}$  assez proche ensuite. C'est proprement inextricable.

La seule façon de s'en tirer est de passer par le théorème des accroissements finis. On pose

$$g(t) = \gamma(t) - \gamma(\alpha_i) - (t - \alpha_i) \gamma'(\alpha_i) .$$

La quantité à évaluer est  $g(\alpha_{i+1}) - g(\alpha_i)$ . Or

$$g'(t) = \gamma'(t) - \gamma'(\alpha_i) .$$

Si  $\|\gamma'(t) - \gamma'(\alpha_i)\| \leq \epsilon$  pour  $t$  dans l'intervalle  $[\alpha_i, \alpha_{i+1}]$ , on aura majoré notre expression en norme par

$$\epsilon(\alpha_{i+1} - \alpha_i) .$$

Il reste donc à assurer  $\|\gamma'(t) - \gamma'(\alpha_i)\| \leq \epsilon$  pour  $t$  dans l'intervalle  $[\alpha_i, \alpha_{i+1}]$ . Là encore, on peut penser à la continuité de  $\gamma'$ , mais il faudrait fixer pour cela  $\alpha_i$ , ce qui conduirait à la même impasse que précédemment. C'est la continuité uniforme qu'il faut invoquer pour  $\gamma'$ , grâce au théorème de Heine. On peut trouver  $\eta > 0$  tel que la condition cherchée soit réalisée dès que  $\alpha_{i+1} - \alpha_i \leq \eta$ .

En faisant la somme sur  $i$ , on voit que nos deux sommes diffèrent en norme de  $\epsilon(b-a)$  au plus dès que  $\alpha_{i+1} - \alpha_i \leq \eta$ .

Construisons maintenant la structure de notre démonstration. Nous nous sommes donnés  $\epsilon$ . Nous en avons tiré une subdivision en supprimant la borne supérieure. Et nous en avons tiré un  $\delta > 0$  pour approcher l'intégrale et un  $\eta > 0$  pour comparer les sommes.

D'abord rien n'empêche de même  $\eta$  et  $\delta$ , en les remplaçant par  $\beta = \min(\eta, \delta)$ . En imposant  $\alpha_{i+1} - \alpha_i \leq \beta$ , on disposera à la fois des deux estimations effectuées.

Le problème est que la subdivision  $\mathcal{S}$  obtenue ne vérifie peut-être pas cette condition. Qu'à cela ne tienne! On rajoutera des points. La longueur de la ligne brisée ne peut qu'en être augmentée, puisque la droite reste "le plus court chemin".

#### 4. Epaissement d'une courbe.

On se place ici dans un plan. Puisqu'il est plus délicat de mesurer des longueurs que des surfaces, pourquoi ne pas ramener le calcul des longueurs à celui des surfaces?

Pour illustrer la difficulté présentée par la définition de la longueur, reprenons le calcul de  $\pi$  par Archimède. Encadrant la longueur du cercle par celle d'un carré inscrit et d'un carré circonscrit, on obtient déjà

$$4\sqrt{2} \leq 2\pi \leq 8 .$$

Cependant, définissant la longueur du cercle comme la borne supérieure de celle des lignes brisées inscrites, s'il est facile de justifier la minoration de gauche, la majoration de droite est plus délicate à établir. Il faut par exemple envisager une ligne brisée arbitraire sur un huitième de cercle et comparer sa longueur avec celle de la portion de tangente en une extrémité qui lui correspond. C'est possible et même élémentaire. On le fera en **exercice**.

**Exercice.** Donner une définition géométrique du nombre  $\pi$ .

**Indication.** Il s'agit de comparer la longueur du cercle et la surface du disque. On partira de la surface pour définir  $\pi$  et on tirera la longueur de la surface d'une couronne.

**Remarque.** Ce n'est pas forcément le meilleur choix pour l'école élémentaire; mieux vaut y partir de la longueur du cercle.

Pour rester dans la définition du nombre  $\pi$ , noter qu'une définition rigoureuse de  $\pi$  à l'université peut être la suivante, dont l'introduction dans l'enseignement est due à Henri Cartan. On définit  $\pi/2$  comme le plus petit  $x > 0$  vérifiant  $\cos x = 0$ , où la fonction cosinus est tirée de l'exponentielle complexe, laquelle est définie par sa série entière.

Jean Dieudonné en donne une autre dans son livre de géométrie élémentaire, qui s'appuie sur une forme différentielle invariante sur le cercle.

De façon générale, une idée possible pour mesurer la longueur d'une courbe est d'épaissir le trait de la courbe. Par exemple, dans le plan, on créera une sorte de boudin  $B_\epsilon$  en réunissant des segments de longueur  $2\epsilon > 0$  obtenus en reportant la longueur  $\epsilon$  de part et d'autre de la normale à la courbe en chaque point. La longueur sera la limite, si elle existe, de la surface de  $B_\epsilon$  divisée par  $2\epsilon$ .

Précisément, si  $\gamma$  est de classe  $\mathcal{C}^1$  et de dérivée  $\gamma'(t)$  partout non nulle, on prendra pour  $B_\epsilon$  l'image du rectangle  $[a, b] \times [-\epsilon, \epsilon]$  par l'application  $\Phi$  qui au point  $(t, u)$  vaut

$$\Phi(t, u) = \vec{\gamma}(t) + u\vec{\nu}(t)$$

où  $\vec{\nu}$  est le vecteur normal unitaire en  $\gamma(t)$  obtenu à partir du vecteur tangent unitaire  $\vec{\tau}$  déjà considéré par une rotation d'angle droit positif.

**Exercice.** On suppose en outre la fonction  $\gamma$  de classe  $\mathcal{C}^2$  et injective; on étend  $\gamma$  de façon  $\mathcal{C}^2$  à un intervalle  $]\alpha, \beta[$  contenant  $[a, b]$  en prenant à gauche de  $a$  (resp. à droite de  $b$ ) son développement de Taylor à l'ordre 2 en  $a$  (resp.  $b$ ); on étend  $\Phi$  à  $]\alpha, \beta[ \times \mathbf{R}$ .

1) Montrer qu'on peut choisir  $\rho > 0$  assez petit pour que  $\Phi$  soit  $\mathcal{C}^1$  difféomorphisme local en tout point de  $]a - \rho, b + \rho[ \times ]-\rho, \rho[$ .

2) Ayant choisi  $\rho$  et posé  $R = [a, b] \times [-\rho/2, \rho/2]$ , dans le produit  $R \times R$  on considère l'ensemble  $A$  des points  $(v, w)$  tels que  $v \neq w$  et  $\Phi(v) = \Phi(w)$ ; montrer que c'est une partie fermée.

3) Montrer que l'application  $\Phi$  est injective sur  $R_\epsilon = [a, b] \times [-\epsilon, \epsilon]$  pour  $\epsilon > 0$  assez petit.

4) Montrer que

$$\text{mes}(B_\epsilon) = \int \int_{[a, b] \times [-\epsilon, \epsilon]} |\det(d\Phi)| dt du$$

pour  $\epsilon$  assez petit également, puis que

$$\lim_{\epsilon \rightarrow 0, \epsilon > 0} \frac{1}{2\epsilon} \text{mes}(B_\epsilon) = L(\gamma) .$$

### Indications.

1) On pourra calculer le déterminant de  $d\Phi$  en un point en se plaçant dans la base  $\vec{\tau}, \vec{\nu}$ ; on prendra ensuite  $\delta > 0$  tel que  $[a - \delta, b + \delta] \subset ]\alpha, \beta[$ .

3) On considère la suite  $K_n = (R_{1/n} \times R_{1/n}) \cap A$  dont on montrera que le terme général est vide pour  $n$  assez grand.

### 5. Surfaces, triangulations.

Les puristes parlent plutôt d'aire, puisque le mot surface est aussi utilisé pour désigner une sous-variété de dimension 2. Mais le commun des mortels en reste aux surfaces. Il n'y jamais d'ambiguïté en pratique.

On peut chercher à remplacer la ligne brisée inscrite dans une courbe par ce qui est appelé une *triangulation* de la surface.

Malheureusement, si l'on ne prend pas de précautions, même pour une surface très régulière comme une tranche de la surface latérale d'un cylindre de révolution, on peut trouver une borne supérieure infinie. Imaginons qu'on ait tracé de petits losanges inscrits  $ABCD$  où la diagonale  $AC$  est parallèle à la base et la diagonale  $BD$  parallèle aux

génératrices, la seconde étant beaucoup plus petite que la première, et cela d'autant que le triangle sera petit. La somme des surfaces des triangles  $BAC$  et  $DAC$  sera beaucoup plus grande que la somme des surfaces des triangles  $ABD$  et  $CBD$ . Le premier choix conduirait à la vraie surface; le second à l'infini.

Evidemment on pourrait mettre des restrictions sur le choix des triangles, mais ce serait compliqué. Mieux vaut chercher ailleurs.

**Remarques.** Triangler par des triangles équilatéraux est exclu : pour la sphère cela aboutirait à un polyèdre régulier; or il n'y en a que cinq, dont seulement deux avec des triangles équilatéraux, le tétraèdre et l'isocaèdre.

De même placer des carrés sur une surface est très risqué; qui n'a pas vu le comportement d'une chaise sur un sol non plan?

## 6. Élément de surface.

Reprenons une idée introduite à propos des courbes. On va se limiter à des sous-variétés de dimension 2 de l'espace euclidien  $\mathbf{R}^3$ .

La définition d'une sous-variété repose sur la notion de fonction différentiable, donc sur le modèle linéaire. On mesurera alors la surface "dans le plan tangent".

La petite question à résoudre est le calcul de l'élément de surface dans un plan qui n'est pas celui de coordonnées comme le plan tangent  $T$  en un point. On montre aisément que c'est la densité différentielle

$$\|d\vec{S}\| = |\vec{n}.d\vec{S}|$$

si  $\vec{n}$  est un vecteur normal unitaire.

Ici encore la norme ou la valeur absolue indiquent que l'on a remplacé l'élément différentiel par une simple mesure; on peut toujours parler de *densité différentielle* pour signifier que les règles de l'image inverse s'appliquent.

Vérifions d'abord l'égalité. La formule

$$(\vec{v}.d\vec{S})(\vec{h}, \vec{k}) = \det(\vec{v}, \vec{h}, \vec{k})$$

où le second membre est nul quand le  $\vec{v}$  est dans le plan  $T$  de  $\vec{h}$  et  $\vec{k}$ , ce qui montre que le vecteur  $d\vec{S}$  est normal au plan; d'où le calcul de sa norme.

Ensuite l'expression dont prend la norme ou la valeur absolue est bilinéaire et alternée dans les vecteurs du plan et sa valeur sur une base orthonormée est 1 en norme ou en valeur absolue. C'est bien l'élément de surface et l'intégrale

$$\int |\vec{n}.d\vec{S}|$$

donne la surface.

**Exercice** (Marc Rogalski). "On considère une conduite d'eau dont la forme est un tube de révolution engendré par rotation autour de l'axe  $Ox$  de la courbe d'équation  $y = 1/x$ , entre les abscisses  $a$  et  $2a$  ( $a > 0$ ). La conduite est poreuse; sur une petite longueur de la canalisation, la fuite est à peu près proportionnelle à la surface du morceau de tube et au débit à travers celui-ci (on notera  $k$  la constante de proportionnalité). Si  $d_0$  est le débit d'entrée, quel est le débit  $d_1$  de sortie?"

Cet exercice se fait en écrivant l'équation différentielle

$$\delta' = -k \frac{\Delta S}{\Delta x} \delta$$

régissant le débit  $\delta$ , où il s'agit d'évaluer la surface  $\Delta S$  de conduite correspondant à la tranche de longueur  $\Delta x$  en  $x$ .

### **7. Epaissement d'une surface.**

On peut aussi reprendre une autre idée introduite à propos des courbes. Il s'agit d'épaissir la surface pour en faire un volume et de passer à la limite. Il n'y a pas de différence sensible avec le cas des courbes.

## Chapitre 11. Propriétés métriques

Dans ce chapitre on se place dans le plan ou dans l'espace euclidiens, le plus souvent orientés. On s'intéresse à des propriétés métriques des courbes et surfaces.

### 1. Courbes planes.

Soit une courbe plane, définie par une application d'un intervalle  $I$  dans le plan  $\mathbf{R}^2$ , de classe  $\mathcal{C}^1$ , injective, de dérivée partout non nulle — donc une immersion; elle est orientée par le paramétrage. On note  $\vec{M}(t)$  le point de paramètre  $t$ .

Etant donnée une origine  $\vec{M}_0 = \gamma(t_0)$  de cette courbe, la *longueur d'arc*, ou *abscisse curviligne*, est la fonction

$$s = \int_{t_0}^t \left\| \frac{d\vec{M}}{dt} \right\| dt .$$

C'est un  $\mathcal{C}^1$ -difféomorphisme de  $I$  sur un intervalle  $J$ . Si, par changement de variable, on remplace le paramètre  $t$  par  $s$ , on dit qu'on a paramétré la courbe par la longueur d'arc. L'orientation est inchangée.

Dans ce cas, si on définit le *vecteur tangent unitaire positif*  $\vec{\tau}$  par

$$\vec{\tau} = \frac{d\vec{M}}{ds} ,$$

on a

$$\|\vec{\tau}\| = 1 .$$

En effet

$$\frac{d\vec{M}}{ds} = \frac{d\vec{M}}{dt} \cdot \frac{dt}{ds} = \left\| \frac{d\vec{M}}{dt} \right\|^{-1} \cdot \frac{d\vec{M}}{dt} .$$

Toujours dans ce cas, la courbe étant supposée de classe  $\mathcal{C}^2$ , le vecteur

$$\frac{d\vec{\tau}}{ds}$$

est orthogonal à  $\vec{\tau}$ . En effet en dérivant  $\vec{\tau} \cdot \vec{\tau} = 1$ , il vient

$$0 = d(\vec{\tau} \cdot \vec{\tau}) = 2\vec{\tau} \cdot \frac{d\vec{\tau}}{ds} .$$

La norme

$$K = \left\| \frac{d\vec{\tau}}{ds} \right\|$$

de ce vecteur normal est, par définition, la *courbure* au point  $s$ ; elle exprime la variation angulaire du vecteur tangent unitaire, donc de la tangente. Le *rayon de courbure* est  $R = 1/K$ .

**Exercice.** Quelle est la courbure d'une droite, d'un cercle?

Si la courbure n'est pas nulle, on introduit le *vecteur normal unitaire*

$$\vec{\nu} = \frac{1}{K} \cdot \frac{d\vec{\tau}}{ds}$$

et le *centre de courbure* est le point  $C$  tel que  $\vec{MC} = R \cdot \vec{\nu}$ . Le cercle de centre  $C$  et de rayon  $R$  est appelé *cercle osculateur*.

**Remarque.** Lorsque le plan est orienté, si  $r$  est la rotation d'angle droit positif, dans la base  $\vec{\tau}, r(\vec{\tau})$ , on peut définir une courbure dotée d'une signe en posant

$$\frac{d\vec{\tau}}{ds} = K.r(\vec{\tau}) .$$

Cette courbure précise dans quel sens la courbe va tourner.

## 2. Expression de la courbure.

On peut utiliser le déterminant ou plonger  $\mathbf{R}^2$  dans  $\mathbf{R}^3$  pour utiliser le produit vectoriel. Le calcul de la norme de  $\vec{\nu}$  est aussi celui de la valeur absolue de  $\det(\vec{\tau}, \vec{\nu})$  ou de la norme de  $\vec{\tau} \wedge \vec{\nu}$  ou. Or

$$\vec{\nu} = \frac{d}{ds} \left( \left\| \frac{d\vec{M}}{dt} \right\|^{-1} \cdot \frac{d\vec{M}}{dt} \right) = \left\| \frac{d\vec{M}}{dt} \right\|^{-1} \frac{d}{dt} \left( \left\| \frac{d\vec{M}}{dt} \right\|^{-1} \cdot \frac{d\vec{M}}{dt} \right) .$$

Dans la parenthèse de droite, le terme obtenu en dérivant le facteur scalaire va se détruire dans le déterminant ou le produit vectoriel avec  $\vec{\tau}$ . Il reste donc

$$\vec{\tau} \wedge \vec{\nu} = \left\| \frac{d\vec{M}}{dt} \right\|^{-3} \cdot \frac{d\vec{M}}{dt} \wedge \frac{d^2\vec{M}}{dt^2} \quad \text{ou} \quad \det(\vec{\tau}, \vec{\nu}) = \left\| \frac{d\vec{M}}{dt} \right\|^{-3} \cdot \det\left(\frac{d\vec{M}}{dt}, \frac{d^2\vec{M}}{dt^2}\right) .$$

Si on désigne par  $x$  et  $y$  les composantes de  $\gamma$ , il vient

$$K(t) = \frac{|x'y'' - x''y'|}{(x'^2 + y'^2)^{3/2}} .$$

Dans le cas où le plan est orienté, la courbure est définie par la même formule, mais sans valeur absolue au numérateur.

**Exercice.** Caractériser les courbes de courbure constante  $K$  donnée, en distinguant  $K = 0$  et  $K > 0$ .

On pourra supposer une telle courbe définie localement par  $y = y(x)$  ou  $x = x(y)$ . Dans le premier cas on résoudra  $y'' = c(1 + y'^2)^{3/2}$  en posant  $y' = \tan t$ .

**Exercice.** Que peut-on dire de la différence  $\vec{\delta}$  entre la courbe  $\gamma$  et son cercle osculateur en un point, les deux étant paramétrés par la longueur d'arc? On pourra considérer la courbe  $\vec{r} = s\vec{\tau}_0 + \vec{\delta}$ , puis  $d\vec{r}/ds$  et  $\vec{\tau}_0 \wedge d^2\vec{r}/dt^2$  en  $t_0$ . Que peut-on encore dire de la distance entre la courbe et son cercle osculateur au voisinage de ce point?

## 3. Courbes gauches.

L'espace euclidien est supposé orienté. La courbe est supposée de classe  $\mathcal{C}^3$ .

La définition du vecteur  $\vec{\tau}$ , de la courbure  $K$  et, si elle n'est pas nulle, du vecteur normal  $\vec{\nu}$  vaut encore pour une courbe de  $\mathbf{R}^3$ . Lorsque le vecteur *normal principal*  $\vec{\nu}$  est défini, le plan qu'il fait avec  $\vec{\tau}$  est appelé *plan osculateur*; on introduit encore alors le *vecteur binormal*

$$\vec{\beta} = \vec{\tau} \wedge \vec{\nu} .$$

Le plan de  $\vec{\tau}$  et de  $\vec{\beta}$  est appelé *plan rectifiant*, le plan de  $\vec{\nu}$  et de  $\vec{\beta}$  est appelé *plan normal* et le repère direct  $\vec{\tau}, \vec{\nu}, \vec{\beta}$  est appelé *repère de Freinet*.

**Remarque.** Dans l'espace la courbure est toujours positive.

**Exercice.** Que peut-on dire de  $\vec{\beta}_0 \cdot \vec{M}$  pour  $\vec{M}$  voisin de  $\vec{M}_0$ , où  $\vec{\beta}_0$  est le vecteur binormal en  $\vec{M}_0$ ? On regardera la valeur et les dérivées en  $t_0$ . Que peut-on dire de la distance entre la courbe et son plan osculateur au voisinage d'un point donné?

On remarque que  $d\vec{v}/ds$  est orthogonal à  $\vec{v}$  puisque que ce dernier est unitaire. De plus

$$\vec{\tau} \cdot \frac{d\vec{v}}{ds} = -K\vec{\tau}$$

se voit en dérivant  $\vec{\tau} \cdot \vec{v} = 0$ .

Ainsi

$$\frac{d\vec{v}}{ds} = -K\vec{\tau} + T\vec{\beta}$$

où  $T$  définit la *torsion*. Alors

$$\frac{d\vec{\beta}}{ds} = \vec{\tau} \wedge \frac{d\vec{v}}{ds} = T\vec{\tau} \wedge \vec{\beta} = -T\vec{v}.$$

En particulier la torsion exprime la variation angulaire du vecteur binormal, donc aussi du plan osculateur; elle a un signe.

On résume tout cela dans les *formules de Frenet*

$$\begin{pmatrix} d\vec{\tau}/ds \\ d\vec{v}/ds \\ d\vec{\beta}/ds \end{pmatrix} = \begin{pmatrix} 0 & K & 0 \\ -K & 0 & T \\ 0 & -T & 0 \end{pmatrix} \begin{pmatrix} \vec{\tau} \\ \vec{v} \\ \vec{\beta} \end{pmatrix}.$$

**Exercice.** Donner la courbure et la torsion de l'hélice circulaire

$$\begin{cases} x = R \cos t \\ y = R \sin t \\ z = \lambda t \end{cases}$$

## Surfaces.

On va donner ici ce qui n'est même pas un véritable aperçu du sujet, car on ne dispose pas des outils nécessaires.

Le plan tangent à une surface de l'espace euclidien  $\mathbf{R}^3$  est muni d'un produit scalaire, induit par celui de l'espace ambiant. C'est une forme quadratique définie positive. On lui donne le nom de première forme fondamentale.

Si l'on prend des coordonnées locales  $u, v$ , le carré scalaire  $\|d\vec{M}\|^2$  peut s'exprimer sous la forme

$$E(du)^2 + 2Fdudv + G(dv)^2$$

où  $EG - F^2 > 0$  et où  $du(h, k) = h$  etc. Les coefficients  $E, F, G$  dépendant du point.

La connaissance de la première forme fondamentale, i.e. la donnée de ce qu'on appelle une structure riemannienne, permet de calculer la longueur d'un arc de courbe et donc de poser le problème des *géodésiques*. On s'intéresse aux courbes tracées sur la surface qui sont critiques dans la recherche du minimum de la distance entre deux points donnés. Cela étant, l'existence d'une courbe réalisant le minimum n'est vraie que localement.

Pour une courbe générale, paramétrée par la longueur d'arc, tracée sur la surface, le vecteur  $d\vec{\tau}/ds$  se décompose en ses composantes suivant la normale à la surface et suivant le plan tangent, donnant naissance à la courbure normale et à la *courbure géodésique*.

Cette dernière est la courbure en projection sur le plan tangent. Les courbes critiques sont caractérisées par le fait que la courbure géodésique y est nulle en tout point; cela signifie que le vecteur  $d\vec{\tau}/ds$  est normal. C'est tout à fait conforme à l'intuition : en projection sur le plan tangent, le plus court chemin est sans courbure.

Disons un mot de la courbure normale en étudiant une surface au voisinage d'un point. Par le choix d'un repère orthonormé convenable, on peut supposer que ce point est l'origine et que la surface est donnée par l'équation

$$z = ax^2 + 2bxy + cy^2 + o(x^2 + y^2) .$$

Coupons cette surface par un plan normal, pour obtenir la courbe d'équation

$$\begin{cases} x = \lambda t \\ y = \mu t \\ z = (a\lambda^2 + 2b\lambda\mu + c\mu^2)t^2 + o(t^2) \end{cases}$$

où  $\lambda^2 + \mu^2 = 1$ .

On obtient ainsi l'expression de la *courbure normale* en fonction de la direction de la tangente à la courbe. C'est la courbure de la courbe considérée, donnée simplement par

$$2(a\lambda^2 + 2b\lambda\mu + c\mu^2) .$$

Cette forme quadratique, qui est l'équation d'une conique, prend des valeurs comprises entre  $K_1$  et  $K_2$ , les valeurs extrêmes, positives ou négatives, étant prises sur les axes.

Les valeurs extrêmes  $K_1$  et  $K_2$  de la courbure normale sont liées au plongement dans  $\mathbf{R}^3$ . En revanche le produit  $K_1K_2$ , appelé *courbure totale* ou *courbure de Gauss*, ne dépend que de la structure riemannienne : cela se démontre.

## Appendice : prérequis généraux d'algèbre linéaire

### Algèbre linéaire - bases

**Exercice.** Soit  $E$  un espace vectoriel réel. On considère un système  $b = (\mathbf{e}_1, \dots, \mathbf{e}_n)$  de vecteurs de  $E$ . On lui associe l'application linéaire

$$u_b : \mathbf{R}^n \rightarrow E$$

définie par

$$u_b(x_1, \dots, x_n) = x_1 \mathbf{e}_1 + \dots + x_n \mathbf{e}_n .$$

1) A quelle condition cette application est-elle surjective, injective, bijective? Que vaut  $u_b$  pour la base canonique de  $\mathbf{R}^n$ ?

**Réponses :** un système générateur, libre, une base; l'application identique.

**Exercice.** Si  $M$  est une matrice  $n \times p$ ; on identifie  $M$  à l'application linéaire de  $\mathbf{R}^p$  dans  $\mathbf{R}^n$  qui au vecteur  $(x_1, \dots, x_p)$  associe le vecteur  $(y_1, \dots, y_n)$  obtenu par

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{np} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$$

suivant les règles usuelles. Montrer que c'est l'application linéaire associée au système des vecteurs colonnes. Quel sens a le produit  $MM'$  où  $M'$  est une matrice  $p \times q$  ?

**Réponse :** la composition des applications.

**Exercice.** Soit  $f : E \rightarrow E'$  une application linéaire. Comment définir sa matrice dans des bases  $b$  de  $E$  et  $b'$  de  $E'$ ?

Soit  $M$  une matrice  $n \times p$ ; comment écrire la matrice obtenue en changeant la base canonique pour une base  $b$  au départ? ou à l'arrivée? Ou les deux?

**Réponses :** voir schéma ci-dessous.

Base  $b$  de  $E$  :

$$\mathbf{R}^n \xrightarrow[u_b]{\cong} E$$

Matrice  $n \times p$  :

$$\mathbf{R}^p \xrightarrow[M]{\rightarrow} \mathbf{R}^n$$

Matrice  $M$  de  $f : E \rightarrow E'$  dans les bases  $b, b'$  :

$$\begin{array}{ccc} \mathbf{R}^n & \xrightarrow[u_b]{\cong} & E \\ M \downarrow & & \downarrow f \\ \mathbf{R}^{n'} & \xrightarrow[u_{b'}]{\cong} & E' \end{array}$$

Changements de base où  $M$  devient  $M'$  :

$$\begin{array}{ccc} \mathbf{R}^p & \xrightarrow[u_b]{\cong} & \mathbf{R}^p & & \mathbf{R}^p \\ M' \searrow & & \downarrow M & & \downarrow M \\ & & \mathbf{R}^n & & \mathbf{R}^n \end{array} \quad \begin{array}{c} M' \\ \swarrow \\ \mathbf{R}^p \end{array} \xrightarrow[u_b]{\cong} \mathbf{R}^p$$

## Algèbre linéaire - groupe orthogonal

L'espace  $\mathbf{R}^n$  est muni du produit scalaire

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + \cdots + x_n y_n .$$

**Exercice.** Soit  $M$  une matrice  $n \times n$ . Montrer que les propriétés suivantes sont équivalentes.

- (i) L'image par  $M$  de la base canonique est orthonormée.
- (ii) L'image par  $M$  de toute base orthonormée est orthonormée.
- (iii)  $M$  conserve le produit scalaire.
- (iv)  ${}^t M \cdot M = 1$ .
- (v)  $M$  est inversible et  $M^{-1} = {}^t M$ .

On montrera les implications de (i) vers (iv), de (iv) vers (iii), de (iii) vers (ii) et de (ii) vers (i); on montrera enfin l'équivalence de (iv) et (v).

**Indication :**  $\langle u\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, {}^t u\mathbf{y} \rangle$ .

Une matrice ayant ces propriétés est dite orthogonale.

**Exercice.** Montrer que la transposée d'une matrice orthogonale est orthogonale.

**Exercice.** Montrer que les matrices orthogonales forment un sous-groupe  $O(n)$  de  $GL(n, \mathbf{R})$ .

**Exercice.** Montrer que  $M$  est orthogonale dès qu'elle transforme une base orthonormée donnée en une base orthonormée.

**Indication :**  $\mathbf{R}^n \xrightarrow{u} \mathbf{R}^n \xrightarrow{M} \mathbf{R}^n$ .

## Algèbre multilinéaire - calcul de déterminants

Il s'agit de déterminants en dimension 2 ou 3, mais dans un contexte géométrique, comme celui d'un déterminant jacobien. On note  $x, y \dots$  les fonctions coordonnées.

L'espace  $\mathbf{R}^2$  (resp.  $\mathbf{R}^3$ ), l'espace  $\bigwedge^2 \mathbf{R}^*$  (resp.  $\bigwedge^3 \mathbf{R}^*$ ) possède une base canonique

$$e^* = dx \wedge dy \quad \text{resp.} \quad dx \wedge dy \wedge dz .$$

C'est, par définition la forme *déterminant* canonique  $\det$ . C'est une forme bilinéaire (resp. trilinéaire) alternée; la valeur sur deux (resp. trois vecteurs)

$$\det(\vec{h}, \vec{k}) = (dx \wedge dy)(\vec{h}, \vec{k}) \quad \text{resp.} \quad \det(\vec{h}, \vec{k}, \vec{l}) = (dx \wedge dy \wedge dz)(\vec{h}, \vec{k}, \vec{l})$$

est égale à 1 sur la base canonique.

Le déterminant  $\det u$  d'un endomorphisme  $u$  est défini par

$$S(u(\vec{h}), u(\vec{k})) = \det u \cdot S(\vec{h}, \vec{k}) \quad \text{resp.} \quad S(u(\vec{h}), u(\vec{k}), u(\vec{l})) = \det u \cdot S(\vec{h}, \vec{k}, \vec{l})$$

pour toute forme alternée  $S$ ; en particulier si  $S$  est la forme déterminant canonique. Ainsi, lorsque  $M$  est une matrice, identifiée à un endomorphisme de  $\mathbf{R}^2$  ou  $\mathbf{R}^3$ , son déterminant  $\det u$  est celui des vecteurs colonnes, images par  $M$  de ceux de la base canonique.

Si l'on change la base canonique pour une base  $b$  quelconque, ce qui revient à se donner un isomorphisme  $u_b$  de  $\mathbf{R}^2$  ou  $\mathbf{R}^3$  dans lui-même, le déterminant dans la base  $b$  est donné par

$$\det_b = (\det u_b)^{-1} \cdot \det$$

où  $\det$  est toujours le déterminant canonique. Cela résulte du diagramme ci-dessous.

$$\begin{array}{ccc} & & \mathbf{R}^n \\ & \swarrow & \downarrow \text{vecteurs} \\ \mathbf{R}^n & \xrightarrow[u_b]{\sim} & \mathbf{R}^n \end{array}$$

On ne change donc pas le déterminant de vecteurs quand l'on remplace la base canonique par une base orthonormée directe; en effet  $u_b$  sera une matrice orthogonale, de déterminant  $\pm 1$ , mais en même temps de déterminant  $> 0$  puisque l'orientation est conservée.

Considérons pour finir le cas de la dimension 3; l'espace  $\mathbf{R}^3$  est muni de sa structure canonique d'espace euclidien orienté. Le produit vectoriel  $\vec{h} \wedge \vec{k}$  est le vecteur caractérisé par la propriété

$$\det(\vec{h}, \vec{k}, \vec{l}) = (\vec{h} \wedge \vec{k}) \cdot \vec{l}$$

dite du produit mixte. On ne le confondra pas avec un produit extérieur, même si le risque n'est pas grand : si règles de calcul sont — heureusement — semblables.

Ainsi le produit vectoriel peut-il servir à un calcul de déterminant. Cela vaut aussi en dimension 2; il suffit d'ajouter une troisième dimension. La formule

$$\det(\vec{h}, \vec{k}) = (\vec{h} \wedge \vec{k}) \cdot \vec{e}_z$$

fait du déterminant la composante suivant  $z$  du produit vectoriel.

## Appendice : longueurs

**Exercice.** Tous les espaces  $\mathbf{R}^n$  sont euclidiens et l'espace  $\mathbf{R}^3$  est orienté. On note  $D_r$  (resp.  $(D'_r)$ ) le disque ouvert (resp. fermé) de centre 0 et de rayon  $r$  de  $\mathbf{R}^2$ . On pose encore  $C_r = [a, b] \times D'_r$

On se donne une courbe gauche paramétrée définie par une fonction  $\vec{\gamma} : [a, b] \rightarrow \mathbf{R}^3$  de classe  $\mathcal{C}^2$ , injective et de dérivée partout non nulle; on étend  $\gamma$  de façon  $\mathcal{C}^2$  à un intervalle  $] \alpha, \beta [$  contenant  $[a, b]$  en prenant à gauche de  $a$  (resp. à droite de  $b$ ) son développement de Taylor à l'ordre 2 en  $a$  (resp.  $b$ ).

1) On admet qu'il existe un vecteur  $\vec{w}$  fixe qui n'est tangent à  $\gamma$  en aucun point. Construire, pour chaque  $t$  dans  $] \alpha, \beta [$ , un repère orthonormé direct  $\vec{\tau}(t), \vec{n}(t), \vec{p}(t)$  dont le premier vecteur soit le vecteur tangent unitaire positif en  $\vec{\gamma}(t)$  et qui dépendent de façon  $\mathcal{C}^1$  en  $t$ .

On définit une application  $\Phi$  de  $] \alpha, \beta [ \times \mathbf{R}^2$  dans  $\mathbf{R}^3$  par

$$\Phi(t, u, v) = \vec{\gamma}(t) + u\vec{n}(t) + v\vec{p}(t) .$$

2) Montrer qu'on peut choisir  $\rho > 0$  assez petit pour que  $\Phi$  soit  $\mathcal{C}^1$  difféomorphisme local en tout point de  $] a - \rho, b + \rho [ \times D_\rho$ .

3) On a choisi  $\rho$  et posé  $C = C_{\rho/2}$ . Dans le produit  $C \times C$ , on considère l'ensemble  $A$  des points  $(z, z')$  tels que  $z \neq z'$  et  $\Phi(z) = \Phi(z')$ ; montrer que c'est une partie fermée du produit.

4) Montrer que l'application  $\Phi$  est injective sur  $C_\epsilon$  pour  $\epsilon > 0$  assez petit; on pourra considérer la suite des  $A \cap (C_{1/n} \times C_{1/n})$  dont on montrera que l'intersection est vide.

5) On pose  $V_\epsilon = \Phi(C_\epsilon)$ . Montrer que

$$\text{mes}(V_\epsilon) = \int \int_{[a, b] \times D'_\epsilon} |\det(d\Phi)| dt du$$

pour  $\epsilon$  assez petit également, puis que

$$\lim_{\epsilon \rightarrow 0, \epsilon > 0} \frac{1}{\pi \epsilon^2} \text{mes}(V_\epsilon) = L(\gamma)$$

où  $L(\gamma)$  est la longueur de la courbe  $\gamma$ .

6) Etablir la dernière limite quand on remplace  $V_\epsilon$  par l'ensemble  $W_\epsilon$  des points dont la distance à la courbe est majorée par  $\epsilon$ .

7) Donner, au passage, l'expression du volume du tore.