Mémo de topologie

Jussieu, LM360

Si le contenu de ce cours est tout à fait classique, sa présentation est un peu inhabituelle. Le cours traditionnel, présenté dans les sections "théorie" de chaque chapitre, est assorti de nombreux commentaires informels ; les plus courts tiennent dans la marge (généreuse), les plus longs sont rangés dans une section à part.
D'autre part, pour aider le lecteur à s'approprier les notions, on a préféré remplacer les traditionnelles preuves par des "recettes de preuves". Celles-ci sont en fait des démonstrations abrégées, à charge pour le lecteur d'en écrire les détails. Ce travail de ré-écriture des preuves est fondamental, c'est probablement la meilleure façon d'assimiler le cours de topologie. Chacun des verbes d'action de ces recettes de preuves **est à prendre au pied de la lettre** !
Assimiler le cours de topologie signifie (1) savoir manipuler les définitions (notamment utiliser à bon escient les quantificateurs et les variables), (2) former des images mentales robustes des différentes notions (ouverts, fermés, intérieur, adhérence, *etc.*), et (3) comprendre l'utilité des différents concepts (espace métrique, compacité, connexité, complétude, *etc.*).

Table des matières

Tous les mots en bleu sont cliquables (dans la version électronique seulement...)

I Espaces métriques **4**
- I.1 Théorie . . . 4
  - (a) Ouverts, fermés, intérieur, adhérence . . . 4
  - (b) Applications continues . . . 6
  - (c) Valeurs d'adhérence des suites . . . 7
  - (d) Sous-espaces métriques . . . 8
  - (e) Distances équivalentes et topologiquement équivalentes . . . 8
  - (f) Produits d'espaces métriques . . . 9
  - (g) Espaces de fonctions . . . 10
  - (h) Espaces topologiques . . . 11
- I.2 Commentaires . . . 11
  - (a) Généralités des espaces métriques . . . 11
  - (b) Détails des esquisses de preuves . . . 13
  - (c) Qu'est-ce qu'un ouvert ? . . . 14
  - (d) Dualité ouverts/fermés . . . 16
  - (e) Notions métriques et topologiques . . . 16
  - (f) Sous-espaces métriques . . . 18
  - (g) Espaces produits, espaces de suites . . . 19
  - (h) Espaces de fonctions . . . 20
- I.3 Exercices . . . 20
- I.4 Aides et Réponses . . . 28

II Complétude **29**
- II.1 Théorie . . . 29
  - (a) Définitions, résultats . . . 29
  - (b) Le théorème du point fixe contractant . . . 30
  - (c) Le théorème de Baire . . . 31
- II.2 Commentaires . . . 31
  - (a) Histoire . . . 31
  - (b) Application I : existence et unicité des solutions d'équations différentielles 32
  - (c) Application II : construction d'ensemble fractals . . . 34
  - (d) Application III : une courbe de Peano . . . 34
  - (e) Applications du théorème de Baire . . . 34
- II.3 Exercices . . . 34

III Compacité **36**
- III.1 Théorie . . . 36
  - (a) Compacité . . . 36
  - (b) Compacité dans $\mathbb{R}^N$ . . . 37
  - (c) Compacité et recouvrements . . . 38
  - (d) Continuité uniforme . . . 40
  - (e) Compacité dans les espaces de fonctions . . . 40
- III.2 Commentaires . . . 41
  - (a) Continuité et continuité uniforme . . . 42
- III.3 Exercices . . . 43

IV Connexité **46**
- IV.1 Théorie . . . 46
  - (a) Connexité par arcs . . . 46
  - (b) Composantes connexes par arcs . . . 47
  - (c) Connexité . . . 47
  - (d) Compacts connexes . . . 48
- IV.2 Commentaires . . . 49
  - (a) Connexité et homéomorphismes . . . 49

IV.3 Exercices . . . 50

V Espaces vectoriels normés, espaces de Banach **51**
V.1 Théorie . . . 51
(a) Définitions . . . 51
(b) Continuité des applications linéaires . . . 52
(c) Dimension finie . . . 53
(d) Espaces de Banach . . . 55
(e) L'algèbre normée $\mathcal{L}(E)$ . . . 56
(f) Le théorème de Riesz . . . 58
V.2 Commentaires . . . 59
(a) Pièges et délices de la dimension infinie . . . 59
V.3 Exercices . . . 60

VI La différentielle **61**
VI.1 Théorie . . . 61
(a) Dérivée . . . 61
(b) Différentielle . . . 62
(c) Exemples I : La différentielle en coordonnées . . . 64
(d) Exemples II : la différentielle sans coordonnées . . . 66
(e) Opérations . . . 68
(f) Applications de classe C^1 . . . 70
VI.2 Commentaires . . . 72
(a) La différentielle . . . 72
(b) Gradient et optimisation . . . 74
(c) Différence entre continu et classe C^1 . . . 74
VI.3 Exercices . . . 75

VII Différentielles d'ordre supérieur **76**
(a) $D(D(f))$ et $D^2(f)$ . . . 76
(b) Formule de Taylor à l'ordre 2 . . . 77
(c) Dérivées partielles secondes . . . 77
(d) Extrema . . . 78
(e) Dérivées d'ordre supérieur . . . 79
VII.1 Exercices . . . 79

VIII Inversion locale, fonctions implicites **80**
VIII.1 Théorie . . . 80
(a) Difféomorphismes, le Théorème d'Inversion Locale . . . 80
(b) Le Théorème des Fonctions Implicites . . . 83
VIII.2 Commentaires . . . 86

IX Sous-variétés, hypersurfaces **87**
IX.1 Sous-variétés, sous-espace vectoriel tangent . . . 87
IX.2 Exemples fondamentaux . . . 88
(a) Graphes . . . 88
(b) Equation implicite d'une hypersurface . . . 88
(c) Hypersurfaces paramétrées . . . 89
IX.3 Extrema liés . . . 89

X Annexe : schémas **90**

I Espaces métriques

I.1 Théorie

(a) Ouverts, fermés, intérieur, adhérence

Soit X un ensemble. Une *distance* sur X est une application d qui associe à deux éléments x, y de X un réel positif ou nul $d(x, y)$ vérifiant les trois axiomes :

1. *séparation :* $d(x, y) = 0 \Leftrightarrow x = y$,
2. *symétrie :* $d(x, y) = d(y, x)$,
3. *inégalité triangulaire :* $d(x, y) \leq d(x, z) + d(z, y)$.

Un espace métrique (X, d) est un ensemble X muni d'une distance d. La distance euclidienne sur $\mathbb{R}^n$ est un exemple fondamental de distance.

Nous verrons plus loin d'autres exemples fondamentaux, plus abstraits, comme les espaces de fonctions, de suites, de matrices.

Dans un espace métrique (X, d) on définit les notions suivantes. La *boule ouverte* de centre x et de rayon $\varepsilon > 0$ est l'ensemble $B(x, \varepsilon) = \{y \in X \mid d(x, y) < \varepsilon\}$. Ainsi, les propriétés "$y \in B(x, \varepsilon)$" et "$d(x, y) < \varepsilon$" sont équivalentes. Une partie O de X est *un ouvert* de X si tout point de O est centre d'une boule incluse dans O :

$$\forall x \in O \quad \exists \varepsilon > 0 \quad B(x, \varepsilon) \subset O.$$

Exercice 1.— Montrer que toute boule ouverte est un ouvert de X. *Aide : étant donné un point dans une boule ouverte B, il s'agit de trouver une boule b centrée en ce point et contenue dans B. Faire un dessin pour trouver un rayon de b qui convient.*

Proposition. *L'intersection d'un nombre fini de parties ouvertes est une partie ouverte. La réunion d'une famille quelconque (finie ou infinie) de parties ouvertes est une partie ouverte.*

Recette de preuve.— L'intersection d'un nombre fini d'ouverts est un ouvert : **prendre** un point dans l'intersection, **écrire** la définition de l'intersection, puis **utiliser** que chacun des ensembles est ouvert. (On peut commencer par deux ouverts).

Détails p13

L'union d'une famille quelconque d'ouverts est ouverte : **prendre** un point dans l'union, **utiliser** la définition de l'union, puis la définition d'un ouvert.

Une suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de X *converge* vers un élément ℓ de X (ou *a pour limite* ℓ) si pour toute boule ouverte centrée en a, tous les termes de la suite à partir d'un certain rang sont dans la boule. Ceci revient à demander que la suite réelle $(d(x_n, \ell))_{n\in\mathbb{N}}$ tend vers 0. Les *fermés* de X sont définis comme les ensembles complémentaires des ouverts.

Proposition (critère séquentiel)**.** *Une partie F de X est fermée si et seulement si pour toute suite (x_n) d'éléments de F qui est supposée convergente, la limite ℓ appartient à F.*

Un critère *séquentiel* est un critère utilisant les suites.

Recette de preuve.— Commençons par le sens direct. Soit F un ensemble fermé, et $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments de F qui converge vers un point ℓ de X ; il s'agit de voir que ℓ appartient à F. On raisonne par l'absurde en supposant que ℓ appartient à l'ensemble $X \setminus F$. **Utiliser** alors la définition des fermés, puis **appliquer** la définition des ouverts au point ℓ pour obtenir

l'existence d'une boule ; **appliquer** enfin la définition de la limite à la suite $(x_n)_{n\in\mathbb{N}}$ à cette boule. **Conclure** que certains termes de la suite ne sont pas dans F, ce qui est absurde.

Pour la réciproque, on commence par montrer que si un ensemble O n'est pas ouvert, alors il existe une suite d'éléments de $X \setminus O$ qui tend vers un élément ℓ de O. Pour montrer ceci, **écrire** la négation de la définition d'ouvert. Ceci fournit un point ℓ de O vérifiant une certaine propriété qui commence par "$\forall\varepsilon > 0$ *etc.*" : **écrire** cette propriété. En appliquant cette propriété avec $\varepsilon = 1/n$ pour chaque entier $n > 0$, **construire** la suite recherchée. **En déduire** enfin que si F n'est pas fermé, il existe une suite d'éléments de F qui converge vers un élément de $X \setminus F$.

On voit ici comment construire une suite à l'aide d'une propriété du type "$\forall\varepsilon > 0$ *etc.*."

Soit E une partie quelconque de X. L'*intérieur* de E est l'ensemble des points de E qui sont le centre d'une boule incluse dans E. L'*adhérence* de E est l'ensemble des points de X tels que toute boule centrée en ce point rencontre E. La *frontière* de E est l'ensemble des points de X tels que toute boule centrée en ce point rencontre à la fois E et son complémentaire. En symboles :

Les commentaires, section (c), peuvent aider à comprendre ces notions.

$$x \in \mathrm{Inte}(E) \Leftrightarrow (\exists\varepsilon > 0 \quad B(x,\varepsilon) \subset E).$$

$$x \in \mathrm{Adhe}(E) \Leftrightarrow (\forall\varepsilon > 0 \quad B(x,\varepsilon) \cap E \neq \emptyset).$$

$$x \in \mathrm{Fr}(E) \Leftrightarrow (\forall\varepsilon > 0 \quad B(x,\varepsilon) \cap E \neq \emptyset \text{ et } B(x,\varepsilon) \cap (X \setminus E) \neq \emptyset).$$

D'après la définition, on a aussi $\mathrm{Fr}(E) = \mathrm{Adhe}(E) \cap \mathrm{Adhe}(X \setminus E)$.

Une partie E de X est *dense dans X* si $\mathrm{Adhe}(E) = X$. Elle est *d'intérieur vide dans X* si $\mathrm{Inte}(E) = \emptyset$.

Exercice 2.—

1. Montrer que $A \subset B \Rightarrow \mathrm{Inte}(A) \subset \mathrm{Inte}(B)$.

2. Montrer que O est ouvert si et seulement si $\mathrm{Inte}(O) = O$.

Proposition (Caractérisation topologique)**.** *L'intérieur de E est ouvert, et il contient tout ouvert inclus dans E. L'adhérence de E est fermée, et elle est incluse dans tout fermé contenant E.*

On résume ces propriétés en disant que l'intérieur de E est le plus grand ouvert inclus dans E, et l'adhérence de E est le plus petit fermé contenant E.

Recette de preuve.—

Quelle est la démarche pour montrer que l'intérieur de E est un ouvert ?

Réponse p28

Soit maintenant x un point de l'intérieur de E : il existe $\varepsilon > 0$ tel que $B(x,\varepsilon) \subset E$. En utilisant les propriétés obtenues dans l'exercice précédant, **montrer** que $B(x,\varepsilon) \subset \mathrm{Inte}(E)$. Conclure.

En utilisant à nouveau les propriétés démontrées en exercice, **montrer** que l'intérieur de E contient tout ouvert O inclus dans E.

Les propriétés de l'adhérence peuvent s'en déduire par passage au complémentaire.

Exercice 3.— (unicité de la limite) Montrer que si une suite (x_n) converge à la fois vers ℓ_1 et vers ℓ_2 alors $\ell_1 = \ell_2$.

(b) Applications continues

Soient X, Y deux espaces métriques, dont les distances seront notées d_X et d_Y. Une application $f : X \to Y$ est *continue* en un point x_0 de X si :

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in X, \quad (d_X(x, x_0) < \delta \Rightarrow d_Y(f(x), f(x_0)) < \varepsilon).$$

En d'autres termes, f est continue en x_0 si toute boule de centre $f(x_0)$ contient l'image d'une boule de centre x_0, ou encore si l'image réciproque de toute boule de centre $f(x_0)$ contient une boule de centre x_0. L'application f est dite *continue* si elle est continue en tout point x_0 de X.

Proposition. *La composée de deux fonctions continues est continue.*

Recette de preuve.— Soient $f : X \to Y$ et $g : Y \to Z$ deux applications (X, Y, Z sont trois espaces métriques). On suppose que f est continue en un point x de X, et que g est continue au point $f(x)$. Il s'agit de montrer, en utilisant la définition de la continuité, que $g \circ f$ est continue, quelle est la démarche ?

Réponse p28

On prend donc un $\varepsilon > 0$. En appliquant l'une des deux hypothèses (f continue ou g continue ?), **obtenir** un $\delta > 0$. En appliquant l'autre hypothèse avec ce δ à la place du ε, **obtenir** un $\gamma > 0$. Si on a fait le bon choix, ce γ doit vérifier ce qui est demandé dans la définition de "$g \circ f$ est continue" ; le **vérifier** (ou changer de choix !).

Proposition (Caractérisation topologique)**.** *$f : X \to Y$ est continue si et seulement si l'image réciproque de tout ouvert de Y est un ouvert de X.*

Recette de preuve.— On suppose f continue, on prend un ouvert W de Y, il s'agit de montrer que $V := f^{-1}(W)$ est un ouvert de X : autrement dit que tout point de V est centre d'une boule incluse dans V. **Partir** d'un point de V, autrement dit d'un point dont l'image est dans W. **Appliquer** la définition des ouverts à W pour obtenir une certaine boule de Y, **appliquer** à cette boule la définition de la continuité de f pour obtenir une boule dans X. **Montrer** enfin que cette boule est incluse dans V.

L'image réciproque d'une partie W de Y est notée $f^{-1}(W)$; par définition, c'est l'ensemble des points de X qui ont leur image dans W.

Pour la réciproque, on suppose que l'image réciproque de tout ouvert est un ouvert, et on veut en déduire que f est continue. Etant donné un x_0 et un $\varepsilon > 0$ il s'agit de trouver un certain δ ; pour obtenir ce δ, **appliquer** l'hypothèse à l'image réciproque de la boule ouverte $B(f(x_0), \varepsilon)$.

Proposition (Caractérisation séquentielle)**.** *$g : Y \to Z$ est continue en un point y si et seulement si pour toute suite d'éléments $(y_n)_{n\in\mathbb{N}}$ de Y convergent vers y, la suite $(g(y_n))$ converge dans Z vers $g(y)$.*

On a choisi de noter la fonction g pour faire ressortir l'analogie avec la composition des fonctions continues.

Recette de preuve.— Rédiger le sens direct en vous inspirant de la preuve de la continuité de la composée de deux fonctions continues (la suite (x_n) joue le rôle de la fonction f, et la convergence de cette suite joue le rôle de la continuité de f). On démontre le sens réciproque par contrapposition. **Écrire** la contrapposée de la réciproque. **Écrire** la négation de la continuité de g. Ceci fournit un $\varepsilon > 0$ qui vérifie une propriété du type "$\forall \delta > 0$ *etc.*" ; **appliquer** cette propriété à $1/n$ pour **construire** une suite (y_n). **Vérifier** qu'on a ainsi construit une suite convergant vers y dont l'image par g ne converge pas vers $g(y)$.

On retrouve la technique de construction de suites utilisée dans la preuve de la caractérisation séquentielle des fermés.

Une application $f : X \to Y$ est *lipschitzienne* s'il existe une constante $k > 0$ telle que, pour tout $x, x' \in X$,

$$d_Y(f(x), f(x')) \leq k d_X(x, x').$$

Une application lipschitzienne "n'éloigne pas trop les points les uns des autres".

Exercice 4.— Montrer que toute application lipschitzienne est continue.

Exercice 5.— Soit $f : I \to \mathbb{R}$ une application de classe C^1 définie sur un intervalle réel I. On suppose que la dérivée de f est bornée. Intérpréter l'inégalité des accroissements finis avec la définition précédente.

Exercice 6.— Nous avons démontré plus haut que la composée de deux fonctions continues est continue. Donner une autre preuve en utilisant la caractérisation séquentielle. Donner une troisième preuve en utilisant la caractérisation topologique.

(c) Valeurs d'adhérence des suites

Soit (x_n) une suite d'éléments de X. Un élément ℓ de X est une *valeur d'adhérence* de la suite s'il existe des termes de la suite, de rangs arbitrairement grands, et arbitrairement proches de ℓ :

$$\forall \varepsilon > 0 \ \forall N > 0 \ \exists n \geq N, \quad d(x_n, \ell) < \varepsilon.$$

Extraire une suite consiste donc à "oublier" certains termes de la suite ; en oubliant les termes de rang pair, par exemple, on extrait la suite $(x_{2n+1})_{n \in \mathbb{N}}$.

On appelle *suite extraite* de la suite (x_n) une suite

$$(x_{n_1}, x_{n_2}, x_{n_3}, \ldots) \text{ avec } n_1 < n_2 < n_3 < \ldots .$$

Proposition (Caractérisation topologique)**.** *Un élément ℓ est une valeur d'adhérence de la suite $(x_n)_{n \in \mathbb{N}}$ si et seulement si il existe une suite extraite $(x_{n_k})_{k \in \mathbb{N}}$ convergent vers ℓ.*

Recette de preuve.— La preuve ressemble à celle du critère séquentiel des fermés : comme dans la réciproque de celui-ci, il s'agit de construire une suite. Supposons que ℓ soit une valeur d'adhérence de (x_n). Pour chaque entier $k > 0$, on applique la définition avec $\varepsilon = 1/k$ et $N = 0$, on obtient un entier que l'on note n_k. De cette façon on construit une suite (x_{n_k}) qui converge vers ℓ. Malheureusement cette construction ne marche pas vraiment, parce que rien n'assure que la suite (n_k) soit strictement croissante, comme il est demandé dans la définition d'une suite extraite. (Il faut dire que c'était un peu étrange d'appliquer la définition de valeur d'adhérence uniquement avec $N = 0$, non ?).

Pour obtenir une suite (n_k) qui soit strictement croissante, **procéder** par récurrence. Après avoir obtenu n_1 comme avant, **appliquer** la définition de valeur d'adhérence avec $\varepsilon = 1/2$ et $N = n_1 + 1$ (au lieu de 0). On obtient un entier que l'on note n_2, qui est strictement plus grand que n_1. **Appliquer** à nouveau la définition de valeur d'adhérence avec $\varepsilon = 1/3$ et $N = n_2 + 1$ pour obtenir un entier n_3, et ainsi de suite. **Construire** ainsi une suite extraite qui converge vers ℓ.

Ceci est une technique classique de construction de suites extraites.

La réciproque est plus facile : lorsqu'on a une suite extraite (x_{n_k}), la suite (n_k) tend nécessairement vers $+\infty$, puisque c'est une suite d'entiers strictement croissante. Si cette suite extraite converge vers ℓ, alors on trouve des termes x_n arbitrairement proches de ℓ avec n arbitrairement grand, comme demandé dans la définition d'une valeur d'adhérence.

(d) Sous-espaces métriques

Soit d'abord (X, d) un espace métrique, et Y une partie de X. La restriction de la distance d à Y vérifie évidemment les trois axiomes de distance ; ainsi, (Y, d) est aussi un espace métrique, on dit que c'est un *sous-espace métrique* de (X, d).

Cette section décrit une façon simple de fabriquer plein de nouveaux espaces métriques.

Proposition.

1. *Une suite $(y_n)_{n\in\mathbb{N}}$ de points de Y converge dans (Y, d) si et seulement si elle converge dans (X, d) vers un point ℓ qui appartient à Y.*
2. *Les ouverts de (Y, d) sont les intersections des ouverts de (X, d) avec Y.*

Recette de preuve.— Pour différencier les deux espaces métriques qui interviennent dans cette proposition, on note B_Y et B_X les boules dans Y et dans X. Pour la suite, remarquons que la définition des boules dans Y est

$$B_Y(y_0, R)=\{y \in Y, d(y, y_0) < R\} = B_X(y_0, R) \cap Y.$$

Le premier point suit immédiatement de la définition de la convergence dans (Y, d). Pour le second point, il y a deux choses à montrer. Tout d'abord, étant donné un ouvert O de (X, d), il faut voir que $O \cap Y$ est un ouvert de (Y, d) : **montrer** ceci à l'aide de la définition d'un ouvert. Enfin, étant donné un ouvert O de (Y, d), il faut trouver un ouvert O' de (X, d) tel que $O = O' \cap Y$. **Montrer** ceci en utilisant que O est une réunion de boules ouvertes de Y (exercice 25).

(e) Distances équivalentes et topologiquement équivalentes

Une distance d est *équivalente* à une autre distance d' définie sur le **même** ensemble X s'il existe une constante $k \geq 1$ telle que l'encadrement

$$\frac{1}{k}d(x, y) \leq d'(x, y) \leq kd(x, y)$$

est vérifié pour tout x, y dans X. Les distances sont dites *topologiquement équivalentes* si l'ensemble des parties ouvertes définies à l'aide de d est le même que l'ensemble des parties ouvertes définies à l'aide de d'. Ces définitions, et la proposition qui suit, nous seront très utiles pour la section suivante sur les espaces produits.

Voir les commentaires p16.

Proposition. *Deux distances d, d' qui sont équivalentes sont topologiquement équivalentes.*

Recette de preuve.— Soient d, d' deux distances équivalentes sur un ensemble X. En utilisant la bonne inégalité dans la définition, **montrer** que toute boule, pour la distance d, centrée en un point x, contient une boule, pour la distance d', centrée au même point. **En déduire** qu'un ensemble ouvert pour l'une des deux distances est aussi ouvert pour l'autre.

(f) Produits d'espaces métriques

Cette section décrit encore une façon de fabriquer de nouveaux espaces métriques. Comme pour les sous-espaces métriques, nous décrirons les suites convergentes et les ouverts de ces nouveaux espaces.

On considère maintenant deux espaces métriques (X, d_X) et $Y, d_Y)$. On peut munir l'ensemble produit

$$X \times Y = \{(x, y) \mid x \in X, y \in Y\}$$

Sur cet ensemble on définit les trois distances suivantes :

$$d_1((x, y), (x', y')) := d_X(x, x') + d_Y(y, y').$$

$$d_2((x, y), (x', y')) := \left((d_X(x, x')^2 + (d_Y(y, y')^2)\right)^{\frac{1}{2}}$$

$$d_\infty((x, y), (x', y')) := \max\left(d_X(x, x'), d_Y(y, y')\right).$$

(x, y) et (x', y') sont deux éléments de l'ensemble $X \times Y$, et chacune des trois formules leur associe un nombre positif, comme attendu d'une distance sur $X \times Y$.

Proposition. *Les applications d_1, d_2, d_∞ sont des distances sur l'ensemble $X \times Y$.*

On obtient ainsi trois espaces métriques appelés *espaces métriques produits* de (X, d_X) et (Y, d_Y).

Recette de preuve.— Dans les trois cas, l'inégalité triangulaire est le seul des trois axiomes qui n'est pas immédiat. **Montrer** l'inégalité triangulaire pour la distance d_1 (bien sûr, il faut utiliser l'inégalité triangulaire dans les espaces métriques (X, d_X) et (Y, d_Y)). Pour la distance d_{infty}, commencer par **écrire** l'inégalité recherchée. Il s'agit de majorer $d_\infty((x, y), (x'', y'')) = \max\left(d_X(x, x''), d_Y(y, y'')\right)$; pour cela, majorer séparément $d_X(x, x'')$ et $d_Y(y, y'')$ par la quantité voulue. Enfin, montrer l'inégalité triangulaire pour la distance d_2, en utilisant l'inégalité de Cauchy-Schwarz.

Proposition. *Les trois distances d_1, d_2, d_∞ définies sur l'espace produit $X \times Y$ sont équivalentes.*

Voir les commentaires sur les espaces produits.

Recette de preuve.— On oublie momentanément le problème de départ ; on considère deux nombres positifs a, b, et les quantités

$$A_1 = a + b, \quad A_2 = (a^2 + b^2)^{\frac{1}{2}}, \quad A_\infty = \max(a, b).$$

Donner une majoration de A_1 par un multiple de A_∞, et une majoration de A_∞ par un multiple de A_1. Revenir au problème de départ, **appliquer** vos majorations à a et b bien choisis, et **en déduire** que pour tous points P_0, P dans $X \times Y$,

$$d_1(P_0, P) \leq 2d_\infty(P_0, P) \text{ et } d_\infty(P_0, P) \leq d_1(P_0, P).$$

Procéder de même pour d_∞ et d_2. En déduire enfin que d_1 et d_2 sont équivalentes.

Un *pavé ouvert* de $X \times Y$ est un ensemble du type $U \times V$ où U est un ouvert de (X, d_X) et V un ouvert de (Y, d_Y).

Puisque les trois distances d_1, d_2, d_∞ sont équivalentes, elles définissent la même notion de suites convergentes et les mêmes ouverts. La proposition est valable pour n'importe laquelle de ces trois distances.

Proposition (suites convergentes et ouverts d'un espace produit)**.**

1. *Une suite $(x_n, y_n)_{n\in\mathbb{N}}$ d'éléments de $X \times Y$ converge vers un élément (x_∞, y_∞) si et seulement si la suite $(x_n)_{n\in\mathbb{N}}$ converge vers x_∞ dans (X, d_X) et la suite $(y_n)_{n\in\mathbb{N}}$ converge vers y_∞ dans (Y, d_Y).*

2. Les ouverts de l'espace produit $X \times Y$ sont les réunions de pavés ouverts.

Recette de preuve.— Déduire le premier point du fait qu'étant données deux suites (a_n) et (b_n) de nombres positifs, la suite $(\max(a_n, b_n))_{n\in\mathbb{N}}$ concerge vers 0 si et seulement si chacune des deux suites converge vers 0.

Pour le second point, on a le choix de la distance : nous choisissons de travailler avec d_∞. **Écrire** la définition d'une boule dans $X \times Y$ pour cette distance, et **vérifier** que c'est le produit d'une boule dans (X, d_X) et d'une boule dans (Y, d_Y) ; en particulier, c'est un pavé ouvert. **En déduire** que tout ouvert de $(X \times Y, d_\infty)$ est réunion de pavés ouverts. Il reste à voir la réciproque, à savoir que toute réunion de pavés ouverts est un ouvert. Il suffit de voir que tout pavé ouvert est un ouvert (pourquoi ?). **Vérifier** qu'un pavé ouvert satisfait à la définition d'un ouvert dans $(X \times Y, d_\infty)$, en utilisant la définition d'un pavé ouvert, la définition d'un ouvert dans (X, d_X) et celle d'un ouvert dans (Y, d_Y).

Les formules définissant d_∞, d_1, d_2, et les propriétés correspondantes, se généralisent sans difficulté au produit d'un nombre fini d'espaces métriques $X_1, ..., X_n$.

Exercice 7.—(utile) Soit (X, d) est un espace métrique. Montrer que l'application distance $d : X \times X \to \mathbb{R}$ est 1-lipschitzienne lorsqu'on munit $X \times X$ de la distance d_1. *Aide : transcrire la définition générale de 1-lipschitzienne à ce contexte particulier (commencer par se demander comment s'exprime la distance au départ, et celle à l'arrivée ?). Le plus dur est fait, on doit maintenant montrer une inégalité impliquant quatre points x_0, y_0, x, y de X. Ecrire cette inégalité, se débarrasser de la valeur absolue. Montrer les deux inégalités.*

La distance associe à deux points x, y un nombre réel, c'est donc bien une application de $X \times X$ dans $\mathbb{R}$.

(g) Espaces de fonctions

Cette section explore encore une façon de fabriquer un nouvel espace métrique.

Soient X un ensemble et (Y, d_Y) un espace métrique. Une application $f : X \to Y$ est *bornée* s'il existe $M > 0$ tel que pour tout $x, x' \in X$, $d_Y(f(x), f(x')) \leq M$. L'ensemble des applications bornées de X dans Y est noté $\mathcal{B}(X, Y)$. Étant données deux applications bornées f et g, on définit

$$d_\infty(f, g) := \sup_{x\in X} d_Y(f(x), g(x)).$$

On est obligé de se restreindre aux applications *bornées* pour que cette formule produise bien un nombre (et pas $+\infty$).

Proposition. *La fonction d_∞ est une distance sur l'ensemble $\mathcal{B}(X, Y)$.*

Recette de preuve.— La symétrie est évidente. La séparation est facile. Il reste l'inégalité triangulaire. **Écrire** l'inégalité recherchée qui concerne trois fonctions f, g, h. Pour un élément x fixé dans X, **majorer** la distance $d_Y(f(x), h(x))$ par la quantité voulue. **Conclure** en "passant au Sup" en x dans l'inégalité obtenue.

Noter la similitude avec la distance définie sur les espaces produit.

La notion de convergence des suites dans l'espace métrique $(\mathcal{B}(X, Y), d_\infty)$ est appelée *convergence uniforme*. Remarquons qu'une suite $(f_n)_{n\in\mathbb{N}}$ d'éléments de cet espace métrique converge (uniformément) vers un élément f si et seulement si

$$\forall \varepsilon > 0 \;\; \exists N > 0 \;\; \forall n \geq N \;\; \forall x \in X \;\; d_Y(f_n(x), f(x)) < \varepsilon.$$

Supposons maintenant que l'ensemble X soit lui aussi muni d'une distance d_X. On définit l'ensemble $C_b(X, Y)$ des fonctions continues et bornées et X dans Y ; c'est un sous-ensemble de $\mathcal{B}(X, Y)$.

Théorème. *Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions continues convergeant uniformément vers une fonction f dans $\mathcal{B}(X,Y)$. Alors f est continue.*

Exercice 8.— En utilisant la caractérisation séquentielle des fermés, donner une autre formulation de ce théorème. Réponse p28

Recette de preuve.— Soit $x_0 \in X$, il s'agit de montrer que f est continue au point x_0. Pour un x quelconque dans X et un entier n quelconque, l'inégalité triangulaire répétée (exercice 18) entraîne

L'exercice 42 peut aider à comprendre cette preuve.

$$d_Y(f(x_0), f(x)) \leq d_Y(f(x_0), f_n(x_0)) + d_Y(f_n(x_0), f_n(x)) + d_Y(f_n(x), f(x)).$$

L'idée de la preuve consiste à remarquer que (1) d'après l'hypothèse de convergence uniforme, lorsque n est très grand, le premier et le dernier terme de cette somme sont petits, et que (2) d'après l'hypothèse de continuité, le terme central est petit si x est assez proche de x_0. Il s'agit maintenant de transformer cette idée en preuve précise, et notamment de choisir tous les objets dans le bon ordre.

On fixe donc un x_0, et aussi un $\varepsilon > 0$, que cherche-t-on alors ? **Appliquer** l'hypothèse de convergence uniforme avec $\varepsilon/3$. Pour le N obtenu, appliquer la continuité de f_N avec $\varepsilon/3$. Conclure. Réponse p28

Dans quel ordre a-t-on introduit les variables $\delta, \varepsilon, N, x, x_0$ de cette preuve ? Réponse p28

(h) Espaces topologiques

Une *topologie* sur un ensemble X est la donnée d'une famille $\mathcal{O}$ de parties qui est stable par intersection finie et par union quelconque et qui contient X et $\emptyset$:

1. $\forall O_1, O_2 \in \mathcal{O}, \quad O_1 \cap O_2 \in \mathcal{O}$;
2. $\forall I \subset \mathcal{O}, \quad \left(\bigcup_{O\in I} O\right) \in \mathcal{O}$;
3. l'ensemble vide et X appartiennent à $\mathcal{O}$.

Les éléments de $\mathcal{O}$ sont appelés *ouverts de la topologie*.

Proposition. *Les ouverts d'un espace métrique (X,d) forment une topologie sur X.*

Cette proposition n'est rien d'autre qu'une reformulation de la première proposition de ce texte.

Les espaces topologiques forment un cadre plus général que les espaces métriques, dans lequel on peut encore définir la plupart des notions vues plus haut : suite convergente, intérieur, adhérence, frontière, application continue, ... Pour cela, on prend comme définition les caractérisations topologiques que nous avons décrites. Pour plus de détails, se référer à Wikipedia ou à un livre de topologie. Ce cadre plus abstrait est souvent indispensable ; par exemple, la notion d'espace topologique quotient n'a pas d'équivalent dans les espaces métriques.

I.2 Commentaires

(a) Généralités des espaces métriques

En étudiant les séries de Fourier, au XIXème siècle, les mathématiciens font d'abord apparaître l'importance des propriétés topologiques de certains sous-ensemble de $\mathbb{R}$ (Dirichlet 1829, Cantor 1872). Plus tard, l'étude des propriétés de

convergence des suites de fonctions révèle des analogies très fortes avec les suites de nombres. Dans sa thèse, soutenue en 1906, Maurice Fréchet définit pour la première fois la notion générale d'espace métrique. Jusque là les mathématiciens travaillaient toujours avec un type d'objet bien défini : nombre, point, courbe, fonction, *etc.*. Ce nouveau cadre abstrait permet de formaliser les analogies entre suites de nombres et suites de fonctions, en englobant en une seule théorie des propriétés communes aux différents types d'objets, vus comme des points d'un espace métrique.

Le premier exemple d'espace métrique est donné par la droite réelle $\mathbb{R}$, muni de la distance définie par $d(x, y) = |x - y|$. Plus généralement, l'ensemble $\mathbb{R}^N$ muni de la distance euclidienne, celle qui correspond à l'intuition de distance qu'on peut avoir en dimension deux ou trois, est un espace métrique.

Nous avons vu dans ce chapitre qu'on pouvait munir certains espaces de fonctions d'une distance ; l'étude des propriétés métriques et topologiques des espaces de fonctions est central dans toute l'analyse moderne. Le théorème du point fixe dans les espaces complets, par exemple, permet de démontrer l'existence et l'unicité des solutions de certaines équations différentielles ou d'équations aux dérivées partielles (voir plus loin).

Voici un exemple assez différent. La distance de Hausdorff est définie sur l'ensemble des fermés de $\mathbb{R}^N$; il s'agit donc d'une façon de préciser l'idée que deux parties du plan sont proches l'une de l'autre. Cet espace métrique est un bon cadre pour construire des objets fractals, et en particulier pour comprendre les techniques de compression fractale.

Voici encore un autre exemple. Les mesures sont des objets fondamentaux en mathématiques, à la base à la fois des théories de l'intégration et des probabilités. La distance de Wasserstein est définie sur des espaces de mesures, et joue un rôle dans la théorie du transport optimal ; intuitivement, on peut voir chaque mesure comme un tas de sable, et la distance de Wasserstein mesure le travail à effectuer pour déplacer un tas de sable d'une position à une autre. La distance de Hausdorff et la distance de Wasserstein apparaîssent par exemple dans cet article de vulgarisation qui traite de l'inférence géométrique, c'est-à-dire de la façon dont on peut retrouver les propriétés géométriques d'un objet à partir d'une photographie ou d'un scanner.

Encore un exemple : la distance de Hamming est définie sur l'ensemble des suites finies de zéros et de uns, elle compte simplement le nombre de différences : par exemple, la distance de Hamming entre les suites $(0, 1, 0, 1, 0, 1, 0, 1)$ et $(0, 0, 0, 1, 1, 1, 1, 1)$ est 3. Cette distance joue un rôle important dans la théorie des codes correcteurs d'erreurs.

Il y a aussi des exemples un peu plus étranges (et moins utiles). En voici un premier. Si X est n'importe quel ensemble, on définit d en disant que $d(x, x) = 0$ pour tout élément x, et $d(x, y) = 1$ pour tous éléments x, y qui sont distincts. Cette fonction d est une distance sur X appelée *distance discrète*. En voici un second (un peu plus utile). Soit δ définie sur $\mathbb{R} \times \mathbb{R}$ par $\delta(x, y) = |\arctan(x) - \arctan(y)|$. Alors δ est une distance sur $\mathbb{R}$. Noter que l'espace métrique $(\mathbb{R}, \delta)$ est borné : pour tous x, y on a $\delta(x, y) < \pi$! ! On peut dire que cette distance "déforme $\mathbb{R}$ en le compressant énormément au voisinage de $\pm\infty$" : par exemple, on a $\delta(10, 100) \simeq 0,1$ et $\delta(100, 1000) \simeq 0,001$. Ces exemples sont un peu surprenants, mais ils rentrent aussi dans le cadre des espaces métriques. La puissance de ce cadre général, c'est

justement que tous les énoncés de ce chapitre sont valables dans tous les exemples que nous venons de mentionner (et bien d'autres encore).

Cette vaste généralité peut donner le vertige ; on ne peut pas vraiment avoir tous ces exemples présents à l'esprit quand on étudie les espaces métriques ! Mais en réalité, pour visualiser les propriétés des espaces métriques généraux, *il suffit la plupart du temps de penser aux sous-espaces du plan euclidien.* Le plan euclidien est un bon modèle de la feuille de papier sur laquelle on écrit, c'est probablement l'espace métrique le plus accessible à l'intuition.

Prenons un exemple. Pour voir que toute boule ouverte est un ouvert (exercice 1), on prend un point quelconque x dans une boule ouverte $B(x_0, R)$, et il s'agit (d'après la définition des ouverts) de trouver un rayon $\varepsilon > 0$ tel que la boule $B(x, \varepsilon)$ est incluse dans $B(x_0, R)$. Notre objectif fait intervenir quatre variables (x_0, R, x, ε), et un dessin est indispensable pour saisir le rôle de chacune. On représente le problème dans le plan (en se souvenant que la distance $d(x_0, x)$ est *strictement* inférieure à R). Le fait de dessiner permet non seulement de voir pourquoi le résultat recherché est plausible, mais aussi d'avoir une idée d'une valeur de ε qui semble convenir : sur ce dessin dans le plan, on voit que la boule de centre x et de rayon $R - d(x_0, x)$ est contenue dans $B(x_0, R)$. Il se trouve que ce qu'on voit ici dans le plan marche dans tout espace métrique, comme on peut le vérifier facilement à l'aide de l'inégalité triangulaire. Le dessin nous permet ici de trouver une valeur de ε, ce qui est la partie la plus difficile de l'exercice ; ensuite "il n'y a plus qu'à" vérifier que cette valeur convient. Bien que le dessin soit dans le plan, la vérification doit être valable pour tout espace métrique, on l'écrit en utilisant seulement nos trois axiomes et les propriétés qui en découlent.

Faire le dessin !

Exercice 9.—

1. Montrer que la distance de Hamming sur l'ensemble des suites de 8 bits (un bit est un symbole 0 ou 1) vérifie les trois axiomes des espaces métriques.

2. Même question pour la distance discrète sur un ensemble X quelconque.

3. Même question pour la distance tordue δ définie sur $\mathbb{R}$.

(b) Détails des esquisses de preuves

Regardons, sur un exemple, ce qu'on attend du lecteur lorsqu'il doit rédiger les détails des "esquisses de preuves". Rappelons-la formulation de la première esquisse de preuve :

Recette de preuve.— L'intersection d'un nombre fini d'ouverts est un ouvert : prendre un point dans l'intersection, écrire la définition de l'intersection, puis utiliser que chacun des ensembles est ouvert. (On peut commencer par deux ouverts).

Ecrivons la preuve pour deux ouverts. Il s'agit essentiellement de prendre des notations, puis de suivre la recette.

Preuve.— Soit O_1, O_2 deux ouverts, montrons que $O_1 \cap O_2$ est un ouvert. Soit x un point de $O_1 \cap O_2$. Par définition de l'intersection, x appartient à la fois à O_1 et à O_2. Puisque O_1 est ouvert et contient x, il existe $\varepsilon_1 > 0$ tel que $B(x, \varepsilon_1) \subset O_1$. Puisque O_2 est ouvert et contient x, il existe $\varepsilon_2 > 0$ tel que $B(x, \varepsilon_2) \subset O_2$.

On a en tête le but : trouver un $\varepsilon > 0$ tel que $B(x,\varepsilon) \subset O_1 \cap O_2$.

Soit $\varepsilon = \min(\varepsilon_1, \varepsilon_2)$. C'est un nombre strictement positif, vérifions que la boule $B(x,\varepsilon)$ est incluse dans $O_1 \cap O_2$. Puisque $\varepsilon \leq \varepsilon_1$ et $\varepsilon \leq \varepsilon_2$ on a

$$B(x,\varepsilon) \subset B(x,\varepsilon_1) \subset O_1 \text{ et } B(x,\varepsilon) \subset B(x,\varepsilon_2) \subset O_2$$

et donc $B(x,\varepsilon) \subset O_1 \cap O_2$ comme voulu.

□

Pour un nombre fini d'ouverts ce n'est pas beaucoup plus difficile à condition de ne pas se noyer dans les notations :

Preuve.— Soit $O_1, \dots O_k$ des ouverts, montrons que l'intersection

$$\bigcap_{i=1,\dots,k} O_i$$

est un ouvert. Soit x un point de l'intersection. Soit $i \in \{1, ..., k\}$; par définition de l'intersection, x appartient à O_i ; puisque O_i est ouvert, il existe $\varepsilon_i > 0$ tel que $B(x,\varepsilon_i) \subset O_i$.

Soit

$$\varepsilon = \min_{i=1,\dots,k}(\varepsilon_i).$$

C'est un nombre strictement positif, vérifions que la boule $B(x,\varepsilon)$ est incluse dans l'intersection. Pour chaque i entre 1 et k, on a $\varepsilon \leq \varepsilon_i$ et par conséquent

$$B(x,\varepsilon) \subset B(x,\varepsilon_i) \subset O_i.$$

Finalement

$$B(x,\varepsilon) \subset \bigcap_{i=1,\dots,k} O_i$$

comme voulu.

□

Qu'est-ce qui ne marche pas si on essaie de faire la même preuve pour une intersection infinie d'ouverts ?

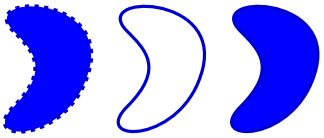
Exercice 10.— Ecrire les détails de la preuve pour une union quelconque d'ouverts.

(c) Qu'est-ce qu'un ouvert ?

Intuitivement, une partie de X est un ouvert si quel que soit le point que l'on choisisse dans cette partie, tous les points assez proches sont aussi dedans ; autrement dit l'ensemble "entoure" chacun de ses points. Lorsqu'on veut dessiner un ouvert, il faut bien spécifier que le bord du dessin ne fait pas partie de l'ensemble considéré, ce qu'on indique par des pointillés. Par exemple le carré sans son bord, $C = [0,1[\times]0,1[$, est un ouvert du plan, mais le carré avec son bord $C' = [0,1] \times [0,1]$ ne l'est pas, parce que les points du bord ne sont pas entouré de points du carré. Le complémentaire de C', par contre, est ouvert, ce qui veut dire que C' est fermé. Il est intéressant de noter qu'il y a plein d'ensemble intermédiaires entre C et C', comme l'ensemble $C \cup \{(1,1)\}$, le carré sans son bord mais avec un coin : ces ensembles donnent des exemples simples qui ne sont ni ouverts ni fermés.

Si le carré fermé C' n'est pas ouvert, c'est que les points du bord du carré, qui font partie de C', ne vérifient pas la définition des ouverts : toutes les boules, aussi petites soient-elles, que l'on peut dessiner autour de ces points, débordent de l'ensemble C'. Il y a donc dans C' deux types de points : ceux qui vérifient la définition des ouverts (ils sont centre d'une boule incluse dans C'), et les autres.

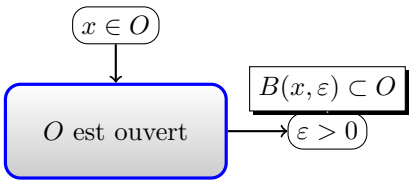
L'**intérieur** de C' est exactement l'ensemble des points de C' qui satisfont à la définition des ouverts.



Intérieur (sans la ligne pointillée), frontière, adhérence

La **frontière** d'une partie, quant à elle, est l'ensemble des points x qui "touchent" à la fois cette partie et son complémentaire ; le sens précis de "toucher" est qu'on doit pouvoir trouver des points aussi proche qu'on veut de x, aussi bien dans la partie que dans son complémentaire. La frontière peut contenir des points qui ne sont pas dans la partie par exemple, la frontière du carré ouvert C coïncide avec ce qu'on appelle couramment le bord du carré, et aucun point du bord n'est dans C. Les ensembles C, C' et $C \cup \{(1,1)\}$ ont tous les trois la même frontière. Intérieur et frontière d'une partie sont toujours disjoints ; leur réunion forme l'**adhérence** de l'ensemble. Finalement, toutes ces notions sont assez simples quand on les appliquent à des ensembles sympathique comme le disque ou le carré (ou la patate dessinée dans la marge). C'est un peu plus compliqué pour les ensembles moins accessibles à l'intuition, comme le sous-ensemble $\mathbb{Q}$ de $\mathbb{R}$ constitué des nombres rationnels. Tout nombre réel peut être approché aussi précisément qu'on le souhaite par un nombre rationnel, mais aussi par un nombre irrationnel : autrement dit, la frontière de $\mathbb{Q}$ est $\mathbb{R}$, ce qui peut paraître un peu suprenant.

Les dessins aident à avoir une intuition des propriétés topologiques des ensembles, mais les preuves doivent utiliser les définitions formelles. La définition d'un ouvert commence par "$\forall x \in O \;\; \exists \varepsilon > 0$". On peut voir cette propriété comme une machine qui produit un nombre ε à chaque fois qu'on lui fournit un élément x de O. On la représente alors par le schéma ci-contre.



Le plus important dans ce schéma, c'est qu'il y une entrée et une sortie, qui ont des rôles très différents. Pour **utiliser** la propriété "O est ouvert", on doit indiquer à quel élément x de O on veut appliquer la définition : la machine ne peut pas fonctionner tant qu'on ne lui a pas fourni un x en entrée ! On reçoit en échange un nombre ε, avec la propriété que $B(x,\varepsilon) \subset O$. Pour **montrer** qu'un ensemble O est ouvert à partir de la définition, on doit "construire une machine", c'est-à-dire expliquer comment, pour chaque élément x de O, on peut trouver un $\varepsilon > 0$ tel que $B(x,\varepsilon) \subset O$. La preuve commencera donc en général par "Soit x un élément de O" , et dira quelque part comment définir ε.

Exercice 11.— Dans la preuve que l'intersection de deux ouverts est un ouvert, le but est de montrer qu'un certain ensemble est ouvert : repérer les phrases "Soit x..." et la phrase qui définit ε. Repérer également l'endroit où on utilise le fait que O_1 est ouvert, et le nombre reçu en échange.

Exercice 12.— **1.** Faire un schéma analogue illustrant la définition de l'intérieur, de l'adhérence, de la continuité en un point x_0, d'une valeur d'adhérence. Dans la suite du cours, illustrer chaque définition par un schéma du même type. **2.** Dans la preuve de la continuité de la composée de deux fonctions continue, faire un schéma pour la propriété "f est continue en x" et un autre pour la propriété "g est continue en $f(x)$". Illustrer la preuve en branchant la sortie de l'une des deux machines (laquelle ?) sur l'entrée de l'autre.

(d) Dualité ouverts/fermés

Dans un espace métrique (ou topologique), les fermés sont exactement les complémentaires des ouverts. L'application $Y \mapsto X \setminus Y$, définie sur l'ensemble des parties de X, est appelée "passage au complémentaire" ; de nombreuses propriétés des fermés peuvent se déduire de celles des ouverts (et réciproquement) en utilisant cette application. Le passage au complémentaire échange en particulier les notions suivantes :

– union et intersection ;
– intérieur et adhérence ;
– être dense, être d'intérieur vide.

Exercice 13.—

1. Donner des énoncés précis correspondant à ces échanges de propriétés par passage au complémentaire, et démontrer-les.

2. Enoncer un critère de continuité des applications portant sur les fermés, puis démontrer-le à partir du critère de continuité portant sur les ouverts.

On verra plus tard d'autres dualités, par exemple la propriété d'être compact peut s'exprimer par une propriété portant sur les ouverts, ou bien par la propriété duale portant sur les fermés.

(e) Notions métriques et topologiques

Les notions de suite convergente, d'intérieur, adhérence, de fonction continue, de valeur d'adhérence d'une suite, peuvent être caractérisées en termes d'ouverts (voir les lemmes ou exercices appelés "caractérisation topologique"). Toute ces notions sont dites "topologiques", elles ne dépendent pas vraiment de la métrique d choisie sur l'ensemble X mais uniquement de la famille des ouverts associée à cette métrique. Ainsi, la proposition disant que la famille des ouverts de l'espace produit $X \times Y$ est la même pour les trois distances d_1, d_2, d_∞ est très important : toutes les notions topologiques vont être les mêmes pour ces trois distances. On peut parler par exemple de la topologie du carré $[0,1]^2$ sans préciser laquelle des trois distances on a choisi.

Un *homéomorphisme* est une application $\Phi : X \to Y$ entre deux espaces métriques (ou juste topologiques) X et Y, qui est bijective, continue, et d'inverse continue. En utilisant le critère topologique de continuité, on voit que la continuité de Φ et de Φ^{-1} revient à dire qu'une partie O de X est ouverte si et seulement si $\Phi(O)$ est ouverte. Lorsqu'il existe un homéomorphisme entre X et Y, on dit que ces deux espaces sont *homéomorphes*. Voici un exemple : l'application

Dessiner le graphe de cette application.

$$\tan : \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[\to \mathbb{R}$$

est un homéomorphisme (dont l'inverse est l'application arctangente). Un exemple un peu plus sophistiqué : soit S la sphère unité dans $\mathbb{R}^3$, et C le bord du cube $[-1,1]^3$ (qui contient la sphère S). Soit f l'application qui "projette S sur C depuis l'origine" : autrement dit, pour chaque point x de S, on trace la demi-droite issue de l'origine et passant par x, cette demi-droite rencontre le bord du cube en un unique point qui est précisément $f(x)$. Alors l'application f est un

homéomorphisme entre S et C. Intuitivement, deux espaces sont homéomorphes lorsqu'on peut déformer le premier jusqu'à ce qu'il prenne la forme du deuxième, sans rien déchirer, ni écraser, ni recoller. La notion d'homéomorphisme est fondamentale en topologie : un topologue considère que deux espaces métriques qui sont homéomorphes ne sont pas vraiment différents. En particulier, toutes les propriétés topologiques de l'un se reflètent alors dans l'autre comme dans un miroir ; cette propriété est formalisée par l'exercice suivant.

Faire un dessin représentant la sphère S, le cube C et l'application f.

Exercice 14.— Soit $\Phi : X \to Y$ un homéomorphisme. Vérifier que :

1. $\Phi(O)$ est ouvert si et seulement si O est ouvert,

2. $\Phi(F)$ est fermé si et seulement si F est fermé,

3. l'intérieur de l'image par Φ d'un ensemble E est égal à l'image de l'intérieur de E,

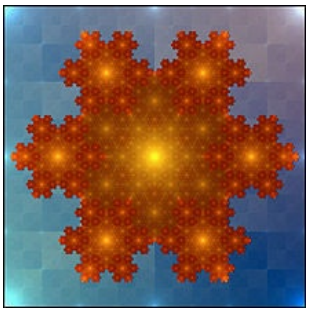
4. l'adhérence de l'image est égale à l'image de l'adhérence,

5. l'image d'une suite convergent vers x est une suite convergent vers $\Phi(x)$,

6. X est compact si et seulement si Y est compact,

7. X est connexe si et seulement si Y est connexe,

8. ...



Le flocon de neige de Von Koch est un sous-espace métrique du plan qui est homéomorphe à un disque (source Wikipedia).

Une *isométrie* est une application bijective $\Phi : X \to Y$ entre deux espaces métriques, qui est compatible avec les distances : la distance entre deux points de X est égale à la distance entre leurs images dans Y. En symboles :

$$\forall x_1, x_2 \in X, \quad d_X(x_1, x_2) = d_Y((\Phi(x_1), \Phi(x_2)).$$

Lorsqu'il existe une telle isométrie, on dit que X et Y sont *isométriques*. Toutes leurs propriété métriques sont alors identiques ; c'est en particulier le cas de la complétude, qui est une notion métrique et non pas topologique (voir plus loin).

Par exemple, les carrés $[0,1] \times [0,1]$ et $[3,4] \times [5,6]$, munis tous deux de la distance induite par la distance euclidienne du plan, sont isométriques. Ces carrés ne sont pas isométriques au rectangle $[10,20] \times [0,1]$; par contre, ils lui sont homéomorphes.

La situation est tout à fait analogue à l'algèbre linéaire, où deux espaces vectoriels sont dits *isomorphes* s'il existe une bijection linéaire entre les deux ; les propriétés des deux espaces vectoriels sont alors identiques ; en particulier, ils ont la même dimension.

Le fait que deux distances d et d' définies sur un ensemble X sont topologiquement équivalentes revient à dire que l'application identité, de (X, d) vers (X, d'), est un homéomorphisme. Toutes les notions topologiques définies à l'aide de d ou de d' seront les mêmes.

Exercice 15.—

1. Construire une isométrie du plan qui envoie le carré $[0,1] \times [0,1]$ sur le carré $[3,4] \times [5,6]$. En déduire que les espaces métriques $([0,1] \times [0,1], d_2)$ et $([3,4] \times [5,6], d_2)$ sont isométriques.

2. Construire un homéomorphisme entre $[0,1] \times [0,1]$ et $[10,20] \times [0,1]$. Démontrer qu'il n'existe pas d'isométrie entre ces deux sous-espaces métriques.

Topologie des surfaces Ce texte décrit la topologie en tant qu'ensemble de concepts fondamentaux pouvant servir d'outils dans toutes les mathématiques. Mais la topopologie elle-même est aussi une branche vivante de la recherche en mathématiques. Le travail du topologue consiste, en gros, à "décrire les espaces topologiques à homéomorphisme près". Dans ce paragraphe, nous allons expliquer un résultat profond de topologie dont la preuve date des années 1920.

Pour des aperçus plus récents, on peut regarder ces deux articles du site *Images des mathématiques* : Les tresses, de la topologie à la cryptographie et Géométriser l'espace, de Gauss à Perelman.

En réalité les topologues ne peuvent pas vraiment classifier tous les espaces topologiques, il y en a beaucoup trop ! Pour espérer démontrer quelquechose d'intéressant, il faut se restreindre à une famille raisonnable d'espaces. Examinons l'exemple des surfaces.

Qu'est-ce qu'une surface ? Intuitivement, c'est un espace qui, près de n'importe lequel de ses point, ressemble au plan. Précisons cette idée en disant qu'une *surface* est un espace métrique ayant la propriété suivante : pour tout point x, il existe un ouvert O qui contient x et qui est homéomorphe au plan $\mathbb{R}^2$. Voici maintenant l'énoncé du théorème de classification des surfaces :

Théorème. *Toute S surface topologique compacte, connexe, orientable est homéomorphe à l'un des modèle suivants : la sphère, le tore, le tore à deux trous, ou plus généralement la "surface de genre g" qui est un tore à g trous (dessinées sur la page Wikipedia genre (mathématiques)).*

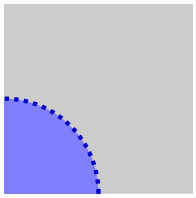
Il nous reste quelques termes à définir pour que l'énoncé soit parfaitement compréhensible. Une surface est dite *orientable* si elle ne contient pas de sous-espace homéomorphe au ruban de Möbius. La compacité et la connexité sont définis dans les chapitres correspondant un peu plus loin.

Une *variété topologique de dimension N* est un espace métrique ayant la propriété suivante : pour tout point x, il existe un ouvert O qui contient x et qui est homéomorphe à $\mathbb{R}^N$. Les surfaces sont donc des variétés de dimension deux. Les intervalles ouverts de $\mathbb{R}$ sont des variétés de dimension un, et ils sont tous homéomorphes.

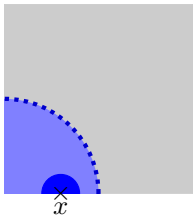
Exercice 16.— Donner un exemple de variété compacte, connexe de dimension un. (On peut en fait montrer que toutes les variétés compactes, connexes de dimension un sont homéomorphes).

(f) Sous-espaces métriques

Dans le plan $\mathbb{R}^2$ muni de la distance habituelle notée d, on considère un carré C. Nous aimerions comprendre le sous-espace métrique (C, d). Pour ça, il suffit de suivre une seule consigne : *oublier complètement le reste du plan.*



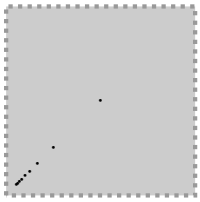
Le dessin ci-contre montre une boule (bleue) dans cet espace métrique (gris). Vous trouvez peut-être que cette boule a une allure un peu louche, mais elle représente bien tous les points du carré à distance strictement plus petite que R du centre (si vous pensez que ce qu'on a dessiné n'est pas la boule toute entière, relisez la consigne !)



Cette boule est-elle un ouvert dans l'espace métrique (C, d) ? Bien sûr que oui : notre carré est un espace métrique comme les autres, tous les résultats du chapitre s'appliquent, et en particulier l'exercice 1 qui dit que toute boule ouverte est un

ouvert. Pourtant elle contient des points du bord du carré, qui ne semblent pas être "entourés" par des points de la boule... Vérifions la définition : en un point x de la boule qui est sur le bord du carré, nous pouvons dessiner une petite boule (bleue foncé) de (C, d) centré en ce point et incluse dans la grande boule. Ces points satisfont donc la définition des ouverts de (C, d).

Considérons maintenant un carré ouvert, par exemple $C' =]0, 2[^2$. La suite $(\frac{1}{n}, \frac{1}{n})_{n\geq 1}$ est-elle convergente dans l'espace métrique (C', d) ?



Réponse p28

(g) Espaces produits, espaces de suites

Le plan $\mathbb{R}^2$ peut être vu comme le produit $\mathbb{R} \times \mathbb{R}$. Chacun des facteurs $\mathbb{R}$ étant muni de la distance usuelle (valeur absolue), nous avons défini trois distances produits d_1, d_2, d_∞ dans le paragraphe sur les espaces produits. On dispose ainsi des trois espaces métriques $(\mathbb{R}^2, d_1)$, $(\mathbb{R}^2, d_2)$, $(\mathbb{R}^2, d_\infty)$. Nous avons vu que ces trois distances sont équivalentes, en particulier les notions d'ouverts, de suites convergente, *etc.* sont les mêmes. Ces trois distances se généralisent sans difficulté à $\mathbb{R}^N$, vu comme le produit de N copies de $\mathbb{R}$.

Une boule ouverte dans $\mathbb{R}^2$ pour chacune des trois distances produits d_1, d_2 et d_∞. Quel dessin va avec quelle distance ?...

Nous voudrions maintenant essayer de faire tendre N vers $+\infty$, c'est-à-dire de définir le **produit d'une infinité (dénombrable) de droites**. Un x élément de $\mathbb{R}^N$ est caractérisé par ses N coordonnées $(x_1, \dots, x_N)$. De même, les éléments de notre produit infini seront caractérisés par une infinité de coordonnées $(x_1, \dots, x_N, \dots)$: autrement dit, ce produit infini n'est rien d'autre que l'ensemble des suites de nombres réels. A cause de ce point de vue "produit infini", cet ensemble des suites réelles est souvent noté $\mathbb{R}^{\mathbb{N}}$. Il nous faut maintenant définir une distance sur cet ensemble. On peut généraliser à $\mathbb{R}^{\mathbb{N}}$ les formules définissant d_1, d_2 et d_∞, mais certains éléments sont à une distance infinie les uns des autres. Prenons le cas de la distance d_∞. On définit

$$d_\infty((x_1, x_2, \dots), (x'_1, x'_2, \dots)) = \sup\{|x_1 - x'_1|, |x_2 - x'_2|, \dots\}$$

mais on a par exemple

$$d_\infty((1, 2, 3, \dots), (0, 0, 0, \dots)) = \sup\{1, 2, 3, \dots\} = +\infty.$$

Cette formules ne définit donc pas une distance sur $\mathbb{R}^{\mathbb{N}}$. Pour contourner ce problème, on élimine les éléments qui sont à distance infinie de l'élément $(0, 0, 0, \dots)$. Autrement dit, on se restreint aux suites bornées. L'ensemble des suites réelles bornées est noté $\ell_\infty(\mathbb{R})$, et le couple $(\ell_\infty(\mathbb{R}), d_\infty)$ est bien un espace métrique.

La même démarche fonctionne avec d_1 : on pose

$$d_1((x_1, x_2, \dots), (x'_1, x'_2, \dots)) = |x_1 - x'_1| + |x_2 - x'_2| + \cdots$$

et on définit l'ensemble

$$\ell^1(\mathbb{N}, \mathbb{R}) = \{(x_1, x_2, \cdots), d_1((x_1, x_2, \dots), (0, 0, \dots)) < +\infty\}$$

qui est en fait l'ensemble des termes généraux des séries absolument convergentes. Le couple $(\ell_1(\mathbb{R}), d_1)$ est alors un espace métrique.

Comme le terme général d'une série convergente est une suite bornée, l'ensemble ℓ_1 est inclus dans l'ensemble ℓ_∞. Autrement dit, la distance d_∞ a aussi un sens sur $\ell_1(\mathbb{R})$, et on peut se demander si les distances d_1 et d_∞ sont équivalentes sur cet ensemble. Contrairement au cas des produits finis, la réponse est négative, et c'est l'une des grosse difficulté des espaces de suites (ou de fonctions) : ils admettent de nombreuses distances différentes (et intéressantes), qui ne sont pas équivalentes ni même topologiquement équivalentes : une suite d'éléments peut converger lorsqu'on considère une certaine distance, et ne plus converger lorsqu'on en considère une autre. Nous verrons des exemples concrets de ce phénomène plus loin, dans la section sur les espaces vectoriels normés de fonctions.

Exercice 17.—

1. Montrer que $(\ell_\infty(\mathbb{R}), d_\infty)$ est un espace métrique.

2. Montrer que $(\ell_1(\mathbb{R}), d_1)$ est un espace métrique.

3. Définir l'espace métrique $(\ell_2(\mathbb{R}), d_2)$.

(h) Espaces de fonctions

***** ex : dans $C_b([-1,1],\mathbb{R})$, $B(0,1)$, $B(x \mapsto x^2, 1)$; dans $\mathcal{B}([-1,1],\mathbb{R})$, $B(\mathbf{1}_{[0,1]}, \varepsilon)$. Cette boule ne contient manifestement aucune fonction continue. Ceci illustre le théorème sur les limites uniformes de suites de fonctions continues. ******

*** CVS : déf, exemple de la bosse voyageuse. Topologie produit. ****

I.3 Exercices

Exercice 18.— Soient $x_1, ..., x_{10}$ dix points dans un espace métrique X. Montrer que

$$d(x_1, x_{10}) \leq d(x_1, x_2) + d(x_2, x_3) + \cdots + d(x_9, x_{10}).$$

Combien de fois avez-vous appliqué l'inégalité triangulaire ?

Exercice 19.— Soit (X, d) un espace métrique. On définit son diamètre

$$\operatorname{diam}(X, d) = \sup_{x,x' \in X} \{d(x, x')\}$$

qui est un élément de $[0, +\infty]$. On dit que (X, d) est borné si son diamètre est fini.

Montrer que les trois propriétés suivantes sont équivalentes :

1. (X, d) est borné,
2. $\forall x_0 \in X \quad \exists M > 0 \quad \forall x \in X \quad d(x, x_0) < M$,
3. $\exists x_0 \in X \quad \exists M > 0 \quad \forall x \in X \quad d(x, x_0) < M$.

Exercice 20.— Enoncer précisément et démontrer l'unicité de la limite d'une suite dans un espace métrique.

Exercice 21.— Montrer que les sous-ensembles suivants de $\mathbb{R}^2$ sont ouverts. Déterminer leur adhérence.

$$\{(x, y) \mid x > 0\}$$

$$\{(x, y) \mid x > 0 \text{ et } y < 0\}$$

$$\{(x, y) \mid x + 2y > 0 \text{ et } y^2 > x\}$$

Exercice 22.— Déterminer (sans preuve, mais en faisant des dessins) l'intérieur, l'adhérence, la frontière des parties suivantes du plan : (a) le disque unité fermé ; (b) la droite $\{y = 0\}$; (c) la réunion des deux précédents.

Exercice 23.— Montrer que le sous-espace vectoriel de $\mathbb{R}^3$ défini par les équations $x + y = 0$, $y + z = 0$ est fermé dans $\mathbb{R}^3$. Montrer plus généralement que tout ensemble dans $\mathbb{R}^n$ de solutions d'un système d'équations linéaires est fermé.

Exercice 24.— Trouver un ouvert O de $\mathbb{R}^2$, différent de $\mathbb{R}^2$, tel que $\mathrm{Inte}(\mathrm{Adhe}(O)) = \mathbb{R}^2$. Trouver d'autres ouverts qui ne sont pas l'intérieur de leur adhérence.

Exercice 25.—(utile) Montrer que les ouverts de (X, d) sont les réunions de boules ouvertes. *Aide. Il y a deux choses à montrer : (1) toute réunion de boules ouvertes est un ouvert, (2) tout ouvert est égal à une réunion de boule ouverte. Le premier découle immédiatement de deux propriétés du cours. Pour le second, étant donné un ouvert O, considérer l'ensemble O' défini comme la réunion de toutes les boules ouvertes incluses dans O, et montrer que $O' = O$.*

Exercice 26.— Que peut-on dire d'une partie O de $\mathbb{R}^2$ qui vérifierait "la définition des ouverts dans laquelle on intervertit les quantificateurs", à savoir :

$$\exists \varepsilon > 0 \;\; \forall x \in O \;\; B(x, \varepsilon) \subset O?$$

Exercice 27.— Montrer que si Y est une partie fermée bornée non vide de $\mathbb{R}$, alors il existe deux éléments a, b **dans** Y tels que $Y \subset [a, b]$.

Exercice 28.— Montrer que la frontière d'une partie E de X est l'intersection de l'adhérence de E et de l'adhérence du complémentaire de E. En déduire que la frontière d'un ensemble coïncide avec la frontière de son complémentaire.

Pour tout ensemble E, montrer que les trois ensembles

$$\mathrm{Inte}(E), \mathrm{Fr}(E), \mathrm{Inte}(X \setminus E)$$

forment une partition de X.

Montrer que l'adhérence est la réunion de l'intérieur et de la frontière.

Exercice 29.— (utile) Montrer la caractérisation suivante de l'adhérence : un point x est dans l'adhérence de E si et seulement si il existe une suite d'éléments de E convergent vers x. (En particulier, E est dense dans X si et seulement si tout élément de X est limite d'une suite d'éléments de E).

Exercice 30.— (utile) Montrer qu'une partie A de (X, d) est dense dans X si et seulement si A rencontre tout ouvert de X.

Exercice 31.— (utile) Donner une caractérisation topologique (c'est-à-dire n'utilisant pas la distance mais seulement les ouverts) d'une suite convergente.

Exercice 32.— Enoncer et démontrer une caractérisation séquentielle des ouverts.

Exercice 33.— En utilisant la continuité de l'application distance (section (f)) et la caractérisation topologique de la continuité, donner un nouvel argument pour montrer que toute boule ouverte est un ouvert.

Exercice 34.—(**caractérisation topologique de la continuité en un point**) Dans un espace métrique, un ensemble V est un *voisinage* d'un point x_0 si x_0 appartient à l'intérieur de V, autrement dit si V contient une boule ouverte centrée en x_0. Montrer que $f : X \to Y$ est continue au point x_0 si et seulement si pour tout voisinage W de $f(x_0)$, l'ensemble $V := f^{-1}(W)$ est un voisinage de x_0.

Exercice 35.—(**caractérisation séquentielle de la continuité**) (utile)

1. Soit $f : X \to Y$, et x un point de X. Montrer le résultat suivante : *l'application f est continue au point x_0 si et seulement si pour toute suite (x_n) de X convergent vers x, la suite $(f(x_n))$ converge vers $f(x)$.*

2. En déduire une nouvelle preuve du fait que la composée de deux fonctions continues est continue.

Exercice 36.— (utile) Montrer qu'une application $f : X \to Y$ est continue si et seulement si l'image réciproque de tout ensemble fermé est fermé.

Exercice 37.—

1. Montrer que l'ensemble des valeurs d'adhérence de $(x_n)_{n\in\mathbb{N}}$ est égal à l'ensemble

$$A := \bigcap_{N\geq 0} \mathrm{Adhe}(\{x_N, x_{N+1}, \ldots\}).$$

Aide : partir de la définition des valeurs d'adhérence, intervertir deux quantificateurs pour faire apparaître la définition de l'adhérence de l'ensemble $\{x_N, x_{N+1}, \ldots\}$. En déduire que cet ensemble est fermé.

2. Déterminer l'adhérence de l'ensemble $\{\frac{1}{n}, n > 0\}$ des inverses des entiers strictement positifs.

3. Tester la formule de la première question sur la suite $(1/n)_{n>0}$. *Aide : il s'agit de décortiquer la formule ; déterminer d'abord ce que vaut* $\mathrm{Adhe}(\{x_0, x_1, \ldots\})$*, puis* $\mathrm{Adhe}(\{x_1, x_2, \ldots\})$*, et en général* $\mathrm{Adhe}(\{x_N, x_{N+1}, \ldots\})$ *; déterminer ensuite l'intersection de ces ensembles.*

Exercice 38.— Dans le sous-espace métrique $(]0,1]^2, d_2)$, on considère la suite $((1/n, 1/n))_{n>0}$. Est-elle convergente ?

Exercice 39.— (utile) Soit X, d un espace métrique, et Y une partie de X.

1. Montrer que les ouverts du sous-espace métrique (Y, d) sont exactement les intersections de Y avec un ouvert de X.

2. En déduire que si Y est ouvert dans X, alors une partie de Y est ouverte dans Y si et seulement si elle est ouverte dans X.

3. Formuler et démontrer des propriétés analogues pour les fermés.

Exercice 40.— (utile) Montrer que les relations "être isométrique" et "être homéomorphes" sont des relations d'équivalence : autrement dit, par exemple pour la relation d'isométrie, pour tous espaces métriques X, Y, Z, on a

- X est isométrique à lui-même ;
- si X est isométrique à Y alors Y est isométrique à X ;
- si X est isométrique à Y et Y est isométrique à Z alors X est isométrique à Z.

Montrer que toute isométrie est un homéomorphisme.

Exercice 41.— Donner un exemple, dans $\mathbb{R}^2$, d'une famille de parties ouvertes dont l'intersection n'est pas ouvertes. Même question pour une union de parties fermés.

Exercice 42.—

1. Dans l'espace métrique $X = C([-1,1], \mathbb{R})$ muni de la distance d_∞, décrire la boule ayant pour centre la fonction nulle et pour rayon 1. Plus généralement, décrire la boule de rayon ε et de contre f_0 pour un élément f_0 quelconque de X (chercher une description utilisant le graphe de la fonction f_0).

2. Soit f_0 la fonction définie sur $[-1,1]$ par $f(x) = 0$ si $x \leq 0$ et $f(x) = 1$ si $x > 0$. Déterminer explicitement un $\varepsilon > 0$ tel que la boule de rayon ε et de centre f_0 ne contient aucune fonction continue.

Exercice 43.— Dans l'espace métrique $X = C_b([0,1], \mathbb{R})$ on considère la partie

$$X_0 = \{f \in C_b([0,1], \mathbb{R}) \ \mid f(0) = 0\}.$$

1. Montrer que X_0 est une partie fermée.

2. Montrer que tout élément de X_0 est limite d'une suite d'éléments de son complémentaire.

3. En déduire qu'il est d'intérieur vide.

4. Montrer de même que l'ensemble $X_{\geq 0}$ des fonctions positives est fermé. (**) Déterminer son intérieur.

Exercice 44.— Soit X l'ensemble des parties mesurables du carré $[0,1]^2$, et m la mesure de Lebesgue. Pour deux éléments A, B de X on pose

$$d(A,B) = m(A\Delta B)$$

où $A\Delta B = (A \setminus B) \cup (B \setminus A)$ est la *différence symétrique* de A et B.

1. Représenter la différence symétrique sur un dessin.

2. Montrer l'égalité $A\Delta C = (A\Delta B)\Delta(B\Delta C)$. En déduire que d vérifie l'inégalité triangulaire.

3. Lequel des axiomes de distance n'est pas vérifié par d ?

(l'application d induit en réalité une distance sur l'ensemble des parties mesurables modulo ensembles de mesure nulle).

Espaces de matrices Soit n un entier positif. L'ensemble $\mathcal{M}_n(\mathbb{R})$ est un espace vectoriel réel de dimension n^2, il s'identifie à $\mathbb{R}^{n^2}$, ce qui en fait un espace métrique lorsqu'on munit $\mathbb{R}^{n^2}$ de l'une des métriques habituelles. Par exemple, l'application

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto (a,b,c,d)$$

identifie $\mathcal{M}_2(\mathbb{R})$ et $\mathbb{R}^4$, et la métrique d_1 s'écrit

$$d_1(M,M') = \max\{|a-a'|, |b-b'|, |c-c'|, |d-d'|\}.$$

Exercice 45.— On rappelle que $GL_n(\mathbb{R})$ désigne le sous-ensemble de $\mathcal{M}_n(\mathbb{R})$ formé des matrices inversibles. On rappelle aussi que l'application déterminant $\det : \mathcal{M}_n(\mathbb{R}) \to \mathbb{R}$ est un polynôme en les coefficients de la matrice, par exemple $\det(M) = ad - bc$ sur $\mathcal{M}_2(\mathbb{R})$; d'autre part une matrice est inversile si et seulement si son déterminant n'est pas nul.

1. Montrer que $GL_n(\mathbb{R})$ est un ouvert de $\mathcal{M}_n(\mathbb{R})$.

2. Soit M_0 une matrice particulière. Rappeler pourquoi il n'existe qu'un nombre fini de valeur de $t \in \mathbb{R}$ qui annule l'expression $\det(M_0 + t\mathrm{Id})$. En déduire qu'on peut trouver une suite (t_n) de réels tendant vers 0 tels que, pour tout n, la matrice $M_0 + t_n\mathrm{Id}$ est inversible. Comment s'interprète ce résultat, en termes de propriété topologique de la partie $GL_n(\mathbb{R})$?

3. Une application. On veut montrer l'identité $\mathrm{Tr}(MN) = \mathrm{Tr}(NM)$ concernant la trace (somme des termes diagonaux) d'une matrice.

a. Il est assez facile de voir que la trace d'une matrice est invariante par conjugaison : autrement dit, pour tout $M \in \mathcal{M}_n(\mathbb{R})$ est tout $P \in GL_n(\mathbb{R})$, on a $\mathrm{Tr}(PMP^{-1}) = \mathrm{Tr}(M)$ (ceci vient par exemple du fait que le polynôme caractéristique d'une matrice est invariant par conjugaison ; mais admettons ce fait). En déduire que l'identité recherchée est vraie lorsque la matrice M est inversible.

b. Utiliser la question 2 pour en déduire que l'identité est encore vraie lorsque N n'est pas inversible.

Exercice 46.— **1.** Soit $SL_n(\mathbb{R})$ l'ensemble des matrices de déterminant 1. Montrer que c'est un ensemble fermé de $\mathcal{M}_n(\mathbb{R})$.

2. Même question pour l'ensemble $SO_n(\mathbb{R})$ des matrices orthogonales, c'est-à-dire telles que $MM^t = \mathrm{Id}$.

Exercice 47.— *(partiel 2013)* Soient (X, d) un espace métrique, et A, B deux parties de X. On suppose que

$$\forall \varepsilon > 0 \quad \exists a \in A \quad \exists b \in B \quad d(a, b) < \varepsilon.$$

Montrer qu'il existe une suite $(a_n)_{n\in\mathbb{N}}$ dans A et une suite $(b_n)_{n\in\mathbb{N}}$ dans B telles que la suite $(d(a_n, b_n))_{n\in\mathbb{N}}$ tend vers 0.

Beaucoup d'étudiants ont donné la réponse suivante :

Soient $(a_n)_{n\in\mathbb{N}}$ une suite de points de A convergeant vers a, et $(b_n)_{n\in\mathbb{N}}$ une suite de points de B convergeant vers b : on a

$$\forall \varepsilon > 0 \;\; \exists n_0 \in \mathbb{N} \;\; \forall n \geq n_0 \;\; d(a_n, a) < \varepsilon,$$

$$\forall \varepsilon > 0 \;\; \exists n_1 \in \mathbb{N} \;\; \forall n \geq n_1 \;\; d(b_n, b) < \varepsilon.$$

On pose $N = \max\{n_0, n_1\}$. Pour tout $n \geq N$, on a

$$d(a_n, b_n) \leq d(a_n, a) + d(a, b) + d(b, b_n) \leq 3\varepsilon.$$

Donc

$$\forall \varepsilon > 0 \;\; \exists N \in \mathbb{N} \;\; \forall n \geq N \;\; d(a_n, b_n) < \varepsilon,$$

ce qui signifie que la suite $d(a_n, b_n)_{n\in\mathbb{N}}$ tend vers 0.

Cette réponse a immédiatement l'air louche : si on peut vraiment prendre n'importe quelle suite (a_n) tendant vers a, pourquoi ne pas prendre la suite constante égale à a ? Et pareil pour la suite (b_n) ? Et finalement, si la première suite tend vers a et la seconde vers b, la seule possibilité pour que la distance entre a_n et b_n tende vers 0 n'est-elle pas qu'on ait $a = b$? Mais comment pourrait-on conclure que $a = b$ si on sait seulement que $d(a, b) < \varepsilon$? ? Examinons maintenant dans le détail les arguments de cette réponse, en essayant de combler au mieux les lacunes de la rédaction. Dès la première ligne, on nous parle d'un point a et d'un point b sans préciser d'où ils viennent ; cependant, on se doute qu'il s'agit des points a et b fournis par l'hypothèse. Si on lit bien l'hypothèse, ces points dépendent de la donnée d'un nombre ε. Il faut donc probablement comprendre ceci :
Soit $\varepsilon > 0$, et $a \in A$ et $b \in B$ deux points donnés par l'hypothèse, vérifiant $d(a, b) < \varepsilon$. *Soient $(a_n)_{n\in\mathbb{N}}$ une suite de points de A convergeant vers a, et $(b_n)_{n\in\mathbb{N}}$ une suite de points de B convergeant vers b...* Jusque là tout va bien : la réponse n'était pas correctement rédigée, mais on peut lui donner un sens en ajoutant la phrase en gras (à vrai dire, c'est même la seule façon de préciser ce qui est écrit). La suite s'interprète alors probablement ainsi :
En appliquant la définition de la limite pour les suites (a_n) et (b_n) au nombre ε, on obtient deux nombres n_0 et n_1 tels que $\forall n \geq n_0 \;\; d(a_n, a) < \varepsilon$ et $\forall n \geq n_0 \;\; d(b_n, b) < \varepsilon$. On pose $N = \max\{n_0, n_1\}$. Pour tout $n \geq N$, on a

$$d(a_n, b_n) \leq d(a_n, a) + d(a, b) + d(b, b_n) \leq 3\varepsilon.$$

Arrêtons-nous pour résumer la situation, en faisant bien attention à l'ordre dans lequel les différents objets mathématiques ont été introduits : **étant donné un $\varepsilon > 0$, nous avons construit deux suites (a_n) et (b_n) pour lesquelles il existe un nombre N tel que pour tout $n \geq N$, $d(a_n, b_n) \leq 3\varepsilon$.**C'est maintenant que les choses se gâtent. On voudrait conclure que la suite $(d(a_n, b_n))$

tend vers 0, mais pour cela il faudrait que nos deux suites aient été construites *avant* le choix du nombre ε. Dans la définition de la limite, on considère une suite ; étant donnée cette suite, à chaque fois qu'on choisit un $\varepsilon > 0$ on doit pouvoir trouver un rang N *etc.*. Ici la situation est inversée : on a d'abord choisi un nombre $\varepsilon > 0$, puis une suite. Pour pouvoir conclure, le nombre ε étant fixé depuis le début, il faudrait considérer un autre nombre quelconque $\varepsilon' > 0$ et trouver alors un N tel que *etc.*.

Mais après tout, peut-être est-ce de cette façon qu'il fallait comprendre la réponse ; puisque le premier ε n'y était pas explicité, peut-être fallait-il comprendre le second ε comme un autre nombre ε'. Essayons :

Soit $\varepsilon' > 0$. En appliquant la définition de la limite pour les suites (a_n) et (b_n) au nombre ε', on obtient deux nombres n_0 et n_1 tels que $\forall n \geq n_0 \quad d(a_n, a) < \varepsilon'$ et $\forall n \geq n_0 \quad d(b_n, b) < \varepsilon'$. On pose $N = \max\{n_0, n_1\}$. Pour tout $n \geq N$, on a

$$d(a_n, b_n) \leq d(a_n, a) + d(a, b) + d(b, b_n) \leq 2\varepsilon' + \varepsilon.$$

Ce n'est pas encore ce qu'on voudrait. On peut bien sûr éliminer le facteur 2 et obtenir l'inégalité $d(a_n, b_n) \leq \varepsilon' + \varepsilon$, mais on ne pourra jamais se débarrasser du ε : pour choisir a et b on a été obligé de fixer ce nombre ε, et on pourra jamais faire mieux que de montrer que la distance de a_n à b_n tend vers le nombre $d(a, b)$, qui est strictement plus petit que ε, mais dont rien ne permet de dire qu'il est nul.

Cette réponse est donc irrécupérable : même un lecteur plein de bonne volonté ne pourra pas y voir un raisonnement correct.

Le problème dans cet argument vient de la "dépendance cachée" entre les différents objets introduits : ici, les suites dépendant du choix antérieur d'un nombre ε. Ce problème est à l'origine de nombreuses erreurs. Comment faire pour les éviter ?

1. *Introduire explicitement tous les objets utilisés* (ici, les points a et b, et plus loin l'entier N, ne sont pas correctement introduits dans la solution proposée initialement par l'étudiant ; une introduction explicite est donnée par la phrase en caractères gras que nous avons ajoutée). Rappelons que dans une phrase avec des quantificateurs, comme

   $$\forall \varepsilon > 0 \quad \exists a \in A \quad \exists b \in B \quad d(a, b) < \varepsilon$$

   aucun objet n'est introduit. Les variables de cette phrases sont muettes. Si on veut utiliser cette phrase pour introduire de nouveaux objets, il faut :

   (a) dire à quel nombre ε on va l'appliquer : ça peut être un nombre ε introduit avant dans le raisonnement (comme $\varepsilon = 1/n$ dans la solution plus bas), ou bien un nombre ε quelconque, qu'on introduit pour l'occasion en écrivant "Soit $\varepsilon > 0$" ;

   (b) dire comment on note les objets obtenus : on utilise parfois les mêmes notations, parfois non (comme a_n, b_n dans la solution plus bas).

2. Lorsque tous les objets ont été introduits explicitement, on peut alors faire attention à *l'ordre dans lequels les objets sont introduits* ; cet ordre entraîne une dépendance entre les objets. On peut même expliciter cette dépendance en écrivant par exemple

Soit $\varepsilon > 0$, et $a_\varepsilon \in A$ et $b_\varepsilon \in B$ deux points donnés par l'hypothèse... Et, plus loin, *Soient* $(a_{n,\varepsilon,a})_{n\in\mathbb{N}}$....

Malheureusement, comme on le voit, les notations deviennent vite très lourdes, ce qui explique pourquoi, en pratique, on n'explicite pas toutes les dépendances dans les notations. Il faut donc faire des bilans où l'ordre des choix apparait explicitement : dans la tentative de solution plus haut on a introduit, dans l'ordre, $\varepsilon, a, b, a_n, b_n, N$; a priori chacun de ces objets dépend du choix des précédents.

Solution de l'exercice Soit $n > 0$ un entier, on applique l'hypothèse au nombre $\varepsilon = 1/n$. On obtient ainsi un élément a_n de A et un élément b_n de b tels que $d(a_n, b_n) < 1/n$. On a alors construit une suite $(a_n)_{n\in\mathbb{N}}$ dans A et une suite $(b_n)_{n\in\mathbb{N}}$ dans B telles que, pour tout n, $d(a_n, b_n) < 1/n$. On en déduit que la suite $(d(a_n, b_n))_{n\in\mathbb{N}}$ tend vers 0.

Ce principe consistant à construire une suite à partir d'une propriété en "$\forall\varepsilon > 0\ \exists$..." est très souvent utilisé ; dans le cours, il a notamment servi dans la preuve de la caractérisation séquentielle des fermés.

I.4 Aides et Réponses

Quelle est la démarche pour montrer que l'intérieur de E est un ouvert ? D'après la définition d'un ouvert, on considère un point x de l'intérieur de E, et on cherche $\varepsilon > 0$ tel que $B(x, \varepsilon) \subset \mathrm{Inte}(E)$.

Quelle est la démarche pour montrer que la composée des deux applications continues $f : X \to Y$ et $g : Y \to Z$ est continue ? On veut montrer que l'application est continue en tout point de X. Il s'agit donc de prendre un $x_0 \in X$ et un $\varepsilon > 0$, et de trouver un $\delta > 0$ vérifiant

$$\forall x \in X, \quad (d_X(x, x_0) < \delta \Rightarrow d_Y(g(f(x)), g(f(x_0))) < \varepsilon).$$

Non ? Si.

Qu'est-ce qui ne marche pas si on essaie de faire la même preuve pour une intersection infinie d'ouverts ? On a une infinité de nombre $\varepsilon_i > 0$, on peut encore définir $\varepsilon = \inf(\varepsilon_i)$, mais ce nombre peut être nul.

En utilisant la caractérisation séquentielle des fermés, donner une autre formulation du théorème sur la continuité de la limite d'une suite de fonctions convergent uniformément. Une formulation équivalente est : $C_b(X, Y)$ est une partie fermée de $\mathcal{B}(X, Y)$.

On veut montrer que f est continue ; on fixe un x_0, et aussi un $\varepsilon > 0$, que cherche-t-on alors ? D'après la définition de la continuité, on cherche un $\delta > 0$ tel que pour tout $x \in X$, $d_Y(f(x), f(x_0)) < \delta$.

Dans quel ordre a-t-on introduit les variables $\delta, \varepsilon, N, x, x_0$ de cette preuve ? L'ordre est $x_0, \varepsilon, N, \delta, x$.

La suite $(\frac{1}{n}, \frac{1}{n})_{n \geq 1}$ est-elle convergente dans l'espace métrique $(]0, 2[^2, d)$? Non : il n'y a aucun point de $]0, 2[^2$ qui satisfait la définition de limite (voir aussi le critère de convergence).

II Complétude

II.1 Théorie

(a) Définitions, résultats

Soit (X, d) un espace métrique. Une suite (x_n) d'éléments de X est une *suite de Cauchy* si :

$$\forall \varepsilon > 0 \ \ \exists n_0 \in \mathbb{N} \ \ \forall p \geq n_0, \forall q \geq n_0, d(x_p, x_q) < \varepsilon.$$

L'exercice 52 propose une reformulation de cette définition.

L'espace métrique (X, d) est *complet* si toute suite de Cauchy d'élements de X est convergente.

Lemme. *Dans tout espace métrique (X, d),*
- *toute suite convergente est de Cauchy,*
- *toute suite de Cauchy est bornée,*
- *toute suite de Cauchy ayant au moins une valeur d'adhérence converge vers cette valeur d'adhérence.*

Recette de preuve.— On considère une suite $(x_n)_{n\in\mathbb{N}}$.

Pour le premier point, on suppose que la suite est de Cauchy. **Appliquer** la définition de suite de Cauchy avec $\varepsilon = 1$ (par exemple). **Ecrire** la propriété vérifée par le n_0 obtenu. On pose $x = x_{n_0}$, et M le maximum des nombres $d(x, x_0), \ldots, d(x, x_{n_0-1})$. **Majorer** alors $d(x, x_q)$, pour tout q (séparer les cas $q < n_0$ et $q \geq n_0$).

Pour le deuxième point, la remarque essentielle est que deux points qui sont dans une même boule de rayon $\varepsilon/2$ sont à distance l'un de l'autre plus petite que ε.

Pour le dernier point, on suppose que la suite est de Cauchy et qu'elle a une valeur d'adhérence ℓ. Étant donné $\varepsilon > 0$, que cherche-t-on ? **Appliquer** d'abord la définition d'une suite de Cauchy (qu'obtient-on ?), puis la définition de valeur d'adhérence (avec quelles données l'applique-t-on ?).

Une suite (x_n) est bornée s'il existe un point x tel que la suite réelle $d(x, x_n)$ est majorée (dans ce cas, c'est vrai pour tout point x).

Théorème. $\mathbb{R}$ *muni de la distance usuelle est un espace métrique complet.*

La démonstration est reportée au chapitre suivant.

Proposition.
- *Soit Y une partie d'un espace métrique complet (X, d). Le sous-espace métrique (Y, d) est complet si et seulement si Y est une partie fermée de X.*
- *Un produit d'espaces métriques complets est complet pour les distances d_∞, d_1, d_2.*
- *Si (Y, d_Y) est complet, alors l'espace $(\mathcal{B}(X, Y), d_\infty)$ des applications bornées de X dans Y, muni de la distance de la convergence uniforme, est complet. Le sous-espace $(C_b(X, Y), d_\infty)$ des fonctions continues bornées de X dans Y est complet.*

Au premier chapitre nous avons vu trois façons de construire de nouveaux espaces métriques. Cette proposition étudie la complétude des espaces obtenus. Le premier point donne plein d'exemples d'espaces métriques complets, et plein d'exemples d'espaces métriques qui ne le sont pas (voir l'exercice 51) !

Recette de preuve.— On considère un fermé Y de X, on veut montrer que (Y, d) est complet. Comment démarre la preuve ? **Utiliser** la complétude de X, puis la fermeture de Y.

Réciproquement, on suppose que (Y, d) est complet, on veut montrer que Y est un fermé de X. Le critère séquentiel s'impose ; que doit-on alors montrer ? Comment démarre la preuve ? **Utiliser** le lemme plus haut, puis la complétude de Y pour **conclure**.

Montrons que le produit de deux espaces métriques complets est complet. On considère d'abord la distance d_∞. On part d'une suite $(x_n)_{n\in\mathbb{N}}$ qui est de Cauchy dans l'ensemble produit $X_1 \times X_2$ muni de la distance d_∞. Notons que chaque terme x_n est un élément $(x_n(1), x_n(2))$ de $X_1 \times X_2$.

1. Pour un $i = 1$ ou 2, **montrer** que la suite $(x_n(i))_{n\in\mathbb{N}}$, qui est une suite d'éléments de X_i, est de Cauchy dans (X_i, d_{X_i}). Qu'en déduit-on ?
2. On note $x(i)$ la limite de la suite $(x_n(i))_{n\in\mathbb{N}}$; on définit ainsi un élément $x = (x(1), x(2))$ de $X_1 \times X_2$.
3. **Montrer** enfin que la suite $(x_n)_{n\in\mathbb{N}}$ tend vers x dans l'espace métrique $X_1 \times X_2$.

Pour voir que les distances d_1, d_2 donnent aussi des espaces complets, on utilise le lemme (facile) suivant.

Lemme. *Soient d, d' deux distances équivalentes sur un ensemble X. Alors (X, d) est complet si et seulement si (X, d') est complet.*

Démontrer le lemme.

La preuve du troisième point est très similaire à celle du second. On part d'une suite $(f_n)_{n\in\mathbb{N}}$ qui est de Cauchy dans l'espace métrique $(\mathcal{B}(X, Y), d_\infty)$.

1. Pour un x fixé dans X, **montrer** que la suite $(f_n(x))_{n\in\mathbb{N}}$, qui est une suite d'éléments de Y, est de Cauchy dans (Y, d). Qu'en déduit-on ?
2. On note $f(x)$ la limite de la suite $(f_n(x))_{n\in\mathbb{N}}$; on définit ainsi une application $f : X \to Y$. **Montrer** que cette application est bornée.
3. **Montrer** enfin que la suite $(f_n)_{n\in\mathbb{N}}$ tend vers f dans l'espace métrique $(\mathcal{B}(X, Y), d_\infty)$.

La preuve de la complétude d'un espace métrique donnée suit très souvent ce schéma général en trois étapes : étant donnée une suite de Cauchy, (1) on construit un candidat à être sa limite, (2) on montre que le candidat est dans le bon espace, (3) on montre que la suite converge effectivement vers le candidat. Par exemple, ce schéma permet de montrer la complétude des espaces de suites, comme $\ell_1(\mathbb{R})$.

Déduire enfin la complétude du sous-espace $C_b(X, Y)$ de celle de $\mathcal{B}(X, Y)$.

Corollaire. *L'espace métrique $(\mathbb{R}^n, d_\infty)$ est complet.*

(b) Le théorème du point fixe contractant

Une homothétie de $\mathbb{R}^N$ de rapport < 1 est un exemple typique d'application contractante. Le centre de l'homothétie est son unique point fixe.

Une application $T : X \to X$ est dite *contractante* s'il existe une constante $k < 1$ telle que, pour tout $x, y \in X$,

$$d(T(x), T(y)) < kd(x, y).$$

Un *point fixe* de f est un élément x de X tel que $f(x) = x$.

Théorème. *Soit X, d un espace métrique complet, et $T : X \to X$ une application contractante. Alors l'application T admet un unique point fixe. De plus, toute suite d'élements de X définie par la donnée d'un x_0 quelconque dans X et la relation de récurrence*

$$x_{n+1} = T(x_n)$$

converge vers ce point fixe.

Ce théorème a d'innombrables applications, notamment en analyse (voir les commentaires).

Exercice 48.— La preuve qui suit utilise les formules sur les sommes de suites géométriques, plus précisément la majoration suivante. Soit $k \in]0,1[$. En utilisant la formule $\sum_{i\geq 0} k^i = \frac{1}{1-k}$, donner une majoration de

$$k^n + k^{n+1} + \cdots + k^{n+p}$$

par un terme ε_n, indépendant de p, et qui tend vers 0 lorsque n tend vers $+\infty$.

Recette de preuve.—

Le dernière phrase de l'énoncé donne la clé pour démontrer l'existence d'un point fixe. On choisi n'importe quel point x_0 dans X, et on considère la suite $(x_n)_{n\in\mathbb{N}}$ comme dans l'énoncé, autrement dit

$$x_1 = T(x_0), \quad x_2 = T(x_1), \quad x_3 = T(x_2), \quad \ldots$$

On pose $C = d(x_0, x_1)$. Montrons que la suite (x_n) converge, en procédant en trois étapes :

1. **Donner** une majoration, pour tout entier n, de $d(x_n, x_{n+1})$, qui montre que cette quantité tend vers 0 lorsque n tend vers $+\infty$.
2. Pour deux entiers n, p, **donner** une majoration de $d(x_n, x_{n+p})$ indépendante de p, avec un majorant qui tend vers 0 lorsque n tend vers $+\infty$.
3. **En déduire** que la suite (x_n) converge dans X.

Montrer ensuite que la limite de cette suite est un point fixe de T.
Enfin, **montrer** que le point fixe de T est unique.

(c) Le théorème de Baire

Théorème. *Soit (X, d) un espace métrique complet, et $(O_n)_{n\in\mathbb{N}}$ une famille dénombrable d'ouverts denses dans X. Alors l'intersection*

$$\bigcap_{n\in\mathbb{N}} O_n$$

est dense dans X.

Ce théorème peut être omis en première lecture.

Recette de preuve.—

******

II.2 Commentaires

(a) Histoire

La méthode du point fixe contractant apparaît pour la première fois en 1890, dans un mémoire du mathématicien français Emile Picard. Sa méthode lui permet de montrer l'existence de solution pour des équations différentielles ou des équations aux dérivées partielles, par une série d'approximations successives. Emile Picard ne peut pas énoncer le théorème du point fixe contractant comme on le fait maintenant, puisque la notion d'espace métrique abstrait ne sera inventée par Maurice Fréchet que 16 ans plus tard. En 1920, Stefan Banach publie un article intitulé "Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales". Les ensembles abstraits munis des opérations de Banach sont ce que nous appelons maintenant des *espaces vectoriels normés complets*, ou encore des espaces de Banach (voir le chapitre qui leur est consacré).

Banach reprend la méthode de Picard pour en faire un théorème abstrait, valable dans tout espace vectoriel normé complet. Le théorème du point fixe contractant est maintenant un outil très important, notamment en analyse : en particulier, il fournit une preuve courte du théorème de Cauchy-Lipschitz (existence et unicité de solutions locales pour les équations différentielles) ; mais il a aussi été appliqué avec succès dans bien d'autres contextes.

Une présentation sympathique du théorème (en anglais...)

(b) Application I : existence et unicité des solutions d'équations différentielles

Nous allons voir, sur un exemple simple, comment la méthode de Picard permet de montrer l'existence de solutions d'une équation différentielle. On cherche les fonctions satisfaisant les deux conditions

$$f' = f \text{ et } f(0) = 1.$$

Bien sûr, le lecteur sait déjà que la fonction exponentielle fournit une solution ; ici le jeu consiste à prétendre qu'on ne connait pas cette fonction, pour comprendre comment on montre son existence.

Le principe est de réussir à transformer le problème en une recherche de point fixe pour une certaine application contractante. Voici une idée naturelle : on pourrait considérer l'application $f \mapsto f'$. Cependant cette idée ne fonctionne pas : on ne sait pas définir un espace métrique de fonction, qui soit complet, et dans lequel cette application serait contractante. Pour remédier à ceci, on intègre l'équation $f' = f$ pour transformer le problème en une équation intégrale :

$$(f' = f \text{ et } f(0) = 1) \Leftrightarrow \forall x, f(x) = 1 + \int_0^x f(t)dt.$$

Autrement dit, les solutions à notre problème sont les points fixes de l'application

$$T : f \mapsto \left(x \mapsto f(x) = 1 + \int_0^x f(t)dt \right).$$

L'application T est un objet un peu abstrait : elle transforme une fonction en une autre fonction, ce qui explique qu'à l'intérieur de la grande parenthèse on trouve la définition d'une fonction. Ce genre d'application est parfois appelé *opérateur*.

Pour préciser les choses, on se place dans l'espace métrique des fonctions continues bornées sur un intervalle ouvert autour de 0, $X = C_b(]-\alpha, \alpha[, \mathbb{R})$ (la valeur de α sera précisée plus tard) ; on le munit de la distance uniforme d_∞. L'application T va de X dans X. Pour voir si elle est contractante, nous devons estimer

$$d_\infty(T(f), T(g)) = \sup_{x \in]-\alpha,\alpha[} d_{\mathbb{R}}(T(f)(x), T(g)(x)).$$

Dans ces formules, $T(f)$ est un élément de X, c'est-à-dire une fonction de $]-\alpha, \alpha[$ dans $\mathbb{R}$; $T(f)(x)$ est un réel.

Or, pour $x \geq 0$,

$$d_{\mathbb{R}}(T(f)(x), T(g)(x)) = \left| \int_0^x f(t)dt - \int_0^x f(t)dt \right| \leq \int_0^x |f(t) - g(t)|\, dt \leq \alpha\, d_\infty(f, g).$$

Pour $x \leq 0$ on obtient la même majoration. Finalement, en choisissant par exemple $\alpha = \frac{1}{2}$, on obtient $d_\infty(T(f), T(g)) \leq \frac{1}{2} d_\infty(f, g)$, et l'application T est contractante. Puisque l'espace métrique (X, d_∞) est complet, nous pouvons appliquer le théorème du point fixe contractant. La conclusion est que T possède un unique point fixe, autrement dit, notre problème de départ admet une unique solution : nous venons de démontrer qu'il existe une unique fonction f définie sur l'intervalle $]-\alpha, \alpha[$ est vérifiant $f(0) = 1$ et $f' = f$. Mais le théorème ne se contente pas de

C'est une limitation de la méthode de ne donner l'existence que sur un intervalle borné. En contrepartie, cette méthode marche pour une classe très générale d'équations différentielles, dont certaines n'ont pas de solution définie sur $\mathbb{R}$ tout entier.

fournir une existence théorique de la solution, il nous donne aussi un moyen de calculer des solutions approchées : il suffit de partir de n'importe quel élément f_0 de X et d'appliquer la transformation T de façon répétée. Choisissons pour f_0 la fonction nulle. Par intégration successive on obtient

$$\begin{aligned}
f_1(x) &= 1 + \int_0^x 0dt = 1 \\
f_2(x) &= 1 + \int_0^x tdt = 1 + x \\
f_3(x) &= 1 + \int_0^x (1+t)dt = 1 + x + \frac{1}{2}x^2 \\
f_4(x) &= 1 + \int_0^x (1 + t + \frac{1}{2}t^2)dt = 1 + x + \frac{1}{2}x^2 + \frac{1}{3!}x^3 \\
f_5(x) &= \dots
\end{aligned}$$

On reconnait le début du développement en série entière de la fonction exponentielle. Par récurrence, on démontre facilement que pour tout entier positif n,

$$f_n(x) = \sum_{p=0}^{n-1} \frac{1}{p!}x^p.$$

Le théorème nous indique que la suite f_n converge uniformément, sur l'intervalle $]-\alpha, \alpha[$, vers l'unique fonction f vérifiant $f(0) = 1$ et $f' = f$.

En résumé, dans un cours d'analyse, il y a deux façons assez différente de définir la fonction éxponentielle : comme une fonction égale à sa dérivée, ou bien comme somme d'une série entière. La méthode du point fixe contractant permet de voir que ces deux définitions sont valides, et qu'elles coincident (au moins sur un intervalle ouvert autour de 0). D'autre part, la méthode se généralise à une vaste famille d'équations différentielles, toutes les équations dont le membre de gauche est f' et celui de droite une formule utilisant f et x, comme par exemple (au hasard)

$$f' = \sin(f), \quad f' = f^2 + x, \quad f' = \arctan(xf^2)\ln(x), \quad \dots$$

Une telle équation s'écrit en toute généralité

$$f' = \Phi(f, x)$$

où la fonction Φ est définie sur un ouvert de $\mathbb{R}^2$, et vérifie une condition appelée "condition de Lipschitz" (voir l'exercice ci-dessous). A vrai dire, la méthode marche aussi en dimension supérieure, ce qui veut dire qu'elle se généralise aux systèmes d'équations différentielles, comme, par exemple, les célèbres équations de Lotka-Volterra utilisées pour modéliser une interaction proies-prédateurs.

Exercice 49.— En reprenant la même démarche, trouver un intervalle $]-\alpha, \alpha[$ sur lequel l'équation différentielle

$$f' = \sin(f)$$

admet une unique solution f vérifiant $f(0) = 1$. Dans les calculs d'intégrale visant à montrer que T est contractant, on utilisera la majoration $|\sin(x) - \sin(y)| \leq |x - y|$, qui est valable pour tout nombres x, y, et qui est une conséquence de l'inégalité des accroissements finis.

Exercice 50.—(Théorème de Cauchy-Lipschitz) Soit Φ une fonction définie sur un ouvert O de $\mathbb{R}^2$, et vérifiant la condition de Lipschitz : il existe une constante $C > 0$ telle que

$$\forall (x, y) \in O, \forall (x, y') \in O, \quad |\Phi(x, y) - \Phi(x, y')| \leq C. |y - y'|$$

(on dit que Φ est "lipschitzienne en y"). Soit (x_0, y_0) un point de O. En généralisant la méthode ci-dessus, trouver un $\alpha > 0$ (dépendant de C) tel qu'il existe une unique fonction f, définie sur $]x_0 - \alpha, x_0 + \alpha[$, vérifiant l'équation différentielle $f' = \Phi(f, x)$ et la condition $f(x_0) = y_0$ (appelée (*condition initiale*). Pour simplifier, on pourra supposer que $O = \mathbb{R}^2$.

(c) Application II : construction d'ensemble fractals

****

(d) Application III : une courbe de Peano

****

(e) Applications du théorème de Baire

Notion topologique d'ensemble "gros". L'ensemble des fonctions continues nulle part dérivables est dense...

II.3 Exercices

Exercice 51.— Montrer que $\mathbb{R} \setminus \{0\}$ n'est pas complet. On donnera trois arguments différents (1) à l'aide de la proposition sur les sous-espaces métriques, (2) à l'aide de la propriété disant qu'une suite convergente est de Cauchy, (3) en utilisant juste la définition d'une suite de Cauchy.

Exercice 52.— Soit (x_n) une suite vérifiant

$$\forall n \in \mathbb{N}, \forall p \in \mathbb{N}, d(x_n, x_{n+p}) \leq \frac{1}{n}.$$

Montrer qu'elle est de Cauchy. Plus généralement, montrer qu'une suite (x_n) est de Cauchy si et seulement si il existe une suite (ε_n) de nombres réels positifs, qui tend vers 0, et telle que,

$$\forall n \in \mathbb{N}, \forall p \in \mathbb{N}, d(x_n, x_{n+p}) \leq \varepsilon_n.$$

Exercice 53.— Soit (x_n) une suite de Cauchy dans un espace métrique X, (y_n) une autre suite dans X, on suppose que la suite réelle $(d(x_n, y_n)$ tend vers 0. Montrer que la suite (y_n) est aussi une suite de Cauchy.

Exercice 54.— Ecrire la version duale du théorème de Baire, concernant les fermés de X.

Exercice 55.— Utiliser le théorème de Baire pour redémontrer que $\mathbb{R}$ n'est pas dénombrable.

Exercice 56.— Soit Φ un homéomorphisme entre $\mathbb{R}$ et l'intervalle $]0,1[$; on peut par exemple prendre

$$\Phi(t) = \frac{2}{\pi}\arctan(t).$$

Soit d la distance définie sur $\mathbb{R}$ en "transportant la distance de $]0,1[$ par Φ^{-1}", c'est-à-dire définie par

$$d(x,y) = |\Phi(x) - \Phi(y)|\,.$$

1. Rappeler pourquoi $]0,1[$, muni de la distance usuelle, b'est pas complet. En déduire que $(\mathbb{R}, d)$ n'est pas complet.

2. La distance d est-elle équivalente à la distance usuelle ? Est-elle topologiquement équivalente à la distance usuelle ?

III Compacité

III.1 Théorie

(a) Compacité

Un espace métrique (X, d) est *compact* si toute suite d'éléments de X admet une valeur d'adhérence ; autrement dit, si de toute suite on peut extraire une sous-suite convergente. Une partie A d'un espace métrique est dite compacte si le sous-espace métrique (A, d) est compact.

Proposition. *Toute partie compacte de (X, d) est fermée. Si de plus (X, d) est compact, alors toute partie fermée est compacte.*

Recette de preuve.— Si A n'est pas fermé dans X il existe une suite d'éléments de A qui converge vers un point de $X \setminus A$. Cette suite n'a pas de valeur d'adhérence dans (A, d).

Réciproquement, soit A une partie fermée de l'espace compact (X, d). Prendre une suite dans A, extraire dans X, conclure.

Proposition. *Le produit d'un nombre fini d'espaces métriques compacts est compact.*

Recette de preuve.— On fait la preuve dans le cas d'un produit $X_1 \times X_2$ de deux espaces compacts. Rappelons-nous que la convergence d'une suite dans cet espace équivaut à la convergence des deux "suites coordonnées". Etant donnée une suite, **extraire une première** fois en utilisant la compacité de X_1, puis **extraire une seconde** fois en utilisant celle de X_2.

Proposition. *("L'image continue d'un compact est compacte") Soit $f : X, \to Y$ une application continue entre deux espaces métriques. Si X est compact, alors $f(X)$ est compact.*

Recette de preuve.— A partir d'une suite dans $f(X)$, **produire** une suite dans X, **utiliser** la compacité de X, puis la continuité de f pour **conclure**.

Corollaire. *Si (X, d_X) et (Y, d_Y) sont deux espaces métriques homéomorphes, le premier est compact si et seulement si le second est compact.*

Cet énoncé dit que la compacité est une propriété topologique.

Recette de preuve.— Appliquer la proposition précédente à f et f^{-1}.

Corollaire. *Si $f : X \to Y$ est une bijection continue entre deux espaces métriques, et si (X, d_X) est compact, alors sa réciproque f^{-1} est continue, et f est un homéomorphisme.*

Recette de preuve.— Utiliser le critère de continuité concernant les parties fermées, et deux des propositions ci-dessus.

(b) Compacité dans $\mathbb{R}^N$

On ne spécifie pas la distance, ce qui signifie qu'il s'agit de l'une des distances usuelles d_1, d_2, d_∞. Comme elles sont topologiquement équivalentes, elles déterminent la même notion de compacité pour les parties de $\mathbb{R}^N$.

Théorème (Bolzano-Weierstrass)**.** *Tout intervalle $[a, b]$ est compact. Autrement dit, toute suite bornée de $\mathbb{R}$ a une valeur d'adhérence.*

Démonstration. Remarquons d'abord qu'il suffit de montrer que $[0, 1]$ est compact, puisque tout intervalle $[a, b]$ est homéomorphe à $[0, 1]$ (utiliser l'application affine $x \mapsto (b - a)x + a$).

La preuve de la compacité dépend de la construction adoptée pour $\mathbb{R}$. Dans tous les cas, on se donne une suite (x_n) dans l'intervalle $[0, 1]$, dont il nous faut extraire une sous-suite convergente.

Preuve basée sur l'axiome de la borne supérieure Cet axiome spécifie que toute partie bornée dans $\mathbb{R}$ admet une borne supérieure, et a pour conséquence le théorème selon lequel toute suite monotone bornée est convergente. Puisque notre suite est bornée, on peut considérer les quantités

$$\begin{array}{rcl} y_0 & = & \sup\{x_0, x_1, x_2, \ldots\} \\ y_1 & = & \sup\{x_1, x_2, x_3, \ldots\} \\ y_2 & = & \sup\{x_2, x_3, x_4, \ldots\} \\ & \ldots & \end{array}$$

On obtient ainsi une suite $(y_k)_{k\in\mathbb{N}}$ qui est décroissante. Puisqu'elle est minorée par 0, elle converge : notons L sa limite. Il reste à voir que L est une valeur d'adhérence de la suite (x_n) (en fait, il s'agit de la plus grande des valeurs d'adhérence, aussi appelée *limite supérieure* de la suite).

Soit $\varepsilon > 0$, il s'agit de trouver un terme de la suite, d'indice arbitrairement grand, dans l'intervalle $]L - \varepsilon, L + \varepsilon[$. On y trouve d'abord un y_k (avec k arbitrairement grand), en utilisant la définition de L comme limite de la suite (y_k)), puis un x_n (avec $n \geq k$), en utilisant la définition de y_k comme borne supérieure de $\{x_k, x_{k+1}, \ldots\}$.

Preuve basée sur l'écriture décimale On utilise que tout nombre $x \in [0, 1]$ admet une écriture décimale $x = 0, \varepsilon_1\varepsilon_2\varepsilon_3 \cdots$ (y compris $1 = 0,999...$) ; et réciproquement, que toute écriture décimale de ce type définit un nombre dans $[0, 1]$.

On considère la première décimale de chacun des termes x_n. Cela nous donne une suite dans l'ensemble fini $\{0, 1, \ldots, 9\}$. D'après le principe des tiroirs, l'un des chiffres de 0 à 9 apparaît une infinité de fois, notons-le α_1 ; autrement dit il existe une suite (x_n^1), extraite de (x_n), dont les termes ont tous α_1 pour première décimale. De même, on peut extraire de la suite (x_n^1) une suite (x_n^2) de termes ayant tous un même chiffre α_2 pour deuxième décimale (remarquons que leur première décimale est bien sûr α_1). Par récurrence, on construit une infinité d'extractions $(x_n^1)_{n\in\mathbb{N}}, (x_n^2)_{n\in\mathbb{N}}, (x_n^3)_{n\in\mathbb{N}}, \ldots$ de la suite initiale, chacune extraite de la suite précédente, les x_n^i ayant pour $i$ème décimale un certain chiffre α_i.

Considérons alors le nombre

$$y = 0, \alpha_1\alpha_2\alpha_3 \cdots .$$

Nous allons vérifier que y est une valeur d'adhérence de la suite initiale ; plus précisément, montrons que la suite extraite $(x_i^i)_{i\in\mathbb{N}}$ converge vers y. Fixons un entier i. Puisque chaque suite est extraite de la précédente, le terme x_i^i apparaît dans chacune des suites $(x_n^1)_{n\in\mathbb{N}}, (x_n^2)_{n\in\mathbb{N}}, \ldots, (x_n^i)_{n\in\mathbb{N}}$. Ses i premières décimales sont donc $\alpha_1\alpha_2\cdots\alpha_i$. On en déduit

$$\left|y - x_i^i\right| \leq 10^{-i}$$

ce qui prouve que la suite $(x_i^i)_{i\in\mathbb{N}}$ converge vers y, comme voulu. □

Corollaire. *L'espace $\mathbb{R}$ est complet.*

Recette de preuve.— **Utiliser** successivement les propriétés suivantes : 1) une suite de Cauchy est bornée ; 2) une suite bornée admet une valeur d'adhérence ; 3) une suite de Cauchy ayant une valeur d'adhérence est convergente.

Proposition. *Les sous-espaces compacts de $\mathbb{R}^N$ sont les parties fermées et bornées.*

Recette de preuve.— 1) Tout cube $[a,b]^N$ est compact. 2) Un fermé borné est inclus dans un cube, il est compact comme partie fermée d'un espace compact.

Proposition. *Soit (X,d) un espace compact et $f : X \to \mathbb{R}$ une application continue. Alors f est bornée et atteint ses bornes : il existe a, b dans X tels que, pour tout x dans X,*

$$f(a) \leq f(x) \leq f(b).$$

Recette de preuve.— Commencer par le cas particulier où X est un sous-espace métrique de $\mathbb{R}$, et f est l'identité : autrement dit, montrer que toute partie compacte A de $\mathbb{R}$ contient deux éléments a, b tels que $A \subset [a,b]$. Pour ce cas, prendre pour a et b les bornes inférieures et supérieures de A, montrer qu'elles sont chacune dans l'adhérence de A, et conclure. En déduire le cas général en appliquant ceci à $A = f(X)$. Lesquelles des propriétés précédentes a-t-on utilisé dans la preuve ?

(c) Compacité et recouvrements

On famille de parties d'un ensemble X dont la réunion est égale à X est appelée *recouvrement* de X.

Lemme. *Soit (X,d) un espace métrique compact. Pour tout $\alpha > 0$ on peut recouvrir X par un nombre fini de boules de rayon ε.*

Recette de preuve.— On raisonne par contrapposée. Si la conclusion n'est pas vérifiée, on peut construire par récurrence une suite de points dont tous les termes sont à distance $\geq \alpha$ les uns des autres. Cette suite ne peut pas avoir de valeur d'adhérence.

Lemme (Lebesgue). *Soit $\{U_i\}_{i\in I}$ une famille d'ouverts qui recouvrent un espace métrique compact (X,d). Alors il existe $\alpha > 0$ tel que toute boule de rayon α est incluse dans l'un des ouverts U_i.*

Recette de preuve.— Sinon il existe une suite de boules (B_n) dont les rayons tendent vers 0, chaque boule n'étant incluse dans aucun des ouverts U_i. La suite des centres admet une valeur d'adhérence, qui est incluse dans l'un des U_i. En utilisant l'ouverture de U_i on en déduit que B_n est incluse dans U_i pour i assez grand, contradiction.

Théorème (Borel-Lebesgue). *("De tout recouvrement par des ouverts on peut extraire un sous-recouvrement fini") Un espace métrique (X,d) est compact si et seulement si, pour toute famille $(U_i)_{i\in I}$ d'ouverts qui recouvre X, il existe une sous-famille finie $U_{i_1},\dots,U_{i_N}$ qui recouvre encore X.*

Recette de preuve.— Soit (X,d) un espace métrique compact. Le lemme de Lebesgue fournit un réel $\alpha > 0$. Le lemme précédent nous dit que X est recouvert par un nombre fini de boules de rayon α. Appliquer la conclusion du lemme de Lebesgue à ces boules.

Réciproquement, supposons que (X,d) ne soit pas compact : il existe une suite $(x_n)_{n\in\mathbb{N}}$ sans valeur d'adhérence. En écrivant qu'un point x de X n'est pas valeur d'adhérence, on obtient une boule $B(x,\varepsilon(x))$ qui ne contient qu'un nombre fini de termes de la suite. La famille de toutes les boules $\{B(x,\varepsilon(x))\}_{x\in X}$ recouvre X, mais on ne peut pas en extraire un recouvrement fini (pourquoi ?...).

Exercice 57.— Déduire du théorème de Borel-Lebesgue une caractérisation duale, par les fermés, des espaces métriques compacts. (Aide : à quelle condition, portant sur les complémentaires des U_i, la famille $(U_i)_{i\in I}$ recouvre-t-elle X ?)

Théorème. *Un espace métrique (X,d) est compact si et seulement si*
- *il est complet,*
- *pour tout $\alpha > 0$ on peut recouvrir X par un nombre fini de boules de rayon α.*

Recette de preuve.— 1) Tout espace métrique compact est complet : partir de la définition de la complétude, utiliser la définition de la compacité, et l'une des propriétés élémentaires des suites de Cauchy. Ceci termine la preuve de l'implication directe, puisqu'on a obtenu la seconde propriété au paragraphe sur les recouvrements.

2) Réciproque. Dans un espace précompact, pour toute suite et tout $\varepsilon > 0$ on peut extraire une sous-suite dont tous les termes sont à distance $< \varepsilon$ les uns des autres (utiliser le principe des tiroirs). Soit maintenant (x_n) une suite dans X. En appliquant récursivement la propriété précédente on construit des suites $(x_n^1),(x_n^2),\dots,(x_n^k),\dots$ dont chacune est extraite de la précédente, la suite (x_n^k) ayant tous ses termes à distance $< 1/k$ les uns des autres. On définit une suite (y_k) extraite de (x_n) en posant $y_k = x_k^k$ (procédé diagonal). On vérifie que la suite (y_n) est de Cauchy (on a donc montré que dans un espace précompact, de toute suite on peut extraire une suite de Cauchy). Puisqu'on a supposé X complet, elle converge.

(d) Continuité uniforme

Une application $f : X \to Y$ est *uniformément continue* si pour tout $\varepsilon > 0$, il existe $\alpha > 0$ tel que pour tous $x_1, x_2 \in X$, .

$$d(x_1, x_2) < \alpha \Rightarrow d(f(x_1), f(x_2)) < \varepsilon$$

Autrement dit, f est uniformément continue si elle est continue et si, dans la définition de la continuité en un point x, on peut choisir le même α pour tous les x

Si f est k-lipschitzienne, alors elle est uniformément continue : en effet pour tout $\varepsilon > 0$, le nombre $\alpha = \frac{\varepsilon}{k}$ vérifie la définition.

Théorème (Heine)**.** *Toute application continue définie sur un espace métrique compact est uniformément continue.*

Recette de preuve.—Etant donné un $\varepsilon > 0$, la continuité de f donne en chaque point x un réel $\alpha(x) > 0$ (écrire la propriété vérifiée par $\alpha(x)$). La famille de boules $B(x, \alpha(x)/2)$ recouvre X. On en extrait un sous-recouvrement fini $B(x, \alpha(x_1)/2), \ldots, B(x, \alpha(x_N)/2)$, ce qui permet de définir un nombre α comme le minimum des nombres $\alpha(x_i)/2$. Il reste à vérifier que ce nombre α convient, ce que l'on fait à l'aide de l'inégalité triangulaire (plus précisément, en utilisant que deux points dans une boule de rayon $\eta/2$ sont à distance plus petite que η l'un de l'autre).

(e) Compacité dans les espaces de fonctions

Exercice 58.— **1.** Montrer que l'espace métrique $X = C([0,1], \mathbb{R})$ muni de la distance d_∞ n'est pas borné. En déduire qu'il n'est pas compact.

2. Soit f_0 la fonction nulle (c'est un élément de X). Montrer que la boule unité fermée centrée en f_0 est égale à la partie $Y = C([0,1], [-1,1])$ de X. On va voir dans la suite que ce fermé borné n'est pas compact.

On travaille dans $Y = C([0,1], [-1,1])$, avec la distance d_∞. Cet espace est-il compact ? Non : il existe des suites, comme $f_n(x) = x^n$, ou encore $g_n(x) = \sin(nx)$, qui n'ont pas de sous-suite convergente. Le problème, c'est que les fonctions f_n ou g_n sont "de moins en moins continues" lorsque n augmente ; en particulier, la pente en certains points devient de plus en plus forte. On peut empêcher ça en prenant des fonctions lipschitziennes.

Théorème (version élémentaire du théorème d'Arzela-Ascoli)**.** *Soit (f_n) une suite d'éléments de X. Supposons qu'il existe $K > 0$ tel que toutes les fonctions f_n sont K-lipschitziennes. Alors cette suite possède une valeur d'adhérence.*

Démonstration. Soit (f_n) vérifiant les hypothèses du théorème, avec $K = 1$ (pour simplifier). Le point-clé consiste à montrer que pour tout $\varepsilon > 0$ on peut extraire une suite dont les termes sont à distance $< \varepsilon$ les uns des autres. Ensuite, comme dans la preuve du théorème (c) (précompact et complet implique compact), on en déduit qu'on peut extraire une suite qui est de Cauchy, puis on conclut en utilisant la complétude de l'espace X.

Soit $\varepsilon > 0$, et N un entier plus grand que $1/(2\varepsilon)$: les "boules" de centres $0, 1/N, \ldots, 1$ et de rayon ε recouvrent l'intervalle $[0,1]$. Puisque la suite $(f_n(0))_{n\ in\mathbb{N}}$ prend ses valeurs dans l'espace compact $[0,1]$, on peut en extraire

une sous-suite dont tous les termes sont à distance $< \varepsilon/2$ les uns des autres. En extrayant successivement, on obtient une suite $(n_k)_{k\in\mathbb{N}}$ telle que pour tout $k, \ell \in \mathbb{N}$ et tout $i \in \{0, \dots, N\}$,

$$d\left(f_{n_k}\left(\frac{i}{N}\right), f_{n_\ell}\left(\frac{i}{N}\right)\right) < \varepsilon \quad (*).$$

Nous allons maintenant utiliser l'hypothèse que les f_n sont 1-lipschitziennes pour en déduire que pour tout $k, \ell \geq 0$,

$$d_\infty(f_{n_k}, f_{n_\ell}) < 3\varepsilon.$$

Soit $k, \ell \geq 0$, et $x \in [0, 1]$. On estime la distance entre $f_{n_k}(x)$ et $f_{n_\ell}(x)$ à l'aide de l'inégalité triangulaire, comme suit :

– x est à distance plus petite que ε de l'un des points i/N ;

– chaque fonction f_n étant 1-lipschitzienne, on en déduit que pour tout n,

$$|f_n(x) - f_n(i/N)| < \varepsilon;$$

– En utilisant cette inégalité pour $n = n_k$ et $n = n_\ell$, et $(*)$, on conclut que

$$|f_{n_k}(x), f_{n_\ell}(x)| < 3\varepsilon.$$

Puisque ceci est vrai pour tout x, on a bien $d_\infty(f_{n_k}, f_{n_\ell}) < 3\varepsilon$. On termine la preuve comme expliqué plus haut. □

Exercice 59.— Adapter la preuve au cas de fonction K-lipschitzienne avec $K > 1$. Adapter la preuve à l'espace $C([a, b], [-M, M])$ pour a, b, M quelconque.

Exercice 60.— Montrer que les suites (f_n) et (g_n) définies ci-dessus n'ont pas de sous-suite convergentes.

III.2 Commentaires

Certains espaces métriques X, comme la droite réelle, sont "trop gros" : il contiennent des suites sans valeurs d'adhérence (toute suite qui tend vers $+\infty$, par exemple), ou ils admettent des fonctions continues $X \to \mathbb{R}$ qui ne sont pas bornées. La compacité est une façon d'empêcher ces deux phénomènes ; on peut donc la voir comme une propriété de petitesse. La propriété disant qu'un compact peut être recouvert par un nombre fini de boules arbitrairement petites va aussi dans ce sens. Attention, cette interprétation peut aussi être trompeuse : $[0, 1]$ est compact alors que $]0, 1]$ ne l'est pas, pourtant le second semble plus "petit" que le premier...

La notion de compacité est très intéressante, d'un part parce qu'il y a beaucoup d'espaces compacts (voir la partie "construction d'espaces compacts"), d'autre part parce que ces espaces ont des propriétés très fortes (voir la partie "applications").

Il s'agit d'une notion intrinsèque : Si A est une partie d'un espace métrique (X, d), la définition de la compacité de (A, d) ne fait pas référence à l'espace ambiant X (autrement dit, quand on réfléchit à la compacité de A il faut "oublier" le reste de X). Par exemple, $]0, 1]$ n'est pas compact, parce que la suite $(1/n)$ n'a pas de valeur d'adhérence dans cet espace (bien qu'elle en ait évidemment dans $\mathbb{R}$).

Le théorème d'Arzéla-Ascoli permet d'obtenir des résultats d'existence de solutions d'équations différentielles, dans des cas qui échappent au théorème du point fixe contractant (théorème de Cauchy-Peano).

(a) Continuité et continuité uniforme

Deux exemples vont nous aider à comprendre les différences entre la continuité, la continuité uniforme, et le caractére lipschitzien.

Considérons d'abord la fonction réelle $x \mapsto x^2$. Cette fonction est continue : pour un $\varepsilon > 0$, en un point x donné, on peut trouver un α vérifiant la définition de la continuité :

$$\forall x', d(x, x') < \alpha \Rightarrow d(x^2, x'^2) < \varepsilon.$$

Le dessin ci-contre **** A FAIRE **** montre les meilleures valeurs possibles pour α en trois points différents (meilleures au sens que ce sont les plus grands α possibles). On voit que plus x est grand plus on doit choisir α petit ; ceci est lié au fait que la pente augmente : en un certain sens, la fonction est "moins continue" en un point x très grand, puisqu'à une petite variation de x correspond une grande variation de x^2. On peut même estimer la valeur de α : en un point x la pente de la courbe est donnée par la dérivée qui vaut $2x$, or cette pente est à peu près égale au rapport $\frac{\varepsilon}{\alpha}$, d'où la meilleure valeur possible $\alpha \simeq \frac{\varepsilon}{x}$, et cette valeur tend vers 0 lorsque x tend vers l'infini. Il est donc impossible de trouver un $\alpha > 0$ qui convienne pour tous les x : on voit que notre fonction n'est pas uniformément continue sur $\mathbb{R}$.

Si on part de $x = 0$ et qu'on l'augmente de $\delta x = 0,1$, la fonction x^2 augmente de $0,01$; si on part de $x = 1$ et qu'on l'augmente à nouveau de $\delta x = 0,1$, la fonction x^2 augmente (environ) de $0,2$; si on fait la même chose en partant de $x = 10$, de combien augmente x^2 ? Et en partant de $x = 100$? de $x = 1000$?

Considérons maintenant la fonction racine carrée. Cette fonction est continue sur $\mathbb{R}^+$. Cette fois-ci la pente décroît, et pour un $\varepsilon > 0$ donné, le α correspondant à la continuité en 0 marche en tous les points de $\mathbb{R}^+$: cette fonction est uniformément continue **** DESSIN ****. Par contre, cette fonction n'est pas lipschitzienne, parce que la pente au point $(0, 0)$ du graphe est $+\infty$. Montrons précisément ceci, en raisonnant par l'absurde : on suppose qu'il existe une constante $k > 0$ telle que, pour tous $x, x' \geq 0$,

$$d(\sqrt{x} - \sqrt{x'}) \leq k d(x, x').$$

Cette inégalité doit en particulier être vérifiée pour $x' = 0$ et pour tout $x \geq 0$, ce qui nous donne $\sqrt{x} \leq kx$. En divisant par x on obtient que pour tout $x > 0$,

$$\frac{1}{\sqrt{x}} \leq k$$

ce qui est manifestement absurde puisque le membre de gauche tend vers $+\infty$ lorsque x tend vers 0.

Exercice 61.— Démontrer précisément que $x \mapsto x^2$ ne vérifie pas la définition de la continuité uniforme sur $\mathbb{R}$. *Aide : on peut prendre $\varepsilon = 1$, raisonner par l'absurde en supposant qu'il existe un α qui marche, et trouver une contradiction en prenant $x' = x + \alpha/2$ et en faisant tendre x vers $+\infty$.*

III.3 Exercices

Exercice 62.— Montrer qu'un espace métrique compact est borné. *Aide : dans un espace X supposé non borné, construire d'abord une suite $(x_n)_{n\in\mathbb{N}}$ pour laquelle $d(x_0, x_n)$ tend vers $+\infty$. Montrer qu'on ne peut pas en extraire une sous-suite convergente.* Donner un exemple d'espace métrique borné qui n'est pas compact.

Exercice 63.— Nous avons vu que si $f : X \to Y$ est une bijection continue et si X est compact, alors f^{-1} est continue. Montrer, par un contre-exemple, que l'énoncé devient faux si on enlève l'hypothèse de compacité sur X. *Aide : considérer la fonction $\theta \mapsto e^{i\theta}$, définie sur $]0, 2\pi]$.*

Exercice 64.— Dans un espace métrique compact, montrer que toute suite ayant une unique valeur d'adhérence est convergente.

Exercice 65.— Donner un exemple de recouvrement de $[0, 1]$ par des boules de rayon $1/n$.

Exercice 66.— Donner un exemple de recouvrement de $[0, 1]$ par deux intervalles ouverts. Pour votre recouvrement, donner un $\alpha > 0$ qui vérifie la conclusion du lemme de Lebesgue. Donner un exemple de recouvrement de $\mathbb{R}$ par des ouverts, qui ne vérifie pas la conclusion du lemme de Lebesgue.

Exercice 67.— Dessiner un recouvrement du carré $[0, 1]^2$ par deux boules. Déterminer géométriquement un nombre $\alpha > 0$ qui satisfait la conclusion du lemme de Lebesgue pour votre recouvrement (quels sont les points du carré qui sont les plus contraignants concernant la valeur de α ?...).

Exercice 68.— Soit $f : X \to Y$ une bijection continue entre deux espaces métriques, et supposons que X est compact. Montrer que f est un homéomorphisme. Donner un contre-exemple sans la compacité.

Exercice 69.—(partiel 2012) Soit X un espace métrique compact, et A une partie de X. On suppose que A est *localement finie* : pour tout point x de X, il existe $\varepsilon > 0$ tel que la boule $B(x, \varepsilon)$ ne contient qu'un nombre fini de points de A.

1. Montrer que A ne contient qu'un nombre fini de points.

2. (*) Montrer que le résultat ne tient plus si, dans la définition de "localement fini", on remplace "pour tout point x de X" par "pour tout point x de A".

Exercice 70.— Trouver l'erreur dans la "solution" suivante de la première question l'exercice précédent. *Par hypothèse, pour tout x de X, il existe $\varepsilon > 0$ tel que l'ensemble $B(x, \varepsilon) \cap A$ ne contient qu'un nombre fini de points. Par compacité, on peut recouvrir X par un nombre fini de boules $B(x_1, \varepsilon), \ldots, B(x_k, \varepsilon)$. On a*

$$A \subset A \cap B(x_1, \varepsilon) \cup \cdots \cup A \cap B(x_k, \varepsilon)$$

et puisque A est une réunion finie d'ensembles finis, il ne contient qu'un nombre fini de points.

Exercice 71.— Soit $f : X \to \mathbb{R}$ une fonction localement bornée : tout point est inclus dans une boule sur laquelle la fonction est bornée. Montrer que si X est compact, alors f est bornée sur X.

Exercice 72.— Montrer que l'ensemble des fonctions continues et affines par morceaux est dense dans $(C([0,1], \mathbb{R}), d_\infty)$. *Aide : utiliser la continuité uniforme.*

Exercice 73.— On se place dans $C_b([0,1], \mathbb{R})$. Deux variantes au choix :

1) Soit n un entier, dessiner une fonction continue $f_n : [0,1] \to [0,1]$, qui s'annule en dehors de l'intervalle

$$\left[\frac{1}{2^{n+1}}, \frac{1}{2^n}\right]$$

et qui prend la valeur 1 au milieu de cet intervalle. Que vaut $d_\infty(f_n, f_m)$ pour $n \neq m$? En déduire que la suite (f_n) n'a pas de valeur d'adhérence.

2) Pour tout n, soit $f_n : x \mapsto x^n$. Pour n fixé, que vaut $d_\infty(f_n, f_m)$ lorsque m est très grand ? Formaliser. En déduire que cette suite n'a pas de valeur d'adhérence.

Exercice 74.— On se place dans un espace métrique. Soit (u_n) une suite convergent vers u. Montrer que la partie $\{u_n \mid n \in \mathbb{N}\} \cup \{u\}$ est compacte.

Exercice 75.— Soit $f : \mathbb{R}^N \to \mathbb{R}$. On suppose que $\lim_{\|x\| \to +\infty} \|f(x)\| = +\infty$; ceci revient à dire que pour toute suite (x_n) dans $\mathbb{R}^N$,

$$\lim_{n \to +\infty} d(x_n, 0) = +\infty \Rightarrow \lim_{n \to +\infty} (f(x_n)) = +\infty.$$

1. Donner des exemples d'applications vérifiant cette hypothèse, et d'applications qui ne la vérifient pas.

2. Montrer que f est minorée et atteint sa borne inférieure.

3. En déduire que pour tout polynôme $P \in \mathbb{R}[X]$ il existe $x_0 \in \mathbb{R}$ tel que, pour tout $x \in \mathbb{R}$, $|P(x)| \geq |P(x_0)|$.

Exercice 76.— Soit ABC un triangle dans le plan. Montrer qu'il existe un point M du plan tel que la somme des distances de M à A, B et C est minimale.

Exercice 77.— Etant donnés deux polynômes $P(X), Q(X)$ à coefficients réels, on définit $d(P, Q)$ comme le maximum des valeurs absolues des coefficients du polynôme $P(X) - Q(X)$. Il est facile de voir que ceci définit une distance sur l'ensemble $\mathbb{R}[X]$.

1. Calculer, pour tout entiers positifs p, q, la distance de 0 au polynôme X^p, puis la distance entre les deux polynômes X^p et X^q.

2. En déduire que la suite $(X^p)_{p\geq 0}$ d'élements de $\mathbb{R}[X]$ est bornée mais n'admet pas de sous-suite convergente.

3. Montrer que, dans cet espace métrique, la boule fermée de centre 0 et de rayon 1 n'est pas compacte.

Exercice 78.— Montrer que $O(n)$ est compact. En déduire que l'application de décomposition polaire est un homéomorphisme (...).

IV Connexité

IV.1 Théorie

(a) Connexité par arcs

Soit (X, d) un espace métrique. Un *chemin dans X joignant un point x_0 à un point x_1*, est une application continue $\gamma : [0, 1] \to X$ telle que $\gamma(0) = x_0$ et $\gamma(1) = x_1$. Un espace métrique (X, d) est *connexe par arcs* si pour tout couple de points x_0, x_1 de X, il existe un chemin joignant x_0 à x_1. Une partie A de (X, d) est dite connexe par arcs si le sous-espace métrique (A, d) l'est.

Par exemple, dans $\mathbb{R}^N$, l'application

$$\gamma : t \mapsto tx_0 + (1 - t)x_1$$

est un chemin joignant le point x_0 au point x_1.

Proposition.

1. *L'espace métrique $\mathbb{R}^N$ est connexe par arcs.*
2. *Toute partie convexe de $\mathbb{R}^N$ est connexe par arcs.*
3. *Les parties connexes par arcs de $\mathbb{R}$ sont les intervalles.*

Le dernier point pourra être redémontré plus bas comme une conséquence du fait que la connexité par arcs entraîne la connexité.

Recette de preuve.— La formule donnée avant l'énoncé pour un chemin joignant deux points quelconque montre que $\mathbb{R}^N$ est connexe par arcs. Le même chemin peut être utilisé pour voir que les parties convexes de $\mathbb{R}^N$ sont connexes par arcs, et en particulier que les intervalles de $\mathbb{R}$ sont connexes par arcs.

Montrer le dernier point en utilisant le théorème des valeurs intermédiaires et la caractérisation suivante des intervalles : une partie A de $\mathbb{R}$ est un intervalle si et seulement si pour tout triplets $a < b < c$ tels que a et c sont dans A, le point b est aussi dans A.

Concaténation Soient γ, γ' deux chemins dans X, et supposons que $\gamma'(0) = \gamma(1)$. On peut alors définir le chemin concaténé $\gamma \star \gamma'$, qui joint le point $\gamma(0)$ au point $\gamma'(1)$, de la façon suivante :

Concaténer signifie "mettre bout à bout".

$$\gamma \star \gamma' : t \longmapsto \begin{cases} \gamma(2t) & \text{si } t \in [0, \frac{1}{2}] \\ \gamma'(2t - 1) & \text{si } t \in [\frac{1}{2}, 1]. \end{cases}$$

Exercice 79.— Vérifier que le chemin concaténé est bien continu (on pourra donner une preuve directe ou utiliser l'exercice **??**).

Proposition IV.1.

- *L'image d'un espace métrique connexe par arcs par une application continue est connexe par arcs.*
- *La réunion de parties connexes par arcs ayant un point commun est connexe par arcs.*
- *Un produit fini d'espaces connexes par arcs est connexe par arcs.*

Recette de preuve.— Si γ est un chemin dans X, et $f : X \to Y$ une application continue, alors $f \circ \gamma$ est un chemin dans Y. On en déduit le premier point. Si A_1, A_2 sont deux parties connexes par arcs ayant un point P en commun, on peut concaténer un chemin de x_1 à P dans A_1 avec un chemin de P à x_2 dans A_2 pour obtenir un chemin de x_1 à x_2 dans $A_1 \cup A_2$. On en déduit le deuxième point. Soit enfin $X = X_1 \times X_2$ un produit d'espaces connexes par arcs, et $x = (x_1, x_2)$, $y = (y_1, y_2)$ deux de ses points. Utiliser un chemin joignant x_1 à y_1 dans X_1 et un chemin joignant x_2 à y_2 dans X_2 pour **fabriquer** un chemin joignant x à y. **En déduire** par récurrence le cas d'un produit d'un nombre fini d'espace connexes par arcs.

Corollaire IV.2 (invariance topologique)**.** *Soient (X_1, d_1) et (X_2, d_2) deux espaces métriques homéomorphes. Alors X_1 est connexe par arcs si et seulement si X_2 est connexe par arcs.*

(b) Composantes connexes par arcs

Dans un espace métrique (X, d), on considère la relation définie par $x \sim y$ s'il existe un chemin de x à y. On a, pour tous x, y, z,

1. (réflexivité) $x \sim x$;
2. (symétrie) $x \sim y \Rightarrow y \sim x$;
3. (transitivité) $x \sim y$ et $y \sim z \Rightarrow x \sim z$.

Etant donné un point x, l'ensemble des points y tels que $x \sim y$ est appelé *composante connexe par arcs* de x. On déduit des trois propriétés ci-dessus que les composantes connexes par arcs forment une *partition* de X, ce qui signifie qu'elle sont deux à deux disjointes et le recouvrent. Par exemple, on peut montrer que les composantes connexes par arcs d'un puvert de $\mathbb{R}$ sont des intervalles ouverts. On en déduit que tout ouvert est réunion disjointe d'un nombre dénombrable d'intervalles ouverts. (Attention, il n'y a pas de propriété analogue pour les fermés).

Exercice 80.— Démontrer mes trois propriétés ci-dessus, et en déduire que les composantes connexes par arcs forment une partition de X.

(c) Connexité

Soit (X, d) un espace métrique. Il est *connexe* si $\emptyset$ et X sont les seules partie de X qui soient à la fois ouvertes et fermées. Lorsque X n'est pas connexe, il existe donc une partie $A \neq \emptyset, X$ qui est à la fois ouverte et fermée ; le complémentaire de A est également une partie non vide qui est à la fois ouverte et fermée. On en déduit :

Proposition. *L'espace métrique X n'est pas connexe si et seulement si il existe une partition $X = O \sqcup O'$ en deux parties ouvertes non vides.*

Exercice 81.— Une application $f : X \to Y$ est dite *localement constante* si $\forall x_0 \in X, \exists \varepsilon > 0 \forall x \in B(x_0, \varepsilon), f(x) = f(x_0)$.

1. Montrer que toute application continue à valeur dans l'espace à deux éléments $Y = \{0, 1\}$ est localement constante.

2. Montrer que, si X est connexe, alors toute application localement constante est constante.

3. Réciproquement, montrer que si toute application continue de X dans $Y = \{0, 1\}$ est constante, alors X est connexe.

Proposition. *L'intervalle* $[0, 1]$ *est connexe.*

Recette de preuve.— Supposer qu'on ait une partition en deux ouverts non vides, $[0, 1] = O \sqcup O'$. Quitte à échanger O et O', on peut supposer que le premier contient 0. Soit

$$t_0 = \sup\{t \in [0, 1], [0, t] \subset O.$$

On suppose (par l'absurde) que $t < 1$. **Montrer** que t n'est pas dans O. **En déduire** que O rencontre O'. **Conclure**.

Corollaire. *Un espace métrique connexe par arcs est connexe.*

Recette de preuve.— On raisonne par contrapposition. Soit X un espace métrique qui n'est pas connexe : il existe une partition $X = O \sqcup O'$ en deux parties ouvertes non vides. Supposons qu'il existe un chemin $\gamma : [0, 1] \to X$ un chemin joignant un point de O à un point de O'. **Fabriquer** alors une partition de $[0, 1]$ en deux ouverts non vides. **Conclure**.

Proposition. *La proposition IV.1 et le corollaire IV.2 sont encore valables en remplaçant partout "connexe par arcs" par "connexe". De plus, l'adhérence d'une partie connexe est connexe.*

On en déduit que l'image d'un espace métrique connexe par une application continue à valeur dans $\mathbb{R}$ est un intervalle.

Recette de preuve.— **********

Exercice 82.— Montrer que les parties connexes de $\mathbb{R}$ sont les intervalles (*Aide : utiliser la caractérisation des intervalles vue plus haut*). Avec le corollaire ci-dessus, on en déduit un nouvel argument montrant qu'une partie connexe par arcs de $\mathbb{R}$ est un intervalle.

Exercice 83.— Montrer qu'une partie A ouverte et connexe de $\mathbb{R}^N$ est connexe par arcs. *Montrer que la composante connexe d'un point dans A est un ouvert. En déduire que c'est aussi un fermé. Conclure*

(d) Compacts connexes

Proposition. *Une intersection décroissante de compacts connexes non vides est un compact connexe non vide.*

Recette de preuve.— **********

IV.2 Commentaires

Nous avons vu que la connexité par arcs entraîne a connexité. La réciproque est fausse, comme le montre le dessin ci contre. On a dessiné le graphe de la fonction $x \mapsto \sin(1/x)$ sur un intervalle du type $]0, \alpha]$, ainsi que le segment vertical de longueur 2 centré en l'origine. Dans la version électronique de ce texte, vous pouvez zoomer sur le dessin pour distinguer plus d'oscillations du graphe ; il faut imaginer qu'au voisinage de 0 le graphe possède une infinité d'oscillations, chacune d'amplitude verticale égale à 2 unités. L'ensemble obtenu est connexe mais n'est pas connexe par arcs, parce qu'un chemin reliant un point du graphe à un point du segment ne peut pas être continu.

Pour éviter les ensembles comme celui-ci, on introduit la notion de connexité locale. On peut alors montrer, sous de bonnes hypothèses, que la connexité est équivalente à la connexité par arcs :

Théorème. *Tout espace métrique compact, connexe et localement connexe est connexe par arcs.*

(Référence ?)

Exercice 84.— On note A le graphe de la fonction $x \mapsto \sin(1/x)$ sur l'intervalle du type $]0, \pi]$.

1. Montrer que sur n'importe quel intervalle du type $[1/(a + 2\pi), 1/a]$, la fonction $x \mapsto \sin(1/x)$ prend toutes les valeurs entre 0 et 1. Dessiner le graphe au-dessus d'un intervalle de ce type lorsque x est très proche de 0.

2. Montrer que $\bar{A} = A \cup \{(0, 0)\} \times [-1, 1]$.

3. Montrer que A est connexe par arcs. En déduire que $\bar{A}$ est connexe.

4. Pour montrer que $\bar{A}$ n'est pas connexe par arcs, on raisonne par l'absurde, en considérant un chemin $\gamma : [0, 1]\ to \bar{A}$ allant du point du graphe d'absisse π à un point $(0, 0)$ du segment vertical ; on note $\gamma_x(t), \gamma_y(t)$ les coordonnées du point $\gamma(t)$ dans le plan.

a. Montrer qu'il existe $\tau \geq 0$ tel que le point $\gamma(t) = (0, y_0)$ est sur le segment vertical mais tous les points $\gamma(t)$ avec $t < \tau$ sont sur le graphe.

b. Soit $\varepsilon > 0$ assez petit, et $x = \gamma_x(\tau - \varepsilon)$. Montrer qu'il existe $t \in [\tau - \varepsilon, t_0]$ tel que

$$\frac{1}{\gamma_x(t)} = \frac{1}{x} + 2\pi.$$

c. En déduire que pour tout point $(0, \alpha)$ du segment, il existe une suite (t_n) décroissante et convergeant vers τ telle que la suite $(\gamma_y(t_n))$ converge vers $(0, \alpha)$.

d. Conclure en montrant que γ ne satisfait pas le critère de continuité séquentiel en τ.

La connexité permet le passage du local au global. Exemple...

Exemple de joli théorème de connexité : le retournement de la sphère...

(a) Connexité et homéomorphismes

Non avons vu que la connexité est une propriété topologique : un espace connexe n'est jamais homéomorphe à un espace qui ne l'est pas. La connexité fournit ainsi une façon de montrer que deux espaces ne sont pas homéomorphes. On peut aussi utiliser cette propriété en l'appliquant à des parties de l'espace. Par exemple, le plan $\mathbb{R}^2$ privé d'un point est connexe par arcs, mais pas $\mathbb{R}$ privé d'un point, ce qui permet de montrer que $\mathbb{R}$ et $\mathbb{R}^2$ ne sont pas homéomorphes.

Exercice 85.— Montrer en détail que $\mathbb{R}$ et $\mathbb{R}^2$ ne sont pas homéomorphes, en suivant l'idée ci-dessus.

En réalité on peut montrer que $\mathbb{R}^p$ et $\mathbb{R}^q$ ne sont homéomorphes pour aucune valeurs de $p \neq q$. Une première façon de faire ça est de généraliser la notion de connexité par arcs. Ainsi, on montre par exemple que $\mathbb{R}^3$ privé d'un point est simplement connexe, alors que $\mathbb{R}^2$ privé d'un point ne l'est pas (mais ce point est plus délicat), ce qui montre que $\mathbb{R}^2$ et $\mathbb{R}^3$ ne sont pas homéomorphe.

IV.3 Exercices

Exercice 86.— Montrer que le graphe d'une fonction $f : I \to \mathbb{R}$ continue sur un intervalle I est une partie connexe par arcs du plan.

Exercice 87.— Soit X le graphe de la fonction $x \mapsto 1/x$, de $\mathbb{R}^*$ dans $\mathbb{R}^*$. Montrer que X n'est pas connexe.

Exercice 88.— Montrer que toute application continue d'un espace métrique connexe dans $\mathbb{Z}$ est constante.

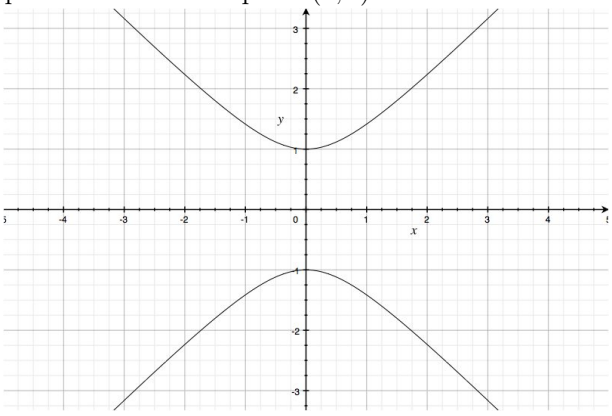
Exercice 89.— (version en dimension deux du théorème de Borsuk-Ulam). On note $\mathbb{S}^1$ le cercle unité de centre 0 dans le plan ; on le voit comme un sous-espace métrique du plan. Soit f une application continue de $\mathbb{S}^1$ dans $\mathbb{R}^2$. Montrer qu'il existe deux points du cercle, diamétralement opposés, qui ont la même image par f. *Aide : utiliser la connexité du cercle et la fonction auxiliaire définie par $g(v) = f(v) - f(-v)$ (noter que v et $-v$ sont diamétralement opposés).*

En déduire le théorème du partage de crêpes : étant données deux régions bornées A, B du plan, il existe une droite qui divise chacune de ces régions en deux parties d'aires égales (pas facile !!).

Exercice 90.—(partiel 2012) Dans le plan $\mathbb{R}^2$, on considère la partie $A = \{(x, y) \mid x^2 - y^2 = -1\}$, qui est dessinée ci-dessous.

1. La partie A est-elle connexe par arcs ? Donner une preuve complète de votre réponse.

2. Déterminer la composante connexe du point $(0, 1)$ dans A.



V Espaces vectoriels normés, espaces de Banach

On va considérer des espaces vectoriels réels ou complexes. On utilise la distance usuelle sur $\mathbb{R}$, et sur $\mathbb{C}$ la distance $d(z, z') = |z - z'|$ (qui s'identifie à la distance euclidienne sur $\mathbb{R}^2$).

V.1 Théorie

(a) Définitions

Soit E un espace vectoriel sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Une *norme sur E* est une application $x \mapsto \|x\|$, de E dans $\mathbb{R}^+$, vérifiant :

1. $\forall x \in E,\ \|x\| = 0$ ssi $x = 0$,
2. $\forall x \in E,\ \forall \lambda \in \mathbb{K},\quad \|\lambda x\| = |\lambda| \,.\, \|x\|$,
3. $\forall x, y \in E,\ \|x + y\| \leq \|x\| + \|y\|$.

Le couple $(E, \|.\|)$ est alors appelé *espace vectoriel normé* (evn).

Proposition. *Si $(E, \|.\|)$ est un evn, la fonction définie sur $E \times E$ par $d(x, y) = \|x - y\|$ est une distance.*

La preuve est immédiate. Le couple (E, d) est appelé *espace métrique associé* à l'evn $(E, \|.\|)$. Toutes les notions métriques et topologiques des chapitres précédents ont donc un sens dans un evn : suites convergentes, ouverts, ...

Un espace vectoriel normé est un espace métrique dans lequel on sait faire des combinaisons linéaires.

Exercice 91.—

1. Soient $(u_n), (v_n)$ deux suites convergentes dans $E, \|.\|$. Montrer que la suite $(u_n + v_n)$ est convergente. Si (λ_n) est une suite convergente de réels, montrer que la suite $(\lambda_n u_n)$ converge. Avec le critère séquentiel de continuité, ces résultats s'interprètent en disant que les applications $(u, v) \mapsto u + v$ et $(\lambda, u) \mapsto \lambda.u$ sont continues.

2. Montrer, pour tout $x, y \in E$, l'inégalité

$$|\|x\| - \|y\|| \leq \|x - y\| \,.$$

Cette inégalité signifie exactement que l'application $x \mapsto \|x\|$ est 1-lipschitzienne de $(E, \|.\|)$ vers $(\mathbb{R}, |.|)$.

Exemples fondamentaux

• Soit E un $\mathbb{R}$-espace vectoriel de dimension N, et $\mathcal{B} = (e_1, \ldots, e_N)$ une base de E. Pour un vecteur $x = \sum_{i=1}^{N} x_i e_i$ de E, on pose

$$\|x\|_1 = \sum_{i=1}^{N} |x_i| \,, \qquad \|x\|_2 = \sum_{i=1}^{N} x_i^2, \qquad \|x\|_\infty = \max_{i=1}^{N} |x_i| \,.$$

Proposition. *Les applications $\|.\|_1$, $\|.\|_2$, $\|.\|_\infty$ sont des normes.*

Recette de preuve.— Exercice. Pour la norme 2, la preuve de l'inégalité triangulaire utilise l'inégalité de Cauchy-Schwarz.

• Soit maintenant $E = \mathcal{B}(X, \mathbb{K})$ l'espace vectoriel des fonctions bornées d'un espace métrique (X, d) dans $\mathbb{K}$. Pour $f \in E$ on pose

$$\|f\|_\infty = \sup_{x \in X} |f(x)| .$$

Proposition. *Cette formule définit une norme sur $\mathcal{B}(X, \mathbb{K})$.*

Recette de preuve.— Exercice.

L'ensemble $F = C_b(X, \mathbb{K})$ des fonctions continues bornées forme un sous-espace vectoriel de $E = \mathcal{B}(X, \mathbb{K})$. Prenons $X = [0, 1]$ et $\mathbb{K} = \mathbb{R}$. L'espace $G = C^1([0, 1], \mathbb{R})$ est un sous-espace vectoriel de $F = C([0, 1], \mathbb{R})$. Les espaces F et G peuvent être munis de la norme $\|.\|_\infty$. Sur l'espace G, on peut aussi définir une autre norme en posant $N(f) = \|f\|_\infty + \|f'\|_\infty$.

(b) Continuité des applications linéaires

On considère deux espaces vectoriels normés E et F.

Proposition. *Soit $L : E \to F$ une application linéaire. Sont équivalentes :*
(i) L est lipschitzienne,
(ii) L est continue,
(iii) L est continue en 0,
(iv) L est bornée sur $B(0, 1)$,
(v) L est bornée sur $S(0, 1)$.
(vi) Il existe $C > 0$ tel que, pour tout x, $\|L.x\| \leq C \|x\|$.

On notera souvent $L.x$ l'image d'un vecteur x de E par l'application linéaire L. Cette notation rappelle le produit matriciel auquel elle correspond en dimension finie.

Lorsque L est continue, on a les égalités

$$\sup_{x \neq 0} \frac{\|L.x\|}{\|x\|} = \sup_{x \in B(0,1)} \|L.x\| = \sup_{x \in S(0,1)} \|L.x\| = \inf\{C \mid \forall x, \|L.x\| \leq C \|x\|\}.$$

Cette égalité dit notamment que les bornes des points (iv) et (v) sont égales, et égales à la meilleure constante C du point (vi).

Recette de preuve.— Le point-clé est de remarquer que (évidemment) pour tout vecteur x non nul, le vecteur

$$\frac{x}{\|x\|}$$

est de norme 1. Utiliser cette remarque pour **montrer** l'implication $(v) \Rightarrow (vi)$.

Pour l'implication $(iii) \Rightarrow (iv)$, **appliquer** la définition de la continuité en 0 avec $\varepsilon = 1$. On en déduit une boule de rayon α sur laquelle L est majorée. **Obtenir** (iv) en utilisant à nouveau la remarque-clé.

Toutes les autres implications sont immédiates.

Les trois premières égalités suivent à nouveau de la remarque-clé. Pour la dernière, appelons

$$A = \left\{ \frac{\|L.x\|}{\|x\|} \mid x \neq 0 \right\}.$$

L'ensemble des majorants de A est

$$B = \{C \mid \forall x, \|L.x\| \leq C \|x\|\} .$$

La borne supérieure étant le plus petit des majorants, on a $\sup A = \inf B$, comme voulu.

Exercice 92.— La matrice $M = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ définit une application linéaire $x \mapsto M.x$ de $\mathbb{R}^2$ dans lui-même. On munit $\mathbb{R}^2$ de la norme $\|.\|_\infty$, montrer par un calcul direct que cette application vérifie le dernier item de la proposition (elle est donc continue).

Exercice 93.— Soient N_1, N_2 deux normes sur un même ev E. Supposons que l'identité soit continue de (E, N_1) vers (E, N_2), ainsi que de (E, N_2) vers (E, N_1). Ecrire la condition (vi) ci-dessus. Qu'obtient-on ?

Exercice 94.— Applications bilinéaires

1. Soit $b : E_1 \times E_2 \to F$ une application bilinéaire. Montrer les équivalences entre :
(i) b est continue,
(ii) b est continue en $(0, 0)$,
(iii) il existe une constante $C > 0$ telle que, pour tout $(x_1, x_2) \in E_1 \times E_2$,

$$\|b(x_1, x_2)\|_F \leq C \|x_1\|_{E_1} \|x_2\|_{E_2} .$$

2. Montrer que la quantité

$$\|(x_1, x_2)\| \stackrel{def}{=} \sup_{x_1 \in B_{E_1}(0,1), x_2 \in B_{E_2}(0,1)} \|b(x_1, x_2)\|_F$$

définit une norme sur l'espace vectoriel des applications bilinéaires continues.

3. Montrer qu'en dimension finie toutes les applications bilinéaires sont continues.

On aurait des résultats analogues pour les applications multilinéaires continues (voir Wikipedia).

(c) Dimension finie

Une *isométrie* entre les deux espaces vectoriels normés $(E, \|.\|_E)$ et $(F, \|.\|_F)$ est une application $\Phi : E \to F$ qui est un isomorphisme d'espaces vectoriels (application linéaire bijective) telle que, pour tout x dans E, $\|\Phi(x)\|_F = \|x\|_E$.

Deux normes $\|.\|_1, \|.\|_2$ sur un même espace vectoriel E sont *équivalentes* s'il existe $K > 0$ telle que, pour tout x dans E,

$$\|x\|_1 \leq K \|x\|_2 \text{ et } \|x\|_2 \leq K \|x\|_1 .$$

Les distances associées sont alors équivalentes ; en particulier la topologie est la même, et la notion de complétude est la même ; il est clair que la notion de partie bornée aussi.

Deux espaces isométriques ont les mêmes propriétés métriques : $(E, \|.\|_E)$ est complet si et seulement si $(F, \|.\|_F)$ est complet, une partie A de E est bornée si et seulement si $\Phi(A)$ est bornée, A est fermée si et seulement si $\Phi(A)$ est fermée, *etc.*. Les preuves de toutes ces propriétés sont immédiates.

Théorème. *Dans un ev de dimension finie, toutes les normes sont équivalentes, et la boule unité fermée est compacte. Tous les evn de dimension finie sont complets.*

Recette de preuve.— Soit $(E, \|.\|)$ un evn de dimension finie : par définition, il admet une base finie $\mathcal{B} = (e_1, \ldots, e_N)$. Cette base permet de définir une norme particulière

$$\left\| \sum_{i=1}^{N} x_i e_i \right\|_{\infty, \mathcal{B}} = \max(|x_1|, \ldots |x_N|).$$

Nous allons montrer que la norme "inconnue" $\|.\|$ est équivalente à cette norme particulière. Il y a deux inégalités à établir ; on va obtenir la première assez facilement, en utilisant seulement la définition d'une norme. De façon un peu surprenante, on pourra ensuite déduire de cette première inégalité la deuxième inégalité.

(1) Soit $\Phi : \mathbb{R}^N \to E$ l'isomorphisme $(x_i) \mapsto \sum x_i e_i$. **Montrer** que Φ est une isométrie de $(\mathbb{R}^N, \|.\|_\infty)$ vers $(E, \|.\|_{\infty,\mathcal{B}})$. **En déduire** que dans l'espace vectoriel normé $(E, \|.\|_{\infty,\mathcal{B}})$ est complet, et que les compacts sont les fermés bornés. En particulier, noter que la sphère unité de cet espace est compacte.

(2) **Trouver** une constante C_1 telle que pour tout x de E,

$$\|x\| \leq C_1 \|x\|_{\infty,\mathcal{B}}.$$

En particulier l'application $\|.\| : E \to \mathbb{R}^+$ est C_1-lipschitzienne de $(E, \|.\|_{\infty,\mathcal{B}})$ vers $\mathbb{R}^+, |.|$, et a fortiori continue.

(3) D'autre part on a vu que la sphère unité de $(E, \|.\|_{\infty,\mathcal{B}})$ est compacte. Cette application y atteint donc son minimum, ce qui nous apprend que ce minimum est strictement positif : il existe $C_2 > 0$ tel que pour tout x de E,

$$\left\| \frac{x}{\|x\|_{\infty,\mathcal{B}}} \right\| \geq C_2$$

ce qui donne la deuxième inégalité recherchée : les deux normes sont donc équivalentes. En particulier, les notions de fermés, de bornés, de compacts sont les même.

(4) Toutes les normes sont équivalentes à la norme particulière $\|.\|_{\infty,\mathcal{B}}$, et par transitivité elles sont donc toutes équivalentes.

(5) Montrons que la boule unité B de $(E, \|.\|)$ est compacte. Par continuité de la norme (exercice 91), c'est un ensemble fermé, et il est clairement borné. Par équivalence des normes, la topologie de $(E, \|.\|)$ est la même que celle de $(E, \|.\|_{\infty,\mathcal{B}})$, et dans cet espace on a vu que les fermés bornés sont compacts. Donc B est compact.

(6) Puisque la norme $\|.\|$ est équivalente à la norme $\|.\|_{\infty,\mathcal{B}}$ et que $(E, \|.\|_{\infty,\mathcal{B}})$ est complet, $(E, \|.\|)$ est aussi complet.

Corollaire. *Tout sev de dimension finie dans un evn est fermé.*

Recette de preuve.— Utiliser la complétude.

Proposition. *Si $(E, \|.\|_E)$ et $(F, \|.\|_F)$ sont deux espaces vectoriels normés et si E est de dimension finie, alors toutes les applications linéaires de E dans F sont continues.*

Recette de preuve.— On reprend les notations des preuves précédentes pour l'espace E, et on considère une application linéaire $L : E \to F$. D'après le critère de continuité obtenu à la section précédente, il s'agit de trouver une constante C telle que, pour tout $x \in E$, $\|L.x\|_F \leq C \|x\|_E$. **Trouver** une constante C' telle que

$$\left\| L.\left(\sum x_i e_i\right) \right\|_F \leq C' \|x\|_{\infty,\mathcal{B}}.$$

Utiliser maintenant l'équivalence des normes en dimension finie pour en déduire la continuité de L.

(d) Espaces de Banach

Un *espace de Banach* est un espace vectoriel normé **complet**. On a vu au paragraphe précédent que tous les evn de dimension finie sont complets. Voici d'autres exemples fondamentaux d'espaces de Banach.

Théorème. *L'espace $\mathcal{B}(X, \mathbb{K})$, muni de la norme $\|.\|_\infty$, est un espace de Banach.*

Recette de preuve.— Déjà vu au chapitre sur la complétude.

Soient E, F deux espaces vectoriels normés, on note $\mathcal{L}(E, F)$ l'ensemble des applications linéaires continues de E dans F. Pour tout élément ℓ de $\mathcal{L}(E, F)$, on pose

$$\|\ell\|_{\mathcal{L}(E,F)} = \sup_{x \in B(0,1)} \|\ell(x)\| .$$

Il s'agit du nombre obtenu à la fin de la proposition sur les applications linéaires continues.

Théorème. *Le couple $(\mathcal{L}(E, F), \|.\|_{\mathcal{L}(E,F)})$ est un espace vectoriel normé. De plus, Si F est un espace de Banach alors $\mathcal{L}(E, F)$ aussi.*

Recette de preuve.— Il n'est pas difficile de montrer que $\mathcal{L}(E, F)$ est un sous-espace vectoriel de l'espace vectoriel des applications linéaires de E dans F.

Pour montrer l'inégalité triangulaire, majorer $\|\ell.x + \ell'.x\|$.

Pour la complétude, commençons par remarquer que la preuve est facile en dimension finie....*****

Pour montrer que c'est complet en toute généralité, on s'inspire des preuves de complétudes des espaces produit ou de fonctions : grâce à la complétude de l'espace d'érrivée on trouve un candidat à être la limite, puis on prouve que le candidat appartient bien à l'espace de fonctions considérées, puis on montre la convergence. On considère une suite de Cauchy (ℓ_n). En évaluant sur chaque $x \in E$ et en utilisant la complétude de F, on définit une application $\ell : E \to F$, candidate à être la limite. On vérifie que ℓ est linéaire (ceci utilise la continuité de l'addition et de la multiplication dans F). La suite des normes triples des ℓ_n est une suite de Cauchy dans $\mathbb{R}$, donc elle est bornée, on en déduit que ℓ est une ALC. Il reste à voir que la suite (ℓ_n) converge vers ℓ : pour cela on écrit la définition d'une suite de Cauchy (... $\|\ell_p - \ell_q\|_{\mathcal{L}(E,F)} < \varepsilon$), on évalue en un x quelconque, on fait tendre p vers $+\infty$, on en déduit que ℓ_q est proche de ℓ.

Séries dans un espace de Banach Etant donnée une suite (x_n) dans un espace vectoriel normé, on définit la *suite des sommes partielles* (S_N) en posant

$$S_N = \sum_{n=0}^{N} x_n.$$

La définition d'une série convergente utilise à la fois la structure d'espace vectoriel (pour définir les sommes partielles) et la structure d'espace métrique (pour la convergence).

Lorsque la suite (S_N) admet une limite, on la note $\sum_{n=0}^{+\infty} x_n$ et on dit que *la série de terme général x_n est convergente*. On dit que la série est *absolument convergente* si la série de terme général $\|x_n\|$ est convergente.

Proposition. *Dans un espace de Banach, toute série absolument convergente est convergente.*

Recette de preuve.— **********

(e) L'algèbre normée $\mathcal{L}(E)$

Soit E un espace vectoriel normé. On note $\mathcal{L}(E)$ l'espace vectoriel des applications linéaires continues de E dans E, et on le munit de la norme définie dans la section précédente,

$$\|L\|_{\mathcal{L}(E)} = \sup_{x \in E \setminus \{0\}} \frac{L.x}{x},$$

ce qui en fait un espace vectoriel normé. La composition des applications fournit en outre un produit sur cet espace. La proposition suivante décrit la façon dont se comporte ce produit vis à vis de la norme.

Proposition.

Pour toutes applications $\ell_1, \ell_2 \in \mathcal{L}(E)$ on a

$$\|\ell_2 \circ \ell_1\|_{\mathcal{L}(E)} \leq \|\ell_2\|_{\mathcal{L}(E)} \|\ell_2\|_{\mathcal{L}(E)}$$

Recette de preuve.— *****

On résume les propriétés de $\mathcal{L}(E)$ en disant qu'il s'agit d'une *algèbre normée*. Lorsque E est un $\mathbb{R}$-espace vectoriel de dimension finie N, les éléments de $\mathcal{L}(E)$ s'identifie aux matrices réelles de taille N, et le produit dans $\mathcal{L}(E)$ s'identifie au produit matriciel. Une fois choisie une norme sur $\mathbb{R}^N$, la norme sur l'espace vectoriel $\mathcal{M}_N(\mathbb{R})$ est définie de façon analogue par

$$\|M\| = \sup_{x \in \mathbb{R}^N \setminus \{0\}} \frac{\|M.x\|}{\|x\|}.$$

L'algèbre normée $\mathcal{M}_N(\mathbb{R})$ est de dimension finie N^2, elle est donc complète.

On suppose dorénavant que E est complet. Etant donné un élément L de $\mathcal{L}(E)$, on peut former la série

$$\sum_{n=0}^{+\infty} L^n.$$

Est-elle convergente, et si oui, vers quoi converge-t-elle ?

Dans toute la suite, on peut garder en tête le cas de la dimension finie, et penser ainsi aux applications linéaires L comme à des matrices.

Exercice 95.—

1. Supposons que $\|L\|_{\mathcal{L}(E)} < 1$. Montrer que dans ce cas, la série converge.

2. Dans le cas de dimension 1, L est un nombre réel dans $]-1, 1[$. Que vaut alors la somme de cette série ?

L'élément L est dit *inversible* s'il existe un autre élément, noté L^{-1}, tel que $L.L^{-1} = L^{-1}L = \mathrm{Id}$, où Id désigne l'identité de E, qui est l'élément neutre pour le produit.

Ceci revient à demander que l'application L soit une bijection de E dans E, mais on préfère faire appel aux propriétés algébriques des éléments de $\mathcal{L}(E)$, plutôt qu'à leur nature d'application linéaire continue.

Proposition. *Soit E un espace vectoriel normé complet.*

1. *Si $\|L\| < 1$, alors $\mathrm{Id} + L$ est inversible, et*

$$(\mathrm{Id} + L)^{-1} = \sum_{k \geq 0} (-L)^k.$$

2. *L'ensemble $GL(E)$ des applications linéaires inversibles est un ouvert de $\mathcal{L}(E)$.*
3. *L'application*

$$\mathrm{Inv} : \left\{ \begin{array}{rcl} GL(E) & \longrightarrow & GL(E) \\ L & \longmapsto & L^{-1} \end{array} \right.$$

est continue.

Exercice 96.— Soit $L_0 \in \mathcal{L}(E)$. Les applications $L \mapsto L_0 L$ et $L \mapsto L L_0$ sont linéaires (de $\mathcal{L}(E)$ dans $\mathcal{L}(E)$) et continues.

Exercice 97.— Si L_1 et L_2 sont inversibles, alors $L_1 L_2$ l'est aussi, et $(L_1 L_2)^{-1} = L_2^{-1} L_1^{-1}$.

L'inverse du produit est le produit des inverses, mais dans l'ordre inverse...

Recette de preuve.— Pour le premier point, commencer par **retrouver** la preuve de la convergence, dans $\mathbb{R}$, de la série $\sum h^k$, lorsque $h \in [0, 1[$ en développant le produit

$$(1-h) \sum_{k=0}^{k_0} h^k$$

avant de faire tendre k_0 vers $+\infty$. **Adapter** ensuite cette preuve à l'algèbre normée $\mathcal{L}(E)$, en utilisant le premier exercice ci-dessus. Remarquons que les matrices qui s'écrivent $\mathrm{Id} + L$ avec $\|L\| < 1$ sont exactement les éléments de la boule ouverte $B(\mathrm{Id}, 1)$. Ainsi, ce premier point s'interprète en disant que $B(\mathrm{Id}, 1) \subset GL(E)$, ce qui montre que l'identité est dans l'intérieur de $GL(E)$.

Pour montrer que $GL(E)$ est ouvert, on considère maintenant un élément inversible L_0, et on cherche une boule centrée en L_0 et incluse dans $GL(E)$. Autrement dit, on considère $L_0 + H$ avec H de petite norme (à déterminer), et on cherche à montrer que $L_0 + H$ est inversible. L'idée est de se ramener à l'identité en factorisant par L_0 : on écrit

$$L_0 + H = L_0(\mathrm{Id} + L_0^{-1} H).$$

Trouver un nombre C tel que l'inégalité $\|H\| < C$ entraine $\left\|L_0^{-1} H\right\| < 1$. **En déduire** que, dans ce cas, $L_0 + H$ est inversible (utiliser le premier point et le second exercice), puis **décrire** une boule ouverte centrée en L_0 est incluse dans $GL(E)$, comme voulu.

Passons au dernier point, la continuité. La formule donnant l'inverse de $\mathrm{Id} + H$, pour $\|H\| < 1$, peut s'écrire

$$(\mathrm{Id} + H)^{-1} = \mathrm{Id} - H \sum_{n=0}^{+\infty} (-H)^n, \qquad \text{ou encore} \qquad \mathrm{Inv}(\mathrm{Id} + H) = \mathrm{Inv}(\mathrm{Id}) + R(H)$$

avec $R(H) = -H \sum_{n=0}^{+\infty} (-H)^n$. Montrons d'abord que Inv est continue au point Id, ce qui revient à voir que $R(H)$ tend vers 0 lorsque H tend vers 0. Pour ceci, en supposant $\|H\| < 1/2$, **trouver** une majoration de $R(H)$ qui permet de conclure. Pour voir que Inv est continue en un point L_0 quelconque, on utilise à nouveau la factorisation pour se ramener en l'identité : d'après la factorisation plus haut et le second exercice, on a

$$(L_0 + H)^{-1} = (\mathrm{Id} + L_0^{-1} H)^{-1} L_0^{-1}.$$

Conclure en utilisant la continuité au point Id et le premier exercice.

(f) Le théorème de Riesz

Théorème. *Soit E un espace vectoriel normé. Alors la boule unité fermée $B_f(0,1)$ est compacte si et seulement si E est de dimension finie.*

Frédéric Riesz (1880-1956), mathématicien hongrois.

Recette de preuve.— Il s'agit seulement de montrer le sens direct puisque nous avons déjà vu la réciproque. On va utiliser au cours de la preuve que tout sev de dimension finie est fermé.

On suppose que la boule unité fermée de l'espace vectoriel normé E est compacte, et on cherche à montrer que E est de dimension finie. On trouve d'abord ***(comment ?)*** un ensemble fini de points $x_1, \dots, x_N$ de la boule unité $B_f(0,1)$ tels que

$$(*) \quad B_f(0,1) \subset B\left(x_1, \frac{1}{2}\right) \bigcup \cdots \bigcup B\left(x_N, \frac{1}{2}\right).$$

Soit F le sev engendré par $\mathcal{F}_1 = \{x_1, \dots x_N\}$ (l'ensemble des combinaisons linéaires de ces points), qui est clairement de dimension finie. Pour conclure, nous allons montrer que $F = E$. Nous commençons par essayer de montrer que F est dense dans $B_f(0,1)$.

Pour décrire la preuve nous avons besoin d'une notion et d'une remarque. La notion est celle d'ε-densité : nous dirons qu'on partie $\mathcal{F}$ de E est *ε-dense dans* $B_f(0,1)$ si tout point de $B_f(0,1)$ est à distance inférieure à ε d'au moins un point de $\mathcal{F}$, autrement dit si

$$B_f(0,1) = \bigcup_{x \in \mathcal{F}} B(x, \varepsilon).$$

Le recouvrement $(*)$ s'interpréte alors en disant que la famille $\mathcal{F}_1$ est $\frac{1}{2}$-dense dans $B_f(0,1)$.

La remarque est la suivante. L'homothétie $x \mapsto \frac{1}{2}x$ de rapport $1/2$ est une application linéaire inversible de E dans E, qui multiplie toutes les distances par $\frac{1}{2}$. En paticulier, elle envoie une boule $B(x,R)$ sur la boule $B(\frac{1}{2}x, \frac{1}{2}R)$. La translation $x \mapsto x + v$ est une isométrie, elle envoie la boule $B(x,R)$ sur la boule $B(x+v,R)$.

Munis de cette notion et de cette remarque, poursuivons la preuve. Considérons l'un des points x_i trouvés plus haut, et appliquons successivement au recouvrement $(*)$ l'homothétie de rapport $1/2$ et la translation de vecteur x_i : nous trouvons que

$$B_f\left(x_i, \frac{1}{2}\right) \subset B\left(x_i + \frac{1}{2}x_1, \frac{1}{4}\right) \bigcup \cdots \bigcup B\left(x_i + \frac{1}{2}x_N, \frac{1}{4}\right).$$

En faisant de même pour chacun des x_i, nous voyons que la famille

$$\mathcal{F}_2 = \left\{ x_{i_1} + \frac{1}{2}x_{i_2} \mid i_1, i_2 \in \{1, \dots, N\} \right\}$$

est $\frac{1}{4}$-dense dans $B_f(0,1)$. En utilisant le même procédé, **montrer** que la famille

$$\mathcal{F}_3 = \left\{ x_{i_1} + \frac{1}{2}x_{i_2} + \frac{1}{4}x_{i_3} \mid i_1, i_2, i_3 \in \{1, \dots, N\} \right\}$$

est $\frac{1}{8}$-dense dans $B_f(0,1)$. Notez que les famille $\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3$ sont incluses dans F ; *a fortiori*, F est donc $\frac{1}{8}$-dense dans $B_f(0,1)$. Par récurrence, on montre ainsi que, pour tout n, F est $\frac{1}{2^n}$-dense dans $B_f(0,1)$.

Ca manque d'un dessin...

On en déduit immédiatement que tout point z de $B_f(0,1)$ est limite d'une suite de points de F. Souvenons-nous maintenant que F, étant de dimension finie, est fermé dans E : on conclut que $z \in F$. Autrement dit F contient $B_f(0,1)$. Comme tout vecteur de E est multiple d'un vecteur de la boule, F est égal à E.

V.2 Commentaires

(a) Pièges et délices de la dimension infinie

Nous avons vu que les espaces vectoriel de dimension finie ont plein de bonnes propriétés : toutes les normes sont équivalentes, ils sont automatiquement complets, les sous-espaces vectoriels sont fermés, et toutes les applications linéaires sont continues. En, dimension infinie, aucune de ces bonnes propriétés ne subsiste. Voyons ceci sur un exemple. On se place dans l'espace vectoriel $E = C([0,1],\mathbb{R})$ des applications continues de l'intervalle dans $\mathbb{R}$, que l'on munit de la norme $\|.\|_\infty$. On a déjà vu que la boule unité fermé de cet espace n'est pas compacte, on peut même donner deux arguments très différents : d'abord c'est une conséquence du théorème de Riesz, puisque cet espace vectoriel est de dimension infinie ; ensuite, la suite d"éléments $(x \mapsto x^n)_{n\in\mathbb{N}}$ n'admet pas de valeur d'adhérence (voir le paragraphe sur la compacité dans les espaces de fonctions).

Considérons maintenant le sous-ensemble F des applications qui sont de classe C^1 (dérivables avec une fonction dérivée qui est continue). La propriété d'être de classe C^1 est conservée par combinaisons linéaires : c'est donc un sous-espace vectoriel. En analyse, on démontre que toute fonction continue est la limite d'une suite de fonctions de classe C^1 ; autrement dit, le sous-espace vectoriel F est dense dans E. En particulier, il n'est pas fermé, et $(F, \|.\|_\infty)$ est un exemple d'espace vectoriel normé qui n'est pas complet. L'application $f \mapsto f'(0)$ est définie de F vers $\mathbb{R}$, elle est clairement linéaire. Considérons la suite de fonctions (f_n) définies par $f_n(x) = \frac{1}{n}\sin(nx)$. On a $f_n'(0) = 1$ pour tout n. D'autre part cette suite converge uniformément vers la fonction nulle dont la dérivée est bien sûr nulle. Ceci montre que l'application linéaire $f \mapsto f'(0)$ n'est pas continue.

On peut munir l'espace vectoriel F d'une autre norme, en posant

$$\|f\| = \|f\|_\infty + \|f'\|_\infty .$$

On montre alors que l'espace vectoriel normé $(F, \|.\|)$ est complet. Ceci montre en particulier que cette nouvelle norme sur F n'est pas équivalente à la norme $\|.\|_\infty$, puisque l'espace $(F, \|.\|_\infty)$ n'est pas complet.

Attention, il ne faut pas déduire de ces exemples qu'en dimension infinie les applications linéaires ne sont jamais continues ! Par exemple, pour tout fonction continue f, on a l'inégalité

$$\left|\int_0^1 f(t)dt\right| \leq \|f\|_\infty$$

qu'on déduit immédiatement, par intégration, de l'inégalité

$$-\|f\|_\infty \leq f(t) \leq \|f\|_\infty$$

valable pour tout t dans $[0,1]$, et qui suit directement de la définition de la norme. Cet inégalité s'interprète en disant que l'application linéaire $f \mapsto \int_0^1 f(t)dt$ est continue de $(E, \|.\|_\infty)$ vers $\mathbb{R}$.

Exercice 98.— Montrer que la formule définissant $\|.\|$ fournit bien une norme sur F. (*) Montrer que l'espace vectoriel $(F, \|.\|)$ est complet.

V.3 Exercices

Exercice 99.— Montrer, à l'aide du théorème de Baire, qu'un espace de Banach n'a pas de base dénombrable.

Exercice 100.— Montrer que l'espace vectoriel $C^1([0,1],\mathbb{R})$ est complet lorsqu'on le munit de la norme $N(f) = \|f\|_\infty + \|f'\|_\infty$. *Aide : suivre le schéma général établi à la fin de la preuve de la complétude de $\mathcal{B}(X,Y)$. On pourra se souvenir que lorsque f et g sont deux fonctions continues, g est la dérivée de f si et seulement si, pour tout x, $f(x) = f(0) + \int_0^x g(t)dt$.* Montrer que ce même espace n'est pas complet pour la norme $\|.\|_\infty$.

VI La différentielle

Dans ce chapitre, on étudie la situation suivante : on a une application f définie sur un ouvert Ω d'un espace vectoriel normé E, à valeurs dans un autre espace vectoriel normé F, et un point a de Ω. On veut étudier le comportement de f au voisinage de a. Le principe du calcul différentiel est d'approcher $f(a+h)$, pour des petites valeurs de h, à l'aide d'une application linéaire appelé différentielle de f au point a. Pour aider le lecteur à repérer le sens de chacun des termes des nombreux développements limités intervenant dans ce chapitre, on utilisera des couleurs différentes pour la partie principale et le reste.

Noter que comme Ω est ouvert, $f(a+h)$ est bien défini pour tout h assez petit. Dans toute la suite, lorsqu'on écrit $f(a+h)$, on suppose implicitement que $a+h$ appartient à Ω.

VI.1 Théorie

(a) Dérivée

On commence par le cas où $E = \mathbb{R}$, autrement dit on considère dans cette section une application $f : \Omega \to F$ où Ω est un ouvert de $\mathbb{R}$ et F un espace vectoriel normé. On dit que f est *dérivable* en un point a de Ω si la limite

$$\lim_{h\to 0} \frac{f(a+h)-f(a)}{h}$$

existe ; cette limite, qui est un élément de F, est alors appelée *vecteur dérivé* au point a. Si f est dérivable au point a, on peut écrire

$$f(a+h) = f(a) + hf'(a) + \varepsilon(h) \qquad \text{avec } \lim_{h\to 0} \frac{\varepsilon(h)}{h} = 0.$$

Cette écriture est appelée développement limité de f au point a à l'ordre 1.

On dit aussi que f est une *courbe dans F paramétrée par les points de Ω*, puisqu'elle modélise la trajectoire d'un objet dans l'espace, $f(t)$ étant la position de l'objet au temps t. Le vecteur dérivée $f'(t_0)$ s'interprète alors comme étant le vecteur vitesse de la courbe au temps t_0.

On exprime cette limite en disant que $\varepsilon(h)$ est *négligeable devant h*.

Exercice 101.— En supposant que la dérivée existe, on obtient le développement limité simplement en définissant $\varepsilon(h)$ pour que l'égalité voulue ait lieu, c'est-à-dire en posant $\varepsilon(h) = f(a+h) - (f(a) + hf'(a))$.

1. Vérifier que $\varepsilon(h)$, définie ainsi, est bien négligeable devant h.

2. Vérifier que, réciproquement, l'existence d'un développement limité à l'ordre 1, c'est-à-dire d'une égalité

$$f(a+h) = f(a) + h\overrightarrow{v} + \varepsilon(h) \qquad \text{avec } \lim_{h\to 0} \frac{\varepsilon(h)}{h} = 0$$

pour un certain vecteur $\overrightarrow{v}$ de F, implique que f est dérivable en a et que $f'(a) = \overrightarrow{v}$.

Théorème (des accroissements finis)**.** *Soit $f : [a,b] \to F$ evn. On suppose f continue sur $[a,b]$ et dérivable sur $]a,b[$.*
(1) Supposons qu'il existe M tel que, pour tout $t \in]a,b[$, $\|f't)\| \leq M$. Alors

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

(2) Plus généralement considérons une fonction continue $g : [a,b] \to \mathbb{R}$ dérivable sur $]a,b[$, et supposons que, pour tout $t \in]a,b[$, $\|f't)\| \leq g'(t)$. Alors

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

Si $f(t)$ est la position au temps t d'une voiture, $\|f'(t)\|$ est la vitesse indiquée au compteur au temps t ; $\|f(b) - f(a)\|$ est la distance (à vol d'oiseau) entre le point de départ et le point d'arrivée ; $b-a$ est le temps de parcours, et l'inégalité (1) ne dit rien d'autre que ceci : en roulant pendant un temps T avec une vitesse au compteur qui ne dépasse jamais la valeur V, on ne peut pas se retrouver à une distance supérieure à VT du point de départ.

Recette de preuve.— On s'intéresse d'abord à la situation (1) du théorème. Pour simplifier un peu, on fait l'hypothèse supplémentaire que f est définie et dérivable sur un intervalle ouvert contenant $[a, b]$. On pose

$$M' = \frac{\|f(b) - f(a)\|}{|b - a|}$$

et on veut montrer que $M' \leq M$.

Soit $m = \frac{a+b}{2}$ le milieu du segment $[a, b]$. **Expliquer** Pourquoi l'une des deux distances $\|f(b) - f(m)\|$ et $\|f(m) - f(a)\|$ est supérieure ou égale à $\frac{M'}{2}$. Noter que ceci signifie que la vitesse moyenne l'un des deux demi-segments est supérieure ou égale à M'.

En utilisant cette idée, **construire** par récurrence deux suites $(t_n), (t'_n)$, adjacentes avec, pour tout n, $t'_n = t_n + \frac{1}{2^n}$, et vérifiant

$$\|f(t'_n) - f(t_n)\| \geq \frac{M'}{2^n}.$$

On est alors dans la situation du lemme suivant.

Lemme. *Soit $f :]a, b[\to F$. On considère deux suites adjacentes (t_n) et (t'_n) qui convergent vers une même limite $t_\infty \in]a, b[$. Si, pour tout n, $t_n \leq t_\infty \leq t'_n$, et si f est dérivable en t_∞, alors la suite*

$$\frac{f(t'_n) - f(t_n)}{t'_n - t_n}$$

converge vers $f'(t_\infty)$.

Démontrer le lemme (**où sert** l'hypothèse $t_n \leq t_\infty \leq t'_n$?). **En déduire** que $f'(t_\infty) \geq M'$, puis que $M \geq M'$.

Pour se passer de l'hypothèse supplémentaire faite au début de la preuve, on peut raisonner par l'absurde en supposant $\|f(b) - f(a)\| > M\,|b - a|$. Alors par continuité de f, pour tout intervalle $[\alpha, \beta] \subset]a, b[$ avec α assez proche de a et β assez proche de b, on a encore $\|f(\beta) - f(\alpha)\| > M\,|\beta - \alpha|$. Mais f est définie et dérivable sur un intervalle ouvert contenant $[\alpha, \beta]$: ceci contredit la propriété que nous venons de démontrer.

Adapter la preuve à la situation (2) du théorème.

(Optionnel) **Donner** un contre-exemple au lemme lorsque t_n et t'_n sont tous les deux du même côté de t_∞ (penser à une fonction dérivable mais dont la dérivée n'est pas continue).

Exercice 102.— Soit t un réel positif, et $\gamma : [0, t] \to F$ une application continue, on suppose que γ est dérivable sur $]0, t[$. Soit v un vecteur de F. Montrer l'inégalité

$$\|\gamma(t) - (\gamma(0) + t.v)\| \leq |t| \sup_{s \in [0t]} \|\gamma'(s) - v\|.$$

Indication : appliquer le théorème à la fonction $f(s) = \gamma(s) - sv$.

(b) Différentielle

On considère maintenant le cas général d'une application $f : \Omega \to F$ où Ω est un ouvert de E. On cherche à approcher $f(a + h)$, pour h assez petit, par une expression du type $f(a) + L.h$, où L est un élément de $\mathcal{L}(E, F)$. On dit que f est *différentiable* en un point a s'il existe un $L \in \mathcal{L}(E, F)$ tel que le *reste*

$$\varepsilon(h) = f(a + h) - f(a) - L.h$$

est *négligeable* devant h, ce qui signifie :

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

L'application L est alors appelée *différentielle* de f au point a et notée $Df(a)$.

La notation $D_a f$ est aussi très courante.

Proposition. *L'application L est unique.*

Cette propriété est essentielle, elle permet de parler de *la* différentielle de f en un point donné.

Recette de preuve.— Soient L_1, L_2 deux applications linéaires continues vérifiant toutes les deux la définition de la différentielle, il s'agit de montrer que $L_1 = L_2$. **Montrer** d'abord que $L_1.h - L_2.h$ est négligeable devant h. Maintenant, on fixe un $h \neq 0$, et on veut montrer que $L_1.h = L_2.h$. Pour ça, on va utiliser un argument d'homogénéité, en considèrant les vecteurs $\alpha_n h$, avec (α_n) une suite de réels qui tend vers 0 (on s'approche de a dans une direction fixée donnée par le vecteur h). Que peut-on dire de

$$\frac{\|L_1.\alpha_n h - L_2.\alpha_n h\|}{\|\alpha_n h\|}$$

lorsque α_n tend vers 0 ? **Montrer** que cette quantité peut s'exprimer sans α_n, **en déduire** qu'elle est nulle.

Lorsque f est différentiable en un point a, on peut écrire

$$f(a+h) = f(a) + Df(a).h + \varepsilon(h) \qquad \text{avec } \lim_{h\to 0} \frac{\varepsilon(h)}{\|h\|} = 0.$$

Cette écriture s'appelle *développement limité* de f au point a à l'ordre 1 ; la partie $f(a) + Df(a).h$ s'appelle *approximation affine* de f au point a.

En pratique, pour montrer qu'une application est différentiable, on cherche à établir l'existence d'un développement limité en calculant $f(a+h)$: voir les exemples.

Proposition. *Si f est différentiable au point a elle y est continue.*

Recette de preuve.— **Écrire** le DL à l'ordre 1. **Conclure** en utilisant le fait que $Df(a)$ est, par définition, une application continue.

On dit que f est *différentiable sur Ω* si elle l'est en tout point de Ω.

Attention, on a alors autant d'applications linéaires $Df(a)$ que de points a dans Ω !

Différentielle et dérivée

• Cas $E = \mathbb{R}$

Soit $\gamma : \Omega \to F$ est définie sur un ouvert de $\mathbb{R}$. Si γ est dérivable en a, le développement limité montre que γ est différentiable en a, et sa différentielle est l'application linéaire définie par

$$D\gamma(a).h = h\gamma'(a_0).$$

Pouvez-vous donner le type de chacun des objets apparaissant dans cette formule ?...

Exercice 103.—

1. Vérifier la réciproque : si γ est différentiable en a, alors elle est dérivable.

2. Vérifier que les applications linéaires de $\mathbb{R}$ dans F sont les applications du type $h \mapsto hv$, où v est un vecteur de F. .

• Cas E quelconque, dérivée selon un vecteur

Sous les hypothèses habituelles, choisissons un vecteur non nul e_1. L'application $t \mapsto a + te_1$ est un paramétrage de la droite passant par a et dirigée par e_1. Si l'application $t \mapsto f(a+te_1)$ est dérivable en 0, on dira que f *admet une dérivée au point a selon le vecteur e_1*, et on notera cette dérivée, qui est un vecteur de F, $\frac{\partial f}{\partial e_1}(a)$.

Dessiner la droite dans l'espace E, et le vecteur dérivé dans l'espace F.

Proposition. *Si f est différentiable en a alors elle y admet une dérivée selon tout vecteur non nul de E, et on a, pour tout vecteur $e_1 \neq 0$,*

$$\frac{\partial f}{\partial e_1}(a) := \lim_{t \to 0} \frac{f(a + te_1) - f(a)}{t} = Df(a).e_1.$$

Recette de preuve.— ´Ecrire le développement limité de f en a donné par la différentielle, **l'appliquer** au vecteur te_1, **en déduire** la limite voulue.

Exercice 104.— La fonction $(x, y) \mapsto \frac{xy^2}{x^2+y^2}$ est dérivable en 0 selon tout vecteur, mais n'est pas différentiable. Elle fournit donc un contre-exemple à la réciproque de cette proposition.

(c) Exemples I : La différentielle en coordonnées

• **La vitesse en coordonnées** En pratique, on calcule le vecteur dérivée d'une application $\gamma : [a, b] \to \mathbb{R}^2$ en dérivant coordonnées par coordonnées.

Exercice 105.— Calculer $\gamma'(t)$ pour $\gamma(t) = (x(t), y(t)) = (t, t^2)$. Faire un dessin !

• **Différentielle et développements limités en dimension un**

Exercice 106.— Rappeler le DL de la fonciton Sinus à l'ordre1 en 0. En déduire que la fonction $(x, y) \mapsto \sin(x + 2y)$ est différentiable en 0 et donner sa différentielle.

• **Différentielle et dérivées partielles** En pratique, lorsqu'on a une fonction de $\mathbb{R}^2$ dans $\mathbb{R}$ comme $f(x, y) = \sin(x + 2y)$, on calcule ses *dérivées partielles*

$$\frac{\partial f}{\partial x}(x, y) = \cos(x + 2y), \quad \frac{\partial f}{\partial y}(x, y) = 2\cos(x + 2y).$$

La première, par exemple, s'obtient en imaginant que y est fixé, et en dérivant la fonction $x \mapsto f(x, y)$ (d'une seule variable !) comme si y était une constante. En écrivant $f(x + t, y) = f((x, y) + t(1, 0))$, on voit que cette dérivée partielle n'est rien d'autre que la dérivée selon le premier vecteur $e_1 = (1, 0)$ de la base canonique :

$$\frac{\partial f}{\partial x}(x, y) = \frac{\partial f}{\partial e_1}(x, y).$$

En décomposant un vecteur (h, k) dans la base canonique, on voit que la différentielle s'écrit

$$Df(x, y).(h, k) = Df(x, y).(he_1 + ke_2) = hDf(x, y)(e_1) + kDf(x, y)(e_2) = h\frac{\partial f}{\partial x}(x, y) + k\frac{\partial f}{\partial y}(x, y).$$

Et le développement limité à l'ordre 1 devient

$$f(x+h, y+k) = f(x, y) + h\frac{\partial f}{\partial x}(x, y) + k\frac{\partial f}{\partial y}(x, y) + \varepsilon(h, k) \quad \text{avec} \quad \lim_{(h,k)\to(0,0)} \frac{\varepsilon(h, k)}{\|(h, k)\|} = 0.$$

• **Gradient**

Dans l'exemple précédent, le terme $h\frac{\partial f}{\partial x}(x,y) + k\frac{\partial f}{\partial y}(x,y)$ peut s'interpréter comme le produit scalaire (usuel) du vecteur (h,k) avec le vecteur "gradient" dont les coordonnées sont les dérivées partielles de f. Plus généralement, si $f : \Omega \to \mathbb{R}$ est définie sur un ouvert de $\mathbb{R}^N$, et si elle est différentiable en un point a, le vecteur *gradient* de f au point a est

$$\nabla_a f = \mathrm{grad} f(a) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_N} \end{pmatrix}$$

et on a, pour tout vecteur h,

$$Df(a).h = \langle \nabla_a f, h \rangle.$$

Le développement limité en a s'écrit alors

$$f(a+h) = f(a) + \langle \nabla_a f, h \rangle + \varepsilon(h).$$

Exercice 107.— Le gradient permet de répondre à la question : "à partir du point a, dans quelle direction est-ce que f augmente le plus ?" L'accroissement de la fonction f lorsqu'on se déplace du point a au point $a+h$ est donné par le terme $Df(a).h = \langle \nabla_a f, h \rangle$ (au moins à l'ordre 1, c'est-à-dire en négligeant le reste $\varepsilon(h)$ du développement limité). Pour préciser la question, disons qu'on s'autorise à se déplacer, à partir du point a, d'une distance c fixée dans l'espace $\mathbb{R}^N$ qui est muni de la norme euclidienne. La question devient alors : parmi tous les vecteurs h tels que $\|h\| = c$, quel est celui pour lequel le nombre $\langle \nabla_a f, h \rangle$ est le plus grand ?

1. Que donne l'inégalité de Cauchy-Schwartz ?

2. Répondre à la question à l'aide des cas d'égalité dans cette inégalité.

3. A quelle condition l'accroissement $\langle \nabla_a f, h \rangle$ est-il nul ? Positif ? Négatif ? Faire un dessin.

• **En Physique** Dans un livre de physique traitant des gaz, on peut lire : *La relation entre la pression P, la température T et le volume V pour un gaz parfait est donnée par $P = nR\frac{T}{V}$. On a donc*

$$dP = \frac{\partial P}{\partial T}dT + \frac{\partial P}{\partial V}dV = nR\frac{1}{V}dT + nR\frac{-T}{V^2}dV.$$

Le mathématicien interprète P comme une fonction de $\mathbb{R}^2$ dans $\mathbb{R}$ (ou fonction de deux variables), $(T,V) \mapsto P(T,V)$. On peut interpréter cette relation comme une expression de la différentielle au point (P,V) calculé sur le vecteur (dT,dV), "dP" étant alors un raccourci pour $DP(T,V).(dT,dV)$. On peut aussi y voir un développement limité où le reste est implicite. En tout cas, les notations dP, dT, dV ont le mérite de nous rappeler qu'il faut y voir des "petits accroissements" des quantité P, V, T, le développement limité ne nous apportant une information seulement lorsque les accroissements dT, dV des variables T et V tendent vers 0.

• **Matrice jacobienne** Pour une fonction de deux variable, la différentielle peut aussi s'écrire matriciellement,

$$Df(x,y).(h,k) = h\frac{\partial f}{\partial x}(x,y) + k\frac{\partial f}{\partial y}(x,y) = \begin{pmatrix} \frac{\partial f}{\partial x}(x,y) & \frac{\partial f}{\partial x}(x,y) \end{pmatrix} . \begin{pmatrix} h \\ k \end{pmatrix},$$

ce qui s'interprète en disant que la matrice $\left(\frac{\partial f}{\partial x}(x,y) \quad \frac{\partial f}{\partial x}(x,y)\right)$ est la matrice de l'application linéaire $Df(x,y) : \mathbb{R}^2 \to \mathbb{R}$ dans la base canonique ; on l'appelle *matrice jacobienne* de f au point (x,y). Regardons un autre exemple, celui des coordonnées polaires, c'est-à-dire de la fonction

$$\Phi : (r,\theta) \mapsto (X(r,\theta), Y(r,\theta)) = (r\cos(\theta), r\sin(\theta))$$

définie sur l'ouvert $\Omega = \{(r,\theta), r > 0, \theta \in \mathbb{R}\}$ de $\mathbb{R}^2$. Admettons provisoirement qu'elle est différentieble sur Ω. Sa différentielle s'écrit alors

$$D\Phi(r,\theta).(dr,d\theta) = \begin{pmatrix} DX(r,\theta).(dr,d\theta) \\ DY(r,\theta).(dr,d\theta) \end{pmatrix}$$

D'où

$$D\Phi(r,\theta).(dr,d\theta) = \begin{pmatrix} dr\frac{\partial X}{\partial r}(r,\theta) + d\theta\frac{\partial X}{\partial \theta}(r,\theta) \\ dr\frac{\partial Y}{\partial r}(r,\theta) + d\theta\frac{\partial Y}{\partial \theta}(r,\theta) \end{pmatrix} = \begin{pmatrix} \frac{\partial X}{\partial r}(r,\theta) & \frac{\partial X}{\partial \theta}(r,\theta) \\ \frac{\partial Y}{\partial r}(r,\theta) & \frac{\partial Y}{\partial \theta}(r,\theta) \end{pmatrix} \cdot \begin{pmatrix} dr \\ d\theta \end{pmatrix}.$$

et sa matrice jacobienne est donc

$$\begin{pmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{pmatrix}.$$

Plus généralement, lorsque $F = \mathbb{R}^{N'}$, se donner une fonction de Ω dans F équivaut à se donner N' fonction $f_1, \ldots, f_{N'}$ de E dans $\mathbb{R}$. La matrice jacobienne de l'application f en a est par définition la matrice de l'application linéaire $Df(a)$ dans les bases canoniques ; elle sécrit

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \ldots & \frac{\partial f_1}{\partial x_N}(x) \\ & \ldots & \\ \frac{\partial f_{N'}}{\partial x_1}(x) & \ldots & \frac{\partial f_{N'}}{\partial x_N}(x) \end{pmatrix}$$

• En dimension finie toutes les normes sont équivalentes, et en regardant les définitions on en déduit immédiatement que la notion de différentiabilité est indépendante des normes choisies.

(d) Exemples II : la différentielle sans coordonnées

• **Applications affines** Soit $L \in \mathcal{L}(E,F)$, et v_0 un vecteur de F. L'application $f : x \mapsto v_0 + L.x$ est appelée *application affine*. Pour a et h dans E on peut écrire

$$f(a+h) = v_0 + L.(a+h) = v_0 + L.a + L.h = f(a) + L.h$$

Ce qui est un développement limité avec un reste nul. Notre application affine est donc différentiable en tout point a de E, et sa différentielle est $Df(a) = L$. Le cas où L est l'identité correspond à une application $T_b : x \mapsto b + x$ appelée *translation de vecteur b*, sa différentielle est donc l'identité.

Pour les applications affines, la différentielle est la même en tout point, ce qui est une propriété très spéciale. En particulier, une application linéaire est sa propre différentielle en tout point.

• **Un exemple en dimension infinie**

Exercice 108.— Calculer la différentielle de $f \mapsto e^f$, de $C(X, \mathbb{R})$ dans lui-même.

• **Applications bilinéaires continues.** Soient E_1, E_2 deux evn, et $E = E_1 \times E_2$, qui est un evn pour la norme définie par $\|(h_1, h_2)\| = \max(\|h_1\|, \|h_1\|)$. Une application $B : E_1 \times E_2 \to F$ est *bilinéaire* si elle vérifie les formules

$$B(\alpha x_1 + \beta y_1, x_2) = \alpha B(x_1, x_2) + \beta B(y_1, x_2),$$

$$B(x_1, \alpha x_2 + \beta y_2) = \alpha B(x_1, x_2) + \beta B(x_1, y_2).$$

Autrement dit, pour tout $x_1 \in E_1$ et $x_2 \in E_2$, les applications $x \mapsto B(x, x_2)$ et $x \mapsto B(x_1, x)$ sont linéaires.

Elle est alors continue si et seulement s'il existe une constante C telle que, pour tout h_1, h_2,

$$\|B(h_1, h_2)\| \leq C \|h_1\| \|h_1\| \qquad (\star).$$

De plus, en dimension finie, toutes les applications bilinéaires sont continues (ces affirmations se démontrent comme pour les applications linéaires, et sont laissées en exercice). Montrons qu'une application bilinéaire continue est différentiable en tout point $a = (a_1, a_2)$ de E : pour (h_1, h_2) dans E, on développe $B(a_1 + h_1, a_2 + h_2) =$ en utilisant la bilinéarité, et on reconnait, à l'aide de la propriété de continuité $(\star)$, un développement limité. La différentielle de B au point a est l'application linéaire $(h_1, h_2) \mapsto B(a_1, h_2) + B(h_1, a_2)$.

Ceci est un exemple de démarche typique pour montrer qu'une application est différentiable en un point a.

Exemple d'applications bilinéaires : le produit de matrices (défini sur $E = \mathcal{M}_{n,p}(\mathbb{R}) \times \mathcal{M}_{p,q}(\mathbb{R})$) ; son équivalent, la composition des applications linéaires continues (définie sur $E = \mathcal{L}(G, H) \times \mathcal{L}(F, G)$) ; en dimension infinie, l'application $(f_1, f_2) \mapsto f_1 f_2$ qui associe à deux application bornée de X dans $\mathbb{R}$ leur produit (ici $E = \mathcal{B}(X, \mathbb{R}) \times \mathcal{B}(X, \mathbb{R})$).

Exercice 109.— Montrer que l'application
Phi : $f \mapsto f^2$, de $(C([0,1], \mathbb{R}), \|.\|_\infty$ dans lui-même, est différentiable en tout f_0, et donner sa différentielle. *Aide : calculer* $\Phi(f_0 + h)$.

La seule difficulté de cet exercice est d'ordre psychologique : il faut accepter de travailler avec une application Φ défini sur un espace dont les points sont des fonctions.

• **Inversion de matrices.** L'application Inv : $L \mapsto L^{-1}$ est différentiable, et sa différentielle en un point L_0 est donnée par

$$D(\mathrm{Inv})(L_0).H = -L_0^{-1} \circ H \circ L_0^{-1}.$$

Cette formule fait l'objet de l'exercice suivant.

Cet exemple sert pour la notion de C^k-difféomorphisme.

Exercice 110.—

1. Rappeler la formule donnant l'inverse de $\mathrm{Id} + H$, en déduire le résultat lorsque $L_0 = \mathrm{Id}$.

2. Exprimer l'inverse de $L_0 + H$ à partir de l'inverse de Id + quelque chose de petit lorsque H est petit. (Indication : factoriser L_0). En déduire une formule pour $(L_0 + H)^{-1}$ à l'aide d'une série, puis démontrer le résultat en vous insiprant de la première question.

3. Autre méthode : à partir de la formule précédente, exprimer $\mathrm{Inv}(L_0 + H)$ comme composée de trois applications, puis appliquer le théorème de composition plus bas.

(e) Opérations

• **Somme.** Soient f_1, f_2 deux applications de Ω dans F. Si toutes les deux sont différentiables au point a, alors f_1+f_2 l'est aussi, et on a $D(f_1+f_2)(a) = Df_1(a)+Df_2(a)$. La preuve est immédiate.

• **Composition**

Proposition. *Soit $f : \Omega \subset E \to F$, $g : \Omega' \subset F \to G$, a un point de Ω tel que le point $b = f(a)$ est dans Ω'. Si f est différentiable en a et g différentiable en d alors $g \circ f$ est différentiable en a et on a*

$$D(g \circ f)(a) = Dg(b) \circ Df(a).$$

Recette de preuve.— Remarquons d'abord que la fonction $g \circ f$ est définie sur $\Omega \cap f^{-1}(\Omega')$, par continuité de f au point a. En particulier elle est définie sur un ouvert de E contenant a (ceci est nécessaire à la différentiabilité, voir la définition).

En négligeant les restes dans les DLs de f et g on a une heuristique de preuve. Pour un argument précis, on écrit les DL à l'ordre 1 de f en a et de g en $b = f(a)$,

$$f(a+h) = f(a) + Df(a).h + \varepsilon_1(h)$$

$$g(a'+k) = g(a') + Dg(b).k + \varepsilon_2(k).$$

On pose $K(h) = Df(a).h+\varepsilon_1(h)$, on remarque que $K(h)$ tend vers 0 lorsque h tend vers 0 (c'est exactement la continuité de f en a). On reporte le premier DL dans le second. Ceci donne un DL de $g \circ f$ en a, à condition de savoir montrer que le reste

$$\varepsilon_3(h) = Dg(b).\varepsilon_1(h) + \varepsilon_2(K(h))$$

est négligeable devant h ; ceci va suivre du fait que les restes $\varepsilon_1(h), \varepsilon_2(k)$ sont négligeables respectivement devant h et k, et que le quotient $\|K(h)\| / \|h\|$ est borné.

Exercice 111.— Généraliser en écrivant la formule donnant la différentielle en un point a de la composée des 3 applications $f_3 \circ f_2 \circ f_1$. *Aide : on appliquera deux fois de suite la formule de composition en écrivant par exemple* $f_3 \circ f_2 \circ f_1 = (f_3 \circ f_2) \circ f_1$. Généraliser à la composée de n applications.

Exercice 112.— Que vaut la différentielle de $T_b \circ f$ lorsque T_b est la translation de F de vecteur b ? de $f \circ T_a$ lorsque T_a est la translation de E de vecteur a ? de $L \circ f$, lorsque L est une application linéaire ? De $f \circ L$?

Ce type de différentielles composées sera souvent utilisé dans les démonstrations qui suivent.

Exercice 113.—

1. Soit I un intervalle ouvert de $\mathbb{R}$, et $\gamma : I \to E$ une application dérivable. Soit $t_0 \in I$. Rappeler le lien entre la différentielle $D\gamma(t_0)$ et le vecteur vitesse $\gamma'(t_0)$.

2. On se donne de plus une application $f : E \to F$ différentiable en $\gamma(t_0)$. Appliquer la formule de composition à $f \circ \gamma$, en déduire le vecteur vitesse $(f \circ \gamma)'(t_0)$ en fonction de $\gamma'(t_0)$.

3. Application : étant donnée des points a, b et un vecteur v, expliciter la dérivée de $t \mapsto f((1-t)a + tb)$, de $t \mapsto f(a + tv)$ (et retrouver ainsi le lien entre la différentielle et la dérivée selon un vecteur).

Exercice 114.— Décomposer l'application $(x_1, x_2) \mapsto \sin(x_1 + 2x_2)$ comme la composée de deux fonctions différentiable. En déduire qu'elle est différentiable sur $\mathbb{R}^2$, et exprimer sa différentielle au point $(4, 5)$.

• **Produit** Différentielle de $B(f_1, f_2)$ lorsque B est bilinéaire. Produit de deux applications différentiables à valeurs dans $F = \mathbb{R}$. (Plus généralement, produit d'une application à valeur réelle et d'une application à valeur dans un evn F quelconque). Produit scalaire.

• **Quotient** Différentielle de $1/f$ lorsque f est une application différentiable à valeurs dans $F = \mathbb{R}$ et ne s'annulant pas.

• **Inverse**

Proposition. *Soit $f : \Omega_1 \to \Omega_2$ un homéomorphisme entre un ouvert de E et un ouvert de F. Si f est différentiable en a et si $Df(a)$ est inversible, alors f^{-1} est différentiable en $b := f(a)$ et $D(f^{-1})(b) = (Df(a))^{-1}$.*

Cet énoncé sera utilisé plus loin dans la preuve du théorème d'inversion local ; on verra aussi que l'hypothèse "Ω_2 est un ouvert" est superflue, puisque $f(a)$ est automatiquement dans l'intérieur de l'image de Ω_1.

Recette de preuve.— Ce qui est difficile c'est de montrer que la réciproque est différentiable ; ensuite, la différentielle s'obtient par la formule de composition. Pour la différentiabilité, on traite d'abord le cas où $E = F$, $a = f(a) = 0$ et $Df(a) = Id$. Dans ce cas, il s'agit de montrer que f^{-1} est différentiable en 0 et que sa différentielle en 0 est l'identité. Pour ceci, on écrit, pour tout h assez petit,

$$f(h) = h + \varepsilon_1(h).$$

Soit k assez petit. En appliquant cette égalité à $h = f^{-1}(k)$, on obtient

$$f^{-1}(k) = k - \varepsilon_1(f^{-1}(k))$$

et il s'agit de montrer que le reste $\varepsilon_2(k) := -\varepsilon_1(f^{-1}(k))$ est négligeable devant k, sachant que le reste $\varepsilon_1(h)$ est négligeable devant h. En utilisant l'hypothèse $Df(a) = Id$, **montrer** que, au voisinage de 0, la norme de $f(h)$ est comparable à celle de h (utiliser $\varepsilon = 1/2$), ce qui revient à dire que la norme de $f^{-1}(k)$ est comparable à celle de k. Puisque $\varepsilon_1(h)/h$ tend vers zéro, **en déduire** que $\varepsilon_2(k)/k$ tend aussi vers zéro, ce qu'on voulait.

On déduit le cas général de ce cas particulier : on translate à la source et au but pour se ramener à $a = f(a) = 0$, puis on compose par $Df(a)^{-1}$ pour se ramener à une différentielle égale à l'identité. Ceci revient à considérer

$$g = Df(a)^{-1} \circ T_{-f(a)} \circ f \circ T_a$$

qui est un homéomorphisme comme composé d'homéomorphismes. **En déduire** f^{-1} en fonction de g^{-1}, puis le fait que f soit différentiable en a par composition.

La composition en coordonnées ("Chain rule")

Exercice 115.— Un mobile se déplace au cours du temps sur une surface dont la hauteur Z est donnée en fonction de la longitude x et la latitude y par la formule $Z(x, y) = xe^{-(x^2+y^2)}$; sa longitude et sa latitude varient au cours du temps selon la formule $\gamma(t) = (x(t), y(t)) = (t, t^2)$. On veut calculer la dérivée de l'altitude en fonction du temps.

1. Exprimer l'altitude en fonction du temps comme composée de deux applications.

2. Calculer la matrice jacobienne de la fonction Z en un point (x, y).

3. En déduire la dérivée de l'altitude par rapport au temps.
4. Vérifier par un calcul direct.

La formule de dérivation composée obtenue dans cet exercice s'écrit (en oubliant les variables pour plus de lisibilité)

$$(Z \circ \gamma)'(t) = \frac{d(Z \circ \gamma)}{dt} = \frac{\partial Z}{\partial x}\frac{dx}{dt} + \frac{\partial Z}{\partial y}\frac{dy}{dt}.$$

Exercice 116.— Soit $f : \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto f(x, y)$. On veut calculer les dérivées partielles de f par rapport aux coordonnées polaires r et θ. On écrit souvent $f(r, \theta)$ pour "l'application f dans les coordonnées polaires", mais en réalité il s'agit d'une fonction différente, notons-là $\hat{f}$, qui est obtenu en pré-composant f par l'application Φ (définie plus haut) qui passe des coordonnées polaires aux coordonnées cartésiennes : $\hat{f} = f \circ \Phi$.

En déduire la relation entre les différentielles de $\hat{f}$, f et Φ. Traduire cette relation sur les matrices jacobiennes. En déduire la formula donnant les dérivées partielles de $\hat{f}$ par rapport à r et θ en fonction des dérivées partielles de f par rapport à x et y. Appliquer à la fonction $f : (x, y) \mapsto \sin(x + 2y)$.

La matrice jacobienne de la composée est le produit des matrices jacobienne, ce qui permet d'écrire les dérivées partielles de la fonction composée $h = g \circ f$ à partir des dérivées partielles des fonctions f et g. Ceci est illustré par l'exercice précédent.

(f) Applications de classe C^1

Théorème (Inégalité des accroissements finis, cas E quelconque)**.** *Soit $f : \Omega \to F$ différentiable sur Ω. On considère deux points a, b de Ω, et on suppose que $[a, b] \subset \Omega$. Alors on a :*

$$\|f(b) - f(a)\| \leq \|b - a\| \sup_{x \in [ab]} \|Df(x)\|.$$

Préciser, pour chacune des trois normes apparaissant dans l'énoncé, l'espace sur lequel elle est définie.

Recette de preuve.— Sous les hypothèse du théorème, notons $M = \sup_{x \in [ab]} \|Df(x)\|\}$. **Relire** la définition de la norme sur $\mathcal{L}(E, F)$, et **en déduire** une majoration de $\|Df(x).h\|$ utilisant M et $\|h\|$ (pour tous $x \in [ab]$ et tout h dans E).

La preuve du théorème consiste à se ramener au cas où $E = \mathbb{R}$ pour pouvoir appliquer l'inégalité des accroissements finis qu'on a démontré plus haut. Pour ceci, on paramètre le segment $[ab]$ par l'application $\gamma : t \mapsto (1 - t)a + tb$. **Calculer** $(f \circ \gamma)'(t)$. **Appliquer** l'inégalité des accroissements finis à $f \circ \gamma$. **Conclure**.

Que peut-on dire d'une fonction f dont la différentielle est nulle en tout point de Ω ?...

Corollaire. *Soit $f : \Omega \to F$ une fonction différentiable sur un ouvert connexe Ω. Supposons que $Df(x) = 0$ pour tout $x \in \Omega$. Alors f est constante sur Ω.*

Recette de preuve.— **Montrer** que, sous l'hypothèse, f est constante sur toute boule $B(x, \varepsilon)$ incluse dans Ω. On dit que f est localement constante.

Soit maintenant c un nombre, et $\Omega_c = \{x \in \Omega \mid f(x) = c\}$ (cet ensemble est appelé *ligne de niveau c*). **Expliquer** pourquoi Ω_c est fermée dans Ω. **Montrer** qu'elle est aussi ouverte (utiliser l'inégalité des accroissements finis). **Conclure**.

Théorème. *On suppose ici que $E = \mathbb{R}^N$. Soit $f : \Omega \to F$. Supposons*
- *que f admet des dérivées partielles en tout point de Ω,*
- *que ces dérivées partielles sont des fonctions continues sur Ω,*

alors f est différentiable sur Ω.

Ce théorème donne un critère pratique pour la différentiabilité, en particulier lorsque f est donné par une formule : par exemple, la fonction $(x, y) \mapsto \sin(x+2y)$ admet des dérivées partielles en tout point (x, y), et celles-ci sont des fonctions continues de (x, y), donc elle est différentiable sur $\mathbb{R}^2$ (nous avons vu deux arguments différents plus haut).

Recette de preuve.— Pour simplifier on travaille en deux variables seulement. Si $Df(a)$ existe alors elle sera donnée par les dérivée partielle, il s'agit donc d'estimer le reste d'un développement limité. On décompose le reste en séparant l'accroissement des deux coordonnées. Il faut alors voir que les restes

$$\varepsilon_1(h_1, h_2) = f(a_1 + h_1, a_2) - \left(f(a_1, a_2) - h_1 \frac{\partial f}{\partial x_1}(a_1, a_2) \right)$$

et

$$\varepsilon_2(h_1, h_2) = f(a_1 + h_1, a_2 + h_2) - \left(f(a_1 + h_1, a_2) - h_2 \frac{\partial f}{\partial x_2}(a_1, a_2) \right)$$

sont négligeables devant h. On estime le premier reste en appliquant au vecteur $v = \frac{\partial f}{\partial x_1}(a)$ et à la fonction $\gamma : s \mapsto f(a+s)$ la version fine de l'inégalité des accroissements finies (voir l'exercice 102 à la fin de la section a), en utilisant que la dérivée partielle sur le segment $[(a_1, a_2), (a_1 + h_1, a_2)]$ est proche de la dérivée partielle en (a_1, a_2) lorsque h_1 est petit (c'est ici qu'on utilise la continuité des dérivées partielles). On procède de façon analogue pour le second reste, à l'aide du vecteur $v = \frac{\partial f}{\partial x_2}(a)$ et de la fonction $\gamma : s \mapsto f(a + (h_1, s))$.

Adapter la preuve pour une fonction de N variables.

Lorsque f est différentiable en tout point de Ω, on peut considérer l'application

$$x \mapsto Df(x)$$

qui va de Ω dans l'espace vectoriel normé $\mathcal{L}(E, F)$. Si elle est continue, on dira que f est *de classe C^1 sur Ω*. On montre immédiatement qu'une somme, composition, produit d'applications de classe C^1 est de classe C^1.

Corollaire. *Lorsque $E = \mathbb{R}^N$, f est de classe C^1 sur Ω si et seulement si les dérivées partielles de f existent et sont continues sur Ω.*

En particulier, lorsque $F = \mathbb{R}^{N'}$, $Df(x)$ dépend continûment de x si et seulement si les coefficients de la matrice jacobienne de f dépendent continûment de x.

Ce résultat permet de montrer très rapidement la différentiabilité d'une application donnée par une formule, comme $\sin(x + 2y)$, à l'aide des dérivées partielles.

Recette de preuve.— Lorsque les dérivées partielles de f existent et sont continues sur Ω, on a déjà vu que f est différentiable sur Ω, il reste à voir que $Df(x)$ dépend continûment de x. Pour cela, **montrer** l'inégalité

$$\|Df(x) - Df(y)\| \leq \max_{i=1,\dots,N} \left\| \frac{\partial f}{\partial x_i}(x) - \frac{\partial f}{\partial x_i}(y) \right\|$$

(on a muni E de la norme $\|.\|_1$).

Exercice 117.—

1. L'inégalité des accroissements finis n'a d'intérêt que si la différentielle est bornée sur $[ab]$. Expiquer pourquoi c'est le cas lorsque f est de classe C^1.

2. En supposant Ω convexe et $\|Df(x)\| \leq M$ pour tout x de Ω, montrer que f est M-lipschitzienne sur Ω.

3. Soit $f : \Omega \to F$ de classe C^1, et K un compact convexe de Ω (par exemple une boule fermée). Montrer que f est lipschitzienne sur K. Plus difficile : montrer que c'est encore vrai lorsque K est un compact connexe.

Exercice 118.—(majoration du reste du développement limité) Sous les hypothèses de l'inégalité des accroissements finis, montrer que

$$\|f(b) - f(a) - Df(a)(b-a)\| \leq \|b-a\| \sup_{x\in[ab]} \|Df(x) - Df(a)\|.$$

Aide : appliquer l'inégalité à la fonction $\hat{f}(x) = f(x) - Df(a)(x)$.

VI.2 Commentaires

(a) La différentielle

On présente généralement la dérivée d'une fonction f en un point a comme un nombre qui mesure le taux de variation de $f(x)$ lorsque x passe au point a. Le but de cette section [1] est de mettre en avant une présentation un peu différente, plus générale. C'est l'idée de la différentiation vue comme une *approximation linéaire* [2], idée qui est au centre d'une grande partie des mathématiques actuelles.

Intuitivement, dire que $f'(a) = m$ revient à dire que si on regarde à travers un puissant microscope le graphe de f dans une petite région autour du point $(a, f(a))$, ce que l'on voit est presque exactement une ligne droite de pente m. En d'autres termes, au voisinage de a, la fonction f est approximativement linéaire. On peut même écrire une formule pour la fonction linéaire g qui approxime f :

$$g(x) = f(a) + m(x-a).$$

Son graphe est la droite de pente m passant par le point $(a, f(a))$. Une façon, un peu plus claire, consiste à écrire

$$g(a+h) = f(a) + mh.$$

Dire que g approxime f au voisinage de a revient à dire que $f(a+h)$ est approximativement égal à $f(a) + mh$ lorsque h est petit.

1. Ce qui suit est adapté du merveilleux livre *The Princeton Companion to Mathematics*.
2. En toute rigueur il faudrait plutôt écrire ici *approximation affine* (une application affine est juste la somme d'une application linéaire et d'une constante).

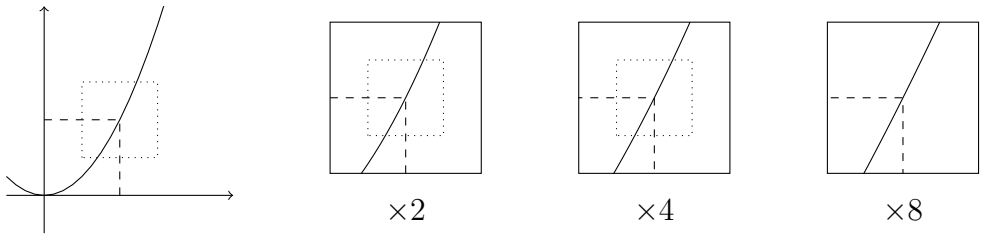


FIGURE 1 – Zooms successifs sur la fonction $x \mapsto x^2$

Ici, il faut faire un peu attention : après tout, si f est continue au point a alors quand h est petit, $f(a+h)$ sera proche de $f(a)$ et mh sera très petit, de sorte que $f(a+h)$ sera proche de $f(a)+mh$. Cette façon de voir semble marcher pour n'importe quelle valeur de m, et pourtant ce que nous nous voulons dire est très spécifique à la valeur $m = f'(a)$. Ce qui n'arrive qu'avec cette valeur de m, c'est que $f(a+h)$ est non seulement proche de $f(a)+mh$, mais tellement proche que la différence $\varepsilon(h) = f(a+h) - f(a) - mh$ est petite *comparée à h*. Autrement dit,

$$\frac{\varepsilon(h)}{h} \text{ tend vers 0 lorsque } h \text{ tend vers 0.}$$

Cette façon de voir peut se généraliser. Les fonctions qui apparaissent en mathématiques, et aussi dans les autres sciences, en ingénierie, en économie, *etc.*, sont souvent des fonctions de plusieurs variables, et peuvent donc être vues comme des fonctions définies sur un espace vectoriel de dimension strictement plus grande que 1. On peut alors se demander si, dans un petit voisinage d'un point, on peut les approcher par des applications linéaires. Lorsque c'est possible, cette approximation est extrêmement utile : une fonction générale peut *a priori* avoir un comportement très compliqué, mais si on peut l'approcher par une fonction linéaire, alors son comportement sera beaucoup plus facile à comprendre, au moins dans de petites régions de l'espace de dimension n. Dans ce cas on peut utiliser toute la machinerie de l'algèbre linéaire et des matrices, qui permet de faire des calculs, surtout si on dispose de l'aide d'un ordinateur.

Imaginez, par exemple, un météorologue s'intéressant à la façon dont la vitesse et la direction du vent varient d'un endroit à l'autre dans une certaine région de l'espace au-dessus de la surface de la Terre. À chaque point (x, y, z) de cette région (x et y représentent par exemple la latitude et la longitude et z l'altitude) on peut associer un vecteur (u, v, w) représentant la vitesse du vent en ce point : u, v, w sont les composantes du vecteur vitesse dans les directions x, y, z.

Déplaçons maintenant très légèrement le point (x, y, z) en choisissant trois petits nombres h, k, l et en considérant le point $(x+h, y+k, z+l)$. En ce nouveau point, nous nous attendons à ce que la vitesse du vent soit différente mais assez proche de celle au point (x, y, z) ; nous l'écrivons donc $(u+p, v+q, w+r)$. Comment la petite variation (p, q, r) du vecteur-vent dépend-elle de la petite variation (h, k, l) de la position du point ? En supposant que le vent n'est pas trop turbulent et que h,k, l sont suffisamment petits, nous nous attendons à ce que cette dépendance soit approximativement linéaire : c'est la façon dont la nature semble fonctionner. Autrement dit, nous nous attendons à ce qu'il existe une application linéaire T telle que (p, q, r) vale approximativement $T(h, k, l)$ lorsque h, k, l sont assez petits. Chacun des nombres p, q, r dépend de chacun des nombres h, k, l, et

il nous faut donc 9 nombres pour exprimer cette dépendance linéaire. Sous forme matricielle, elle s'écrit

$$\begin{pmatrix} p \\ q \\ r \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} . \begin{pmatrix} h \\ k \\ l \end{pmatrix} .$$

Chaque entrée a_{ij} de la matrice exprime comment l'un des trois nombres p, q, r dépend de l'un des trois nombres h, k, l. Par exemple, si l'on fixe x et z, ce qui revient à poser $h = l = 0$, on obtient $p = a_{12}k$: le coefficient a_{12} représente donc le taux de variation de u lorsque y change. Techniquement, a_{12} est la dérivée partielle $\partial u/\partial y$ au point (x, y, z).

Ceci nous dit comment calculer la matrice de T (appelée matrice jacobienne), mais d'un point de vue conceptuel il vaut mieux éviter les coordonnées et raisonner de façon purement vectorielle. En notant $\mathbf{x}$ pour (x, y, z), $\mathbf{u}(\mathbf{x})$ pour (u, v, w) et $\mathbf{h}$ pour (h, k, l), tout ceci peut se résumer en écrivant la relation

$$\mathbf{u}(\mathbf{x} + \mathbf{h}) = \mathbf{u}(\mathbf{x}) + T(\mathbf{h}) + \varepsilon(\mathbf{h}),$$

où $\varepsilon(\mathbf{h})$ est petit comparé à $\mathbf{h}$. Ceci nous dit que si nous ajoutons un petit vecteur $\mathbf{h}$ à $\mathbf{x}$, la variation de $\mathbf{u}(\mathbf{x})$ sera approximativement $T(\mathbf{h})$. Cette formule est bien sûr très similaire à la formule du début, $f(a + h) = f(a) + mh + \varepsilon(h)$.

Plus généralement, soit $\mathbf{u}$ une application de $\mathbb{R}^n$ dans $\mathbb{R}^m$. On dit que cette application est différentiable s'il existe une application linéaire $T : \mathbb{R}^n \to \mathbb{R}^m$ telle que, à nouveau, la formule $\mathbf{u}(\mathbf{x} + \mathbf{h}) = \mathbf{u}(\mathbf{x}) + T(\mathbf{h}) + \varepsilon(\mathbf{h})$ soit vérifiée avec $\varepsilon(\mathbf{h})$ petit devant $\mathbf{h}$. L'application linéaire T est appelée différentielle de $\mathbf{u}$ au point $\mathbf{x}$.

Le cas $m = 1$ est un cas particulier important. Si $f : \mathbb{R}^n \to \mathbb{R}$ est différentiable au point $\mathbf{x}$, alors la différentielle de f est une application linéaire T de $\mathbb{R}^n$ dans $\mathbb{R}$. La matrice de T est un vecteur ligne de taille n, qui est souvent noté $\nabla f(x)$ et appelé gradient de f au point $\mathbf{x}$. Ce vecteur pointe dans la direction dans laquelle f augmente le plus vite, et sa longueur est égale au taux de variation dans cette direction.

(b) Gradient et optimisation

La notion de vecteur gradient est à la base d'algorithmes de recherche du minimum d'une fonction, comme l'algorithme du gradient, cf Wikipedia. La méthode fournit une bonne approximation d'un minimum local.

(c) Différence entre continu et classe C^1

On a vu qu'il existe des courbes *continues* qui remplissent le carré. Ceci ne peut pas arriver pour une courbe de classe C^1. En fait, on peut montrer que l'image d'une courbe de classe C^1 dans $\mathbb{R}^n$, avec $n \geq 2$, est de mesure de Lebesgue nulle. C'est une conséquence du lemme de Sard qui dit que l'image des points où la vitesse s'annule est de mesure nulle, et du fait qu'au voisinage d'un point en lequel la vitesse ne s'annule pas, la courbe est une sous-variété (voir plus loin...).

VI.3 Exercices

Exercice 119.— Relier chaque objet à son type.

Objets : $D(f)$; $D(f)(a)$; $D(f)(a).\overrightarrow{h}$.

Types : F ; $\mathcal{L}(E,F)$; applications de Ω vers $\mathcal{L}(E,F)$.

Bonus : Quel type d'objet est le "D" de $D(f)$?

Exercice 120.— Calculer la différentielle de l'application qui à un élément de $C^1([0,1],\mathbb{R}^3$ associe sa longueur, donnée par la formule

$$\text{Longueur}(\gamma) = \left(\int_0^1 \|\gamma'(t)\|^2 \, dt\right)^{\frac{1}{2}}$$

où la norme utilisée est la norme euclidienne sur $\mathbb{R}^3$. On utilisera la bilinéarité du produit scalaire euclidien et la formule de composition.

VII Différentielles d'ordre supérieur

NON REDIGEE.

(a) $D(D(f))$ et $D^2(f)$

On remarque que $D(D(f))(h)(k)$ est linéaire n h et en k. Autrement dit, si on note $D^2f(h,k) = D(D(f))(h)(k)$, D^2f est une forme bilinéaire.

De façon général, on peut identifier $\mathcal{L}(E_1, \mathcal{L}(E_2, F))$ avec $\mathcal{L}(E_1, E_2; F)$ en tant qu'espaces vectoriels. Lorsque E_1, E_2 et F sont de dimension finie, il en est de même pour $\mathcal{L}(E_1, \mathcal{L}(E_2, F))$ (en particulier les normes sont équivalentes).

Exemple : si B est une application bilinéaire continue, alors on a vu que $DB(a_1, a_2).(h_1, h_2) = B(a_1, h_2) + B(h_1, a_2)$. On a $D^2B(a_1, a_2).((h_1, h_2), (k_1, k_2)) = B(h_1, k_2) + B(k_1, h_2)$. (Donc la différentielle troisième sera nulle, et B est de classe C^∞).

Le lemme de Schwartz dit que la différentielle seconde est une forme bilinéaire *symétrique*.

Théorème VII.1. *Si f est deux fois différentiable en a alors*

$$D^2(f)(a)(v,w) = D^2(f)(a)(v,w)$$

pour tout vecteurs v, w.

Recette de preuve.— Pour tous vecteurs v, w on définit

$$\Delta(v,w) = f(a+v+w) - f(a+vv) - f(a+w) + f(a).$$

Il vaut mieux lire d'abord la preuve de la formule de Taylor ci-dessous (la technique utilisée est la même, mais la preuve du Lemme de Schwartz est plus subtile).

Lemme.

$$\lim_{s\to 0} \frac{\Delta(sv, sw)}{s^2} = D^2f(v,w).$$

On applique ensuite le lemme au couple (v,w) du théorème, et aussi au couple (w,v). Comme on a évidemment $\Delta(v,w) = \Delta(w,v)$, on obtient l'égalité voulue.

Montrons le lemme. On fixe d'abord deux vecteurs v, w et on pose, pour $t \in [0,1]$,

$$\gamma(t) = f(a+tv+w) - f(a+tv) - tD^2(f)(a)(w)(v)$$

On a $\gamma(1) - \gamma(0) = \Delta(v,w) - D^2(f)(w,v)$ qu'on veut estimer. Or

$$\gamma'(t) = L.v \text{ avec } L = Df(a+tv+w) - Df(a+tv) - D(D(f))(a)(w).$$

On écrit maintenant que f est deux fois différentiable en a, autrement dit $x \mapsto D(f)(x)$ est différentiable en a :

$$D(f)(a+h) = D(f)(a) + D(D(f)(a)(h) + \varepsilon(h)$$

avec $\varepsilon(h)$ négligeable devant h. On applique cette égalité deux fois, la première fois avec $h = tv + w$, puis avec $h = tv$. La différence des deux égalités obtenues donne

Ici $\varepsilon(h)$ est un élément de $\mathcal{L}(E,F)$.

$$\gamma'(t) = (\varepsilon(tv+w) - \varepsilon(tv)).v.$$

Soit maintenant $\varepsilon > 0$. Puisque $\varepsilon(h)$ est négligeable devant h, il existe $\alpha > 0$ tel que

$$\|h\| < \alpha \Rightarrow \|\varepsilon(h)\| < \|h\|\,\varepsilon.$$

Choisissons v, w tels que $\|v\|+\|w\| < \alpha$, et considérons à nouveau $t \in [0,1]$. On a alors $\|tv+w\| < \alpha$ et $\|tv\| < \alpha$, et on obtient

$$\|\gamma'(t)\| \leq < (\|tv+w\| + \|tv\|)\varepsilon \|v\| \leq (\|v\| + \|w\|)^2 \varepsilon.$$

D'où la même majoration pour $\gamma(1) - \gamma(0)$. On a montré que pour tout $\varepsilon > 0$ il existe $\alpha > 0$ tel que

$$\|v\| + \|w\| < \alpha \Rightarrow \left\|\Delta(v,w) - D^2(f)(w,v)\right\| < \varepsilon(\|v\| + \|w\|)^2,$$

autrement dit

$$\lim_{\|v\|+\|w\|\to 0} \frac{\left\|\Delta(v,w) - D^2(f)(w,v)\right\|}{(\|v\| + \|w\|)^2} = 0$$

On en déduit la limite voulue en fixant v, w et en appliquant ceci à sv et sw avec s qui tend vers 0.

(b) Formule de Taylor à l'ordre 2

Théorème. *Si $f : E \to F$ est deux fois différentiable au point a, alors*

$$f(a + \overrightarrow{h}) = f(a) + Df(a).\overrightarrow{h} + \frac{1}{2}D^2(f)(\overrightarrow{h}, \overrightarrow{h}) + \varepsilon(\overrightarrow{h})$$

avec $\varepsilon(h)$ négligeable devant $\|h\|^2$ (c'est-à-dire $\varepsilon(\overrightarrow{h})/\|h\|^2$ qui tend vers 0 lorsque $\overrightarrow{h}$ tend vers 0).

Recette de preuve.—

– On définit $\varepsilon(h)$ pour que la formule soit vraie, et on vérifie que $\varepsilon(0)$, $D\varepsilon(0)$, $D^2\varepsilon(0)$ sont nuls. C'est essentiellement une application répétée du théorème de composition. En particulier, il faut calculer la différentielle seconde d'une application bilinéaire (symétrique).

La preuve se ramène alors au lemme suivant.

Lemme VII.2. *Soit g une application telle que $g(0)$, $Dg(0)$, $D(D(g))(0)$ sont nuls. Alors $g(\overrightarrow{h})/\|h\|^2$ tend vers 0 lorsque $\overrightarrow{h}$ tend vers 0.*

Montrons le lemme.

– On écrit la définition de $D(D(g))(0)$, différentielle de $h \mapsto D(g)(h)$ en 0 : on obtient que, pour ε fixé et pour k assez petit,

$$\|D(g)(k) - D(g)(0) - D(D(g))(0)(k)\| = \|D(g)(k)\| < \varepsilon \|k\|.$$

– Pour h fixé assez petit, on a donc d'après l'inégalité des accroissements finis

$$\|g(h)\| \leq \left(\mathrm{Sup}_{k\in[0h]} \|Dg(k)\|\right) \|h\| \leq \varepsilon \|h\|^2$$

d'où le résultat.

(c) Dérivées partielles secondes

On considère $f : \mathbb{R}^p \to F$, base (e_i).

Exemple. Définition : si $f$admet des DP au voisinage de a, et si ces DP admettent elles-mêmes des DP, alors on dit que f admet des DP seconde, et on définit

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right)(a) = \lim \frac{1}{t}\left(\frac{\partial f}{\partial x_j}(a + te_i) - \frac{\partial f}{\partial x_j}(a)\right)$$

Proposition :

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = D^2 f(e_i, e_j).$$

En particulier, les formules de Schwartz deviennent...

Preuve en écrivant la définition de D(D(f)) ; lemme : si (A_n) tend vers A dans $\mathcal{L}(E, F)$ alors pour tout vecteur v, $A_n(v)$ tend vers $A(v)$. Ou plutôt : l'application $(L, v) \mapsto L.v$, de $\mathcal{L}(E, F) \times E$ dans F est continue.

Exercice 121.— Soit $\gamma(t) = f(a + t\overrightarrow{h})$. Montrer avec la même technique que $\gamma''(0) = D^2 f(\overrightarrow{h}, \overrightarrow{h})$.

Lorsque $F = \mathbb{R}$, définition de la matrice Hessienne dans la base canonique, formule donnant $D^2(f)$ en coordonnées. Exemple : formule de Taylor en coordonnées.

(d) Extrema

Définitions : minimum, minimum strict ; minimum local, minimum local strict. Maximum... A ECRIRE.

Exercice 122.—cas $E = \mathbb{R}$ Soit Ω un ouvert de $\mathbb{R}$, $f : \Omega \to \mathbb{R}$ et a un point de Ω. Avant de regarder la suite du cours, discuter des liens logiques entre les propriétés suivantes :

1. f a un minimum local au point a,
2. f a un minimum local strict au point a,
3. $f'(a) = 0$,
4. $f''(a) \geq 0$,
5. $f''(a) > 0$.

On pourra s'aider des exemples $x \mapsto x^2$, $x \mapsto x^3, x \mapsto x^4$.

Théorème 1 : Si a est un minimum local alors $f'(a) = 0$ et $f''(a) \geq 0$.

Théorème 2 : Si $f'(a) = 0$ et $f''(a) > 0$ alors f a un minimum local strict en a.

Ces deux théorèmes se généralisent (le second seulement en dimension finie).

Preuves en regardant $f(a+t\overrightarrow{h})$, ce qui nous ramène en dimension 1. Preuve de 3 par compacité de la sphère unité, et en utilisant la formule de Taylor ci-dessus.

Utilisation des théorèmes :

1) Méthode. Pour déterminer les max et min locaux d'une fonction, (1) on trouve les points critiques, (2) on écrit la formule de Taylor à l'ordre 2 en calculant les dérivées secondes, (3) on trouve le signe de la forme quadratique (4) on réfléchit.

2) Pour trouver le signe d'une forme quadratique : rappels sur la réduction, lien avec les valeurs propres. En pratique, plutôt la méthode de Gauss pour trouver la signature.

3) Exemples en dimension deux.

4) Exemple d'extremum global : $x^2 + y^2 - xy - 2x + y$. Montrer que $f(x, y)$ tend vers $+\infty$ lorsque $\|(x, y)\|$ tend vers $+\infty$. En déduire qu'il y a (au moins) un minimum global. Trouver ce(s) minimum(s).

(e) Dérivées d'ordre supérieur

Tous les résultats se généralisent aux différentielles d'ordre supérieures ; donnons rapidement quelques énoncés.

Soit $f : \Omega \to F$. On définit de façon récursive les applications de classe C^k : si f est différentiable sur Ω et si $D(f)$ est de classe C^2, alors f est dite de classe C^3 ; si $D(f)$ est de classe C^3, f est dite de classe C^4, etc.. Comme pour l'ordre 2, la différentielle d'ordre k en un point a s'identifie à un élément de $\mathcal{L}(E^k; F)$. Si f est de classe C^k, $D^k f$ est donc une application continue de Ω dans l'evn $\mathcal{L}(E^k; F)$. Le lemme de Schwartz se généralise : $D^k f(a)$ est symétrique.

Lorsque $E = \mathbb{R}^p$, pour que f soit de classe C^k sur Ω, il faut et il suffit qu'elle admettent en tout point x des dérivées partielles successives

$$\frac{\partial^k f}{\partial x_{i_k} \cdots \partial x_{i_1}}(x) = \frac{\partial}{\partial x_{i_k}}\left(\frac{\partial}{\partial x_{i_{k-1}}}\left(\cdots \frac{\partial}{\partial x_{i_1}} f\right)\right)(x)$$

qui sont des fonctions continues de x. En coordonnées, on a

$$D^k f(a)(h, \ldots, h) = \sum_{i_1, \ldots, i_k} \frac{\partial^k f}{\partial x_{i_k} \cdots \partial x_{i_1}}(a) h_{i_1} \cdots h_{i_k}.$$

La composée de deux applications de classe C^k est de classe C^k. Toute application polynômiale est de classe C^∞.

Formule de Taylor avec $o(\left\|\overrightarrow{h}\right\|^k)$.

VII.1 Exercices

Exercice 123.— Relier chaque objet à son type.

Objets : $D^2(f)$; $D((D(f))$; $D^2 f(a)$; $D(D(f))(a)$; $D(D(f))(a).\overrightarrow{h}$; $(D(D(f))(a).\overrightarrow{h}).\overrightarrow{k}$; $D^2 f(a).(\overrightarrow{h}, \overrightarrow{k})$.

Types : F ; $\mathcal{L}(E, F)$; $\mathcal{L}(E, \mathcal{L}(E, F))$; $\mathcal{L}(E, E; F)$; applications de Ω vers $\mathcal{L}(E, F)$, *etc.*.

Bonus : Quel type d'objet est le "D^2" de $D^2(f)$?

VIII Inversion locale, fonctions implicites

VIII.1 Théorie

(a) Difféomorphismes, le Théorème d'Inversion Locale

Soient E, F deux espaces vectoriels normés. Une application $\Phi : U \to V$ entre un ouvert U de E et un ouvert V de F est un *C^k-difféomorphisme* si elle est de classe C^k, elle est bijective, et sa bijection réciproque est aussi de classe C^k.

Remarques.
1) D'après le théorème de composition des différentielles, puisque $f \circ f^{-1}$ est l'identité, on voit que $Df(a)$ est une application linéaire inversible. En particulier, si E est de dimension finie, E et F ont la même dimension.
2) Les applications linéaires continues inversibles sont des C^∞-difféomorphismes. C'est aussi le cas des applications affines $x \mapsto a + L.x$ avec L continue inversible.
3) La réciproque d'un difféomorphisme est un difféomorphisme, la composée de deux difféomorphismes en est un. En particulier, l'ensemble $\mathrm{Diff}^k(\mathbb{R}^N)$ des difféomorphismes de $\mathbb{R}^N$ de classe C^k est un groupe pour la loi de composition.

Le théorème suivant montre que la différentielle permet de savoir si f est un difféomorphisme au voisinage d'un point a.

Théorème VIII.1. *Supposons que E et F sont des espaces de Banach, que $f : \Omega \to F$ est de classe C^1 sur l'ouvert Ω, et que a est un point de Ω en lequel la différentielle $Df(a)$ est inversible dans $\mathcal{L}(E, F)$.*

Alors il existe un ouvert U de E contenant a, et un ouvert V de F contenant $f(a)$, tels que la restriction de f à U soit un C^1-difféomorphisme entre U et V.

Le point le plus frappant du théorème dit que si $Df(a)$ est injective, alors f est également injective sur un petit voisinage U du point a.

Recette de preuve.—

0) En composant par des translations et des applications linéaires à la source et au but, on se ramène (comme dans la preuve ...) au as où $E = F$, $a = 0 = f(0)$ et $Df(0) = Id$.

1) Montrons d'abord l'énoncé avec "homéomorphisme" à la place de "difféomorphisme".

Le DL de f en 0 s'écrit

$$f(x) = x + \varepsilon(x)$$

avec $\varepsilon(x)$ négligeable devant x. Etant donné un y (dans un voisinage de 0 à préciser), on cherche à montrer que l'équation $x + \varepsilon(x) = y$, d'inconnue x, a une unique solution. Ceci revient à dire que la fonction $\phi_y : x \mapsto y - \varepsilon(x)$ a un unique point fixe. On va appliquer le théorème du point fixe de Picard-Banach.

Comme f est C^1, ε aussi, et on a $D\varepsilon(0) = 0$. Si $B_f(0, \delta)$ est une boule sur laquelle $\|D\varepsilon(x)\| \leq 1/2$, on en déduit, en utilisant le théorème des accroissements finis, que ε est 1/2-lipschitzienne sur cette boule ; par conséquent ϕ_y aussi. De plus, si $y \in B(0, \delta/2)$ alors $\phi_y(B_f(0, \delta)) \subset B(0, \delta)$. Le théorème de Picard s'applique : pour tout $y \in B(0, \delta/2)$ il existe un unique $x \in B(0, \delta)$ tel que $f(x) = y$. Notons $g(y)$ ce point x. On définit ainsi une fonction $g : B(0, \delta/2) \to B(0, \delta)$ telle que

$$\forall y \in B(0, \frac{\delta}{2}), f(g(y)) = y.$$

On pose également $V = B(0, \delta/2)$ et $U = g(B(0, \delta/2))$.

– $U = f^{-1}(B(0, \delta/2)) \cap B(0, \delta)$, ce qui montre que c'est un ouvert ; il contient bien sûr 0.

– Sachant ε est 1/2-lipschitzienne, on encadre $\|f(x_1) - f(x_2)\|$ par 1/2 et 3/2 fois $\|x_1 - x_2\|$, ce qui permet de voir que g est injective et 2-lipschitzienne, donc continue. (Remarquons aussi que f est injective sur $B(0, \delta)$).

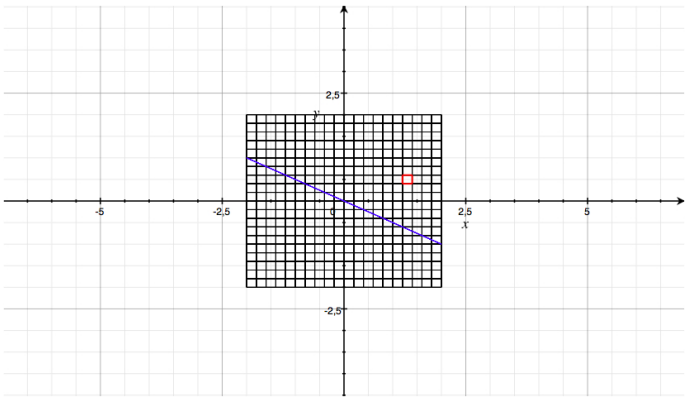
Terence Tao, médaille fields 2006, est un mathématicien aux capacités de travail étonnantes. Le 9 septembre 2011, il poste un message sur le forum Mathoverflow demandant s'il existe une version de ce théorème pour les applications qui sont seulement différentiables sur Ω (et pas de classe C^1. Le 12 septembre, à 19h21, il reçoit une réponse lui indiquant un article de Jean Saint-Raymond de 18 pages. Le soir même, à 00h10, il indique qu'il a posté un billet sur son blog expliquant la démonstration.

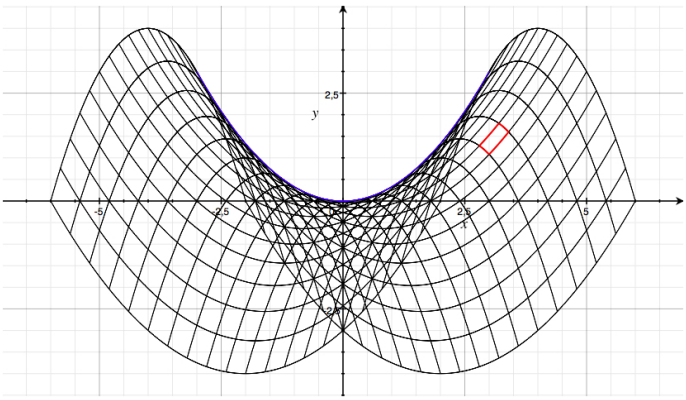
– f est une bijection entre U et V (f est injective sur la boule de rayon δ, et surjective par existence de g). Sa réciproque est g, qui est continue. Donc f induit un homéomorphisme entre U et V.

2) Montrons que, quitte à restreindre U, et V, f est un C^1-difféomorphisme. Soit U' l'ensemble des $x \in U$ en lesquels $Df(x)$ est inversible. L'ensemble des éléments inversibles de $\mathcal{L}(E, E)$ étant un ouvert, U' est ouvert et contient a. On pose aussi $V' = f(U')$, puisque $f : U \to V$ est un homéo, c'est aussi un ouvert. On a vu (juste après le théorème de composition au chapitre 6) que pour tout point x de U', f^{-1} est différentiable en $f(x)$, et on a la formule $D(f^{-1}) = ...$, et le théorème de composition montre que $D(f^{-1})$ est une application continue. $f : U' \to V'$ est donc un C^1-difféomorphisme.

Exemples **Exercice 124.—(Illustration du théorème d'inversion locale)**

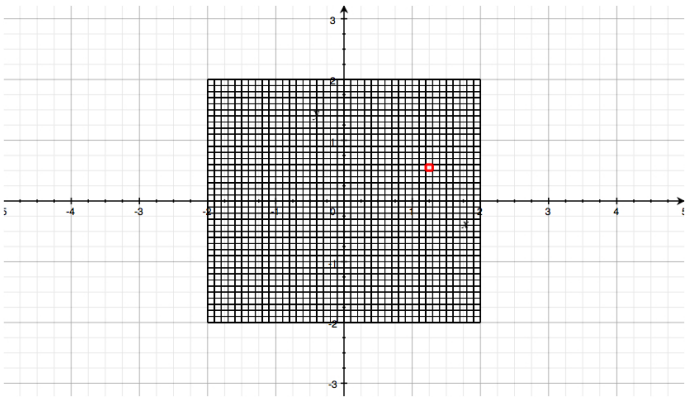
1. On considère l'application $F_1 : \mathbb{R}^2 \to \mathbb{R}^2$ définie par $F_1(x, y) = (2x + y, x^2 - y^2)$. Le dessin suivant représente une grille et son image par l'application F. La droite en bleu, à gauche, est envoyé sur la parabole en bleu, à droite.

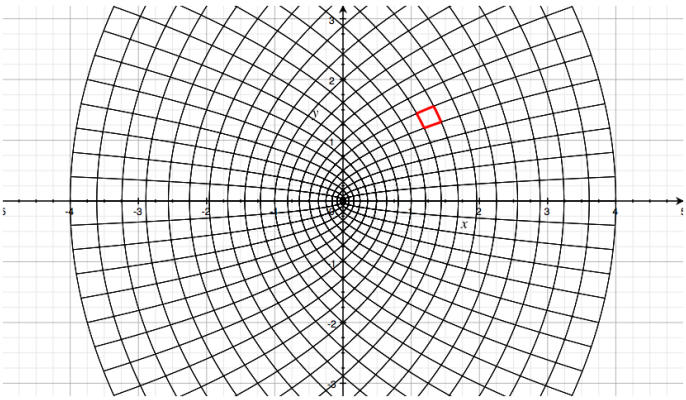




a. Déterminer les "bons" points a, ceux en lesquels l'hypothèse du théorème d'inverison locale est vérifiée. **b.** Soit a un "mauvais" point. L'application F_1 est-elle localement injective en a ? Localement surjective en $F_1(a)$? On pourra s'aider du dessin.

2. Mêmes question pour l'application $F_2 : \mathbb{R}^2 \to \mathbb{R}^2$ définie par $F_2(x, y) = (x^2 - y^2, 2xy)$, représentée ci-dessous.





3. Mêmes questions pour l'application $x \mapsto x^3$, de $\mathbb{R}$ dans $\mathbb{R}$.

Pour en savoir plus, on peut lire le joli article Le pli et la fronce sur le site Images des mathématiques.

Exercice 125.— **1.** On considère l'application $M \mapsto \exp(M)$ de $\mathcal{M}_N(\mathbb{R})$ dans $GL_N(\mathbb{R})$. Montrer qu'elle est différentiable en l'identité, et calculer sa différentielle. On admet que cette application est de classe C^∞ ; en déduire que toute matrice N assez proche de l'identité est l'exponentielle d'une matrice M, et que M est unique si on la suppose assez proche de l'identité.

2. Montrer de même que toute matrice N assez proche de l'identité est le carré d'une unique matrice $M(N)$ proche de l'identité. Donner un développement limité à l'ordre 1 de $M(Id + H)$ lorsque H tend vers 0.

Conséquences

1- Si $f : \Omega \to F$ est de classe C^1, et si $D(f)(a)$ est inversible en tout point de Ω, alors $f(\Omega)$ est un ouvert. En effet, en appliquant le TIL en un point a de l'ouvert de départ on obtient un ouvert V contenant $f(a)$ et qui est dans l'image.

2- Si de plus f est injective, alors c'est un C^1-difféomorphisme sur son image. En effet, f^{-1} est de classe C^1 d'après le TIL.

Proposition. *Tout C^1-difféomorphisme de classe C^k est un C^k-difféomorphisme.*

Recette de preuve.— On a la formule

$$D(f^{-1})(y) = (Df)^{-1}(f^{-1}(y)) = \mathrm{Inv} \circ Df \circ f^{-1}(y).$$

On sait, par hypothèse, que f^{-1} est de classe C^1. L'application Inv est de classe C^∞ (voir plus bas), et si $k \geq 2$ Df est de classe C^{k-1}, donc par composition $D(f^{-1})$ est de classe C^1, et f^{-1} est de classe C^2. Si $k \geq 3$, on applique à nouveau le théorème de composition, on obtient que f^{-1} est de classe C^3, etc..

Lemme. *L'application* $\mathrm{Inv} : u \mapsto u^{-1}$ *est C^∞, et sa différentielle est donnée par*

$$D(\mathrm{Inv})(u_0).\overrightarrow{h} = -\mathrm{Inv}(u_0) \circ \overrightarrow{h} \circ \mathrm{Inv}(u_0).$$

Recette de preuve.— On montre d'abord que f est différentiable. D'abord en $u_0 = Id$ par l'expression comme série, puis en tout u_0 par composition en écrivant $h^{-1} = (u_0^{-1}h)^{-1}u_0^{-1}$, soit $\mathrm{Inv} = A \circ \mathrm{Inv} \circ B$ où A et B sont les multiplications à gauche et à droite par u_0^{-1}, qui sont linéaires. Ensuite, la formule et une récurrence facile montre que l'inversion est C^∞.

(Autre argument en dimension finie : on a des formules).

– Traduisons l'énoncé en dimension 1. Si $f : \mathbb{R} \to \mathbb{R}$ est de classe C^1, sa différentielle en un point a s'écrit $h \mapsto D(f)(a).h = f'(a).h$; elle est inversible ssi $f('(a) \neq 0$. L'énoncé devient donc : si $f : \Omega \subset \mathbb{R} \to \mathbb{R}$, de classe C^1, vérifie $f'(a) \neq 0$, alors c'est un difféo local.

– Les coordonnées polaires donnent un exemple d'application qui vérifie en tout point non nul les hypothèses du TIL, et qui est donc un difféo local. Ce n'est pas un difféo, puisqu'elle n'est pas injective. Ceci montre la nécessité de l'hypothèse d'injectivité dans le point 2 ci-dessus (contrairement au cas de la dimension 1).

Difféomorphismes et champs de vecteurs A ECRIRE.

Définition d'un champ de vecteurs. Théorème de Cauchy-Lipschitz pour les champs de vecteurs de classe C^1 (admis). On dit que C est solution de l'EDO $x' = X(x)$ avec la condition initiale $x(0) = x_0$. Exemple des champs de vecteurs constants.

Définition de l'image $Y = f_*X$ d'un champ de vecteurs X par un difféomorphisme f. Propriété fondamentale : γ est solution de $x' = X(x)$ si et seulement si $f \circ \gamma$ est solution de $y' = Y(y)$. Théorème de redressement du flot (admis).

(b) Le Théorème des Fonctions Implicites

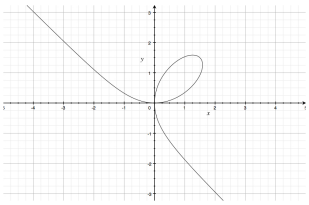
On dit qu'une équation du type

$$f(x, y) = 0$$

détermine y en fonction de x sur un domaine $\Omega = \Omega_1 \times \Omega_2$ si, pour tout $x \in \Omega_1$ donné, il existe un unique $y \in \Omega_2$ tel que $f(x, y) = 0$. Dans ce cas, en notant ϕ la fonction de Ω_1 dans Ω_2 qui associe à x cet unique y, on a

$$\forall (x, y) \in \Omega_1 \times \Omega_2, \quad f(x, y) = 0 \Leftrightarrow y = \phi(x).$$

Exercice 126.— Pour chacune de ces fonctions f, l'équation $f(x, y) = 0$ détermine-t-elle y en fonction de x sur le domaine de définition de f ? Sinon, trouver un domaine plus petit où c'est le cas. **1.** $f(x, y) = y - x^2$. **2.** $f(x, y) = y^2 - x$. **3.** $f(x, y) = x^2 + y^2 - 1$ **4.** $f(x, y) = x^3 + y^3 - 3xy$ (pour cette dernière équation, on s'aidera de l'ensemble des solutions dessiné ci-contre). **5.** $f(x, y) = y^3 - x$.



Le théorème des fonctions implicites permet de montrer qu'une équation détermine **localement** y en fonction de x.

Théorème VIII.2 (Théorème des fonctions implicites dans $\mathbb{R}^2$)**.** *Soit $f : \mathbb{R}^2 \to \mathbb{R}$ de classe C^1, et (x_0, y_0) un point tel que $f(x_0, y_0) = 0$. Supposons que*

$$\frac{\partial f}{\partial y}(x_0, y_0) \neq 0.$$

alors il existe $\alpha, \beta > 0$ tels que l'équation $f(x, y) = 0$ détermine y en fonction de x sur $]x_0 - \alpha, x_0 - \beta[\times]y_0 - \beta, y_0 + \beta[$:

$$\forall x \in]x_0 - \alpha, x_0 - \beta[\quad \exists! y \in]y_0 - \beta, y_0 + \beta[\quad f(x, y) = 0.$$

De plus, la fonction $\phi :]x_0 - \alpha, x_0 - \beta[\to]y_0 - \beta, y_0 + \beta[$ définie par

$$f(x, y) = 0 \Leftrightarrow y = \phi(x)$$

est de classe C^1.

On dit que l'équation $f(x, y) = 0$ détermine localement y en fonction de x au voisinage du point (x_0, y_0).

Exercice 127.— Soit f définie par $f(x,y) = x^2+y^2-1$. Déterminer les points (x_0, y_0) du cercle d'équation $f(x,y) = 0$ en lesquels l'hypothèse du théorème est vérifié. Donner des valeurs de α et β qui conviennent. Soit $(x_0, y_0) = (1, 0)$; l'équation détermine-t-elle localement y en fonction de x au voisinage de ce point ? Montrer que l'équation détermine localement x en fonction de y au voisinage de ce point.

Exercice 128.— Soit f définie par $f(x,y) = x^3 + y^3 - 3xy$. Montrer que l'ensemble C des solutions de l'équation $f(x,y) = 0$ est localement le graphe d'une fonction (y fonction de x ou x fonction de y) au voisinage de tout point autre que $(0,0)$.

Exercice 129.— Sous les hypothèses générales du théorème, calculer $\phi'(x_0)$. *Aide : dériver la relation* $f(x, \phi(x)) = 0$.

Plus généralement, on considère $f : \Omega \to F$, avec E et F deux evn de dimension finie (ou même deux espaces de Banach). On se donne deux espaces vectoriels supplémentaires $E = E_1 \oplus E_2$. L'espace E s'identifie alors au produit $E_1 \times E_2$, et on veut comprendre si une ligne de niveau de f est le graphe d'une fonction de E_1 dans E_2. On note $a = a_1 + a_2 = (a_1, a_2)$, et $C = f(a)$.

On peut avoir en tête les deux cas particulier, dans $E = \mathbb{R}^3$, (1) $E_1 = \{(x, y, 0)\}$ est le plan horizontal et $E_2 = \{(0, 0, z)\}$ est l'axe vertical, ou bien (2) $E_1 = \{(x, 0, 0)\}$ est un axe horizontal et $E_2 = \{(0, y, z)\}$ est un plan vertical.

Théorème VIII.3 (Théorème des fonctions implicites)**.** *Soit $k \geq 1$. Supposons que f est de classe C^k sur Ω, et que l'application linéaire*

$$Df(a)_{|E_2} : E_2 \to F$$

est une bijection. Alors l'équation $f(x_1, x_2) = C$ détermine localement x_2 en fonction de x_1 au voisinage du point (a_1, a_2) : il existe un ouvert Ω_1 de E_1 contenant a_1, un ouvert Ω_2 de E_2 contenant a_2 tels que $\Omega_1 \times \Omega_2 \subset \Omega$, et une fonction $\phi : \Omega_1 \to \Omega_2$ de classe C^k telle que, pour tout $(x_1, x_2) \in \Omega_1 \times \Omega_2$,

$$f(x_1, x_2) = C \quad \Leftrightarrow x_2 = \phi(x_1).$$

Cette équivalence revient à dire que, pour chaque x_1 fixé dans Ω_1, $\phi(x_1)$ est l'unique solution de l'équation, d'inconnue x_2, $f(x_1, x_2) = C$.

Il faut noter qu'on a $\phi(a_1) = a_2$ puisque $f(a_1, a_2) = C$. La conclusion dit qu'au voisinage du point a, l'ensemble $L_C = \{x \mid f(x) = C\}$ est le graphe d'une application de E_1 dans E_2 de classe C^k ; plus précisément,

$$L_C \cap (\Omega_1 \times \Omega_2) = \{(x_1, \phi(x_1) \mid x_1 \in \Omega_1\}.$$

On applique souvent cet énoncé dans le contexte suivant : on prend $E = \mathbb{R}^N$ et on note (e_i) la base canonique ; on choisit un entier p, et on pose

$$E_1 = \mathrm{Vect}(e_1, \ldots, e_p) = \{(x_1, \ldots, x_p, 0, \ldots, 0)\},$$

$$E_2 = \mathrm{Vect}(e_{p+1}, \ldots, e_N) = \{(0, \ldots, 0, x_{p+1}, \ldots, x_N)\},$$

qui sont deux sous-espaces vectoriel supplémentaires dans E. On choisi $F = \mathbb{R}^{N-p}$, qu'on munit également de la base canonique. Sous les hypothèses de l'énoncé, la matrice de $Df(a)$ dans la base canonique est alors rectangulaire avec $N-p$ lignes et N colonnes ; la matrice de la restriction $Df(a)_{|E_2}$ est la sous-matrice carrée de

L'énoncé postule l'existence d'une application linéaire bijective entre E_2 et F ; pour que cette hypothèse puisse être vérifiée, il faut que F soit de dimension $N - p$.

taille $(N-p)\times(N-p)$ à droite ; l'hypothèse essentielle de l'énoncé est que cette matrice est inversible.

Dans la situation la plus simple qui correspond à l'énoncé précédent, $E=\mathbb{R}^2$ et E_1, E_2 et F sont de dimension 1. L'application $Df(a)$ s'écrit en coordonnées

$$(h,k)\mapsto \frac{\partial f}{\partial x}(a)h+\frac{\partial f}{\partial y}(a)k$$

sa restriction à E_2 est simplement $k\mapsto \frac{\partial f}{\partial y}(a)k$ et la condition d'inversibilité équivaut à

$$\frac{\partial f}{\partial y}(a)\neq 0.$$

On retrouve ainsi l'énoncé dans $\mathbb{R}^2$.

Exercice 130.— Complément : exprimer la différentielle de ϕ en a_1 en fonction de la différentielle de f en (a_1,a_2). Exprimer ceci en terme de matrices.

Exercice 131.— En appliquant le théorème des fonction simplicites, montrer que l'équation matricielle $M^3+N^3-3MN=\mathrm{Id}$ définit localement N en fonction de M au voisinage de la solution $(\mathrm{Id},0)$.

Recette de preuve.— On définit $\overline{f}:\Omega\to E_1\times F$ en posant $\overline{f}(x_1,x_2)=(x_1,f(x_1,x_2))$. La différentielle de $\overline{f}$ au point a est

$$D\overline{f}(a)(h,k)=(h,Df(a).h+Df(a).k)=(h,Df(a)_{|E_1}.h+Df(a)_{|E_2}.k)$$

est inversible, comme on le voit en résolvant le système $D\overline{f}(a)(h,k)=(u,v)$ d'inconnu (h,k). Le TIL fournit une réciproque locale $G:V\to U_1'\times U_2$. On définit

$$U_1=\{x_1\in U_1', (x_1,C)\in V\}.$$

On a

$$f(x_1,x_2)=C\Leftrightarrow F(x_1,x_2)=(x_1,C)\Leftrightarrow (x_1,x_2)=G(x_1,C)\Leftrightarrow x_2=G_2(x_1,C).$$

Application fondamentale : courbes et surfaces implicites dans $\mathbb{R}^3$.

– Soit $f:\Omega\to\mathbb{R}$ une fonction de classe C^k définie sur un ouvert de $\mathbb{R}^3$. Soit $S=\{(x,y,z), f(x,y,z)=0\}$. Supposons que la différentielle de f ne s'annule pas sur S. Alors, au voisinage de tout point, S est localement le graphe d'une application de classe C^k.

NB : toute équation peut se mettre sous la forme ... = 0...

– Soit $f_1, f_2:\Omega\to\mathbb{R}$ deux fonctions de classe C^k définies sur un ouvert de $\mathbb{R}^3$. Soit $C=\{(x,y,z), f(x,y,z)=C\}$. Sous quelle conditions le TFI permet-il de dire que C est localement le graphe d'une fonction de classe C^k $\mathbb{R}$ dans $\mathbb{R}^2$ au voisinage d'un point a ? Soit $f=(f_1,f_2):\Omega\to\mathbb{R}$, et $L=Df(a)$. On peut appliquer le TFI dès lors que l'une des trois sous-matrices carrées 2×2 de L est inversible. Ceci équivaut à dire que L est de rang 2, c'est-à-dire surjective.

Exercice 132.—(interlude d'algèbre linéaire) Soient $p \leq N$, et L la matrice d'une application linéaire de $\mathbb{R}^N$ dans $\mathbb{R}^p$ (de taille $p \times N$, donc). On rappelle la propriété fondamentale du rang : la dimension du sous-espace vectoriel de $\mathbb{R}^N$ engendré par les lignes de L est égal à la dimension du sous-espace vectoriel de $\mathbb{R}^p$ engendré par les colonnes de L ; ce nombre s'appelle le rang de L.

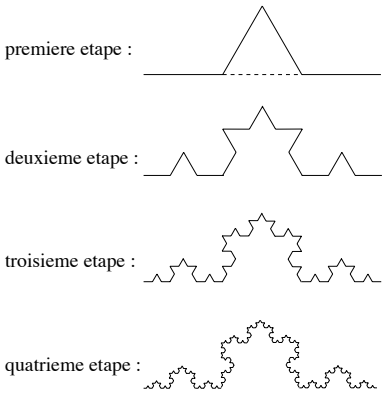
1. Montrer, à l'aide de la propriété fondamentale du rang, que L est surjective si et seulement si il existe dans L une sous-matrice carré inversible de taille $p \times p$.

2. Enoncer et démontrer un résultat analogue pour les matrices des applications linéaires de $\mathbb{R}^p$ dans $\mathbb{R}^N$.

VIII.2 Commentaires

Les difféomorphismes comme isomorphismes de la géométrie différentielle. Difféomorphismes vs homéomorphismes. A ECRIRE.

Une courbe de classe C^1 admet une tangente en chacun de ses points : pour faire court, nous dirons qu'elle est lisse. Un homéomorphisme préserve les propriétés topologiques, mais pas la propriété d'avoir des tangences. Le dessin ci-contre montre comment construire une courbe de VonKoch : cette courbe qui n'a aucune tangente est l'image du segment $[0, 1]$ de l'axe réel par un homéomorphisme du plan. Les C^1-difféomorphismes sont des homéomorphismes qui ont la propriété d'envoyer les courbes lisses sur des courbes lisses.



et ainsi de suite...

IX Sous-variétés, hypersurfaces

IX.1 Sous-variétés, sous-espace vectoriel tangent

Dans tout ce chapitre k est un entier ≥ 1. Dans $\mathbb{R}^n$, la partie

$$P_d = \{(x_1, \ldots, x_d, 0, \ldots, 0)\} = \{(x_1, \ldots, x_n) \mid x_{d+1} = 0, \ldots, x_n = 0\} = \mathrm{Vect}(e_1, \ldots, e_d)$$

(où d est un entier entre 0 et n) est un sous-espace vectoriel de dimension d.

Exercice 133.—(interlude d'algèbre linéaire) Montrer que pour tout sous-espace vectoriel E de $\mathbb{R}^n$ de dimension d, il existe un isomorphisme Φ de $\mathbb{R}^n$ tel que $\Phi(E) = E_d$.

Définition. Une partie S de $\mathbb{R}^n$ est appelée *hypersurface de classe C^k* ($k \geq 1$) si pour tout point p_0 de S il existe un voisinage U de p_0, un voisinage V de 0 dans $\mathbb{R}^n$, et un C^k-difféomorphisme $\Phi : U \to V$ tel que $\Phi(p_0) = 0$ et $\Phi(U \cap S) = V \cap P_{n-1}$.

Autrement dit, S est une hypersurface si elle est localement difféomorphe à l'hyperplan P_{n-1} (rappelons qu'un hyperplan est un set de dimension $n-1$). Plus généralement, S serait une sous-variété de dimension d si elle était localement difféomorphe à P_d (exercice : écrire la définition). On dira que le difféomorphisme Φ redresse localement S au voisinage du point p_0.

Lorsque γ est une courbe tracée sur S, les vecteurs vitesses de γ sont dits tangents à S.

Définition. Soit S une partie de $\mathbb{R}^n$, et $p_0 \in S$. Un vecteur $\overrightarrow{v} \in \mathbb{R}^n$ est *tangent à S au point p_0* s'il existe une application $\gamma : I \to \mathbb{R}^n$ définie sur un voisinage $I =]-\varepsilon, \varepsilon[$ de 0, telle que $\gamma(I) \subset S$, et $\gamma'(0) = \overrightarrow{v}$. On appelle *espace tangent à S en p_0*, et on note $T_{p_0}S$ l'ensemble des vecteurs tangents à S en p_0.

Exercice 134.—

1. Vérifier que si F est un sous-espace vectoriel, alors $T_{p_0}F = F$ pour tout $p_0 \in F$.

2. Vérifier que la notion de plan tangent est locale : pour tout ouvert U contenant p_0, $T_{p_0}(U \cap S) = T_{p_0}S$.

Une animation montrant les plans tangents à un ballon de rugby.

Théorème.

1. *Si S une hypersurface de $\mathbb{R}^n$ de classe C^k, alors en tout point p_0 de S l'espace tangent $T_{p_0}S$ est un hyperplan vectoriel de $\mathbb{R}^n$.*
2. *Si de plus $\Phi : U \to V$ est un C^k-difféomorphisme défini sur un ouvert U contenant S, alors $\Phi(S)$ est une hypersurface, et l'hyperplan tangent au point $\Phi(p_0)$ est l'image de $T_{p_0}S$ par la différentielle de Φ au point p_0 :*

$$T_{\Phi(p_0)}\Phi(S) = D\Phi(p_0)(T_{p_0}S).$$

Recette de preuve.— Soit S_0 une partie quelconque de $\mathbb{R}^n$, et $\Phi : U \to V$ comme dans l'énoncé, p_0 un point de S_0. Posons $S_1 = \Phi(S_0)$ et $p_1 = \Phi(p_0)$, et montrons d'abord l'égalité

$$T_{p_1}S_1 = D\Phi(p_0)(T_{p_0}S_0).$$

Si $\overrightarrow{v}$ est un vecteur tangent à S_0 en p_0, il existe une courbe $\gamma_0 :]-\varepsilon, \varepsilon[\to S_0$ tracée sur S_0 dont $\overrightarrow{v}_0$ est le vecteur vitesse en 0. La courbe $\gamma_1 = \Phi \circ \gamma_0$ est alors tracée sur $S_1 = \Phi(S_0)$; d'après la formule de composition des différentielles, son vecteur vitesse $\overrightarrow{v}_1 = \gamma_1'(0)$ est l'image par $D\Phi(p_0)$ de $\overrightarrow{v}_0$ (vérifier). Ceci montre une inclusion (laquelle ?). L'inclusion réciproque se montre de façon symétrique, puisque S_0 est l'image de S_1 par le difféomorphisme Φ^{-1} (écrire les détails ; il faut notamment utiliser la formule d'inversion des différentielles, $D(\Phi^{-1})(p_1) = (D\Phi(p_0))^{-1}$).

Supposons maintenant que S est une hypersurface, soit p_0 un point de S, et Φ un difféomorphisme redressant S au voisinage du point p_0. On a $\Phi(U \cap S) = V \cap P_{n-1}$. D'après ce qui précède, on en déduit que $D\Phi(p_0)$ envoie $T_{p_0}S$ sur $T_0 P_{n-1}$. Puisque P_{n-1} est un sous-espace vectoriel, il est son propre espace tangent. On en déduit

$$T_{p_0}S = (D\Phi(p_0))^{-1}(P_{d-1})$$

qui est bien un hyperplan.

IX.2 Exemples fondamentaux

Pour simplifier on énonce les résultats pour les hypersurfaces, mais la généralisation aux sous-variétés de dimension quelconque est immédiate.

(a) Graphes

Proposition. *Le graphe d'une fonction f de classe C^k de $\mathbb{R}^{n-1}$ dans $\mathbb{R}^n$ est une sous-variété de $\mathbb{R}^n$ de dimension $n-1$ et de classe C^k.*

Preuve : comme pour les surfaces de $\mathbb{R}^3$.

Exercice 135.— Déterminer son sous-espace vectoriel tangent en un point $p = (x, f(x))$.

(b) Equation implicite d'une hypersurface

Soit $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ une application de classe C^k, C une constante. Soit $S = \{p \mid f(p) = C\}$.

Théorème IX.1. *Si $Df(p) : \mathbb{R}^n \to \mathbb{R}^{n-d}$ est non nulle en tout point p_0 de S, alors S est une hypersurface, et*

$$T_p N = \mathrm{Ker}(Df(p)).$$

Complément : l'équation de l'hyperplan vectoriel tangent au point p s'écrit $Df(p).\overrightarrow{h} = 0$, ou en coordonnées,

$$\sum_{i=1}^{n} \frac{\partial f}{x_i} h_i = 0$$

qu'on écrit encore, en utilisant le produit scalaire et en introduisant le vecteur gradient de l'application f au point p, $< \overrightarrow{\mathrm{grad}} f(p), \overrightarrow{h} >= 0$. Ce vecteur est donc orthogonal à l'hyperplan tangent. [3]

3. Le résultat se généralise pour les sous-variétés, en considérant une fonction $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^d$ dont la différentielle est surjective en tout point.

Recette de preuve.— Si $Df(p) : \mathbb{R}^n \to \mathbb{R}^{n-d}$ est surjective pour tout point p de N, alors N est une sous-variété (exercice : généraliser la preuve donnée dans le cas des surfaces). Déterminons T_pN. Si γ est une courbe vérifiant $\gamma(0) = p$, alors $f(\gamma(t)) = C$ pour tout t, et donc par dérivation composée $Df(p).\gamma'(0) = 0$. Ceci montre que T_pN est inclus dans le noyau de $Df(p)$. Puisque cette application linéaire est surjective, son noyau est de dimension d, et on conclut.

Exemple : la sphère. Hyperplan affine tangent. Le cône n'est pas une surface de $\mathbb{R}^3$ si on enlève pas l'origine, car $T_0\mathcal{C}$ n'est pas un plan.

(c) Hypersurfaces paramétrées

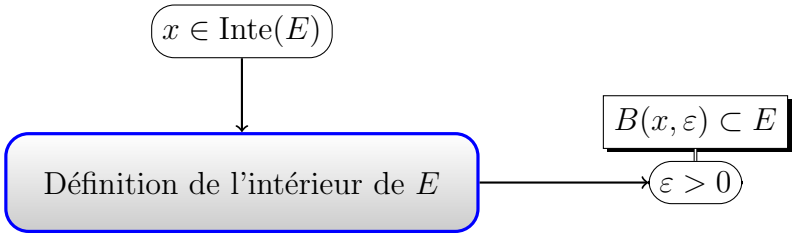
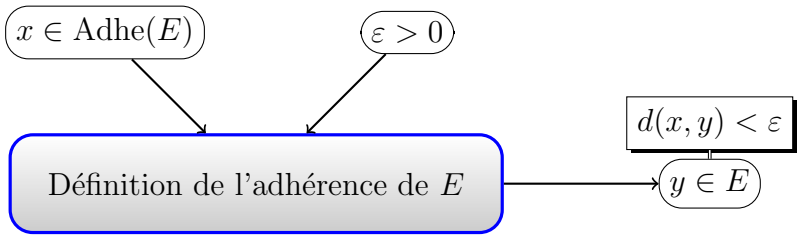
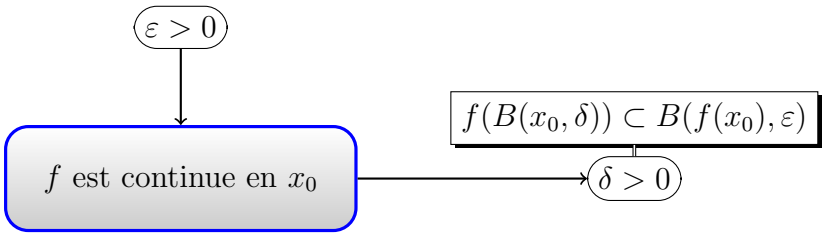
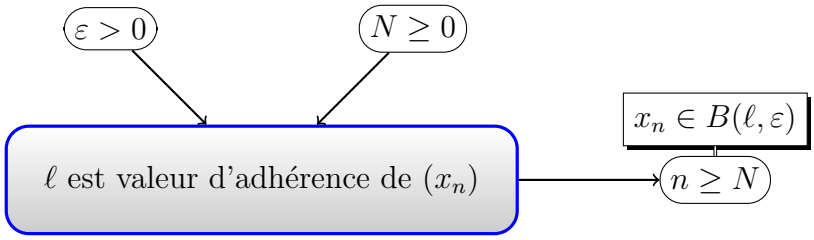
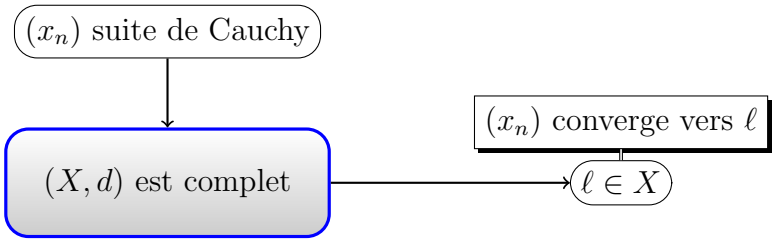
Soit $\Gamma : \Omega \subset \mathbb{R}^{n-1} \to \mathbb{R}^n$ une application de classe C^1.

Théorème IX.2. *Si la différentielle de Γ en tout point de Ω est injective, alors tout point a de Ω a un voisinage U tel que $S_U = \Gamma(U)$ est une sous-variété de dimension d, et*

$$T_pS_U = \mathrm{Im}(D\Gamma(a)).$$

Recette de preuve.— Si la différentielle de Γ en tout point de Ω est injective, alors tout point a de Ω a un voisinage U tel que $N = \Gamma(U)$ est une sous-variété de dimension d (exercice : généraliser le cas des surfaces de $\mathbb{R}^3$). Soit $p = \Gamma(a)$, déterminons T_pN. En considérant les courbes $t \mapsto \Gamma(t\overrightarrow{h})$ pour un vecteur $\overrightarrow{h}$ de $\mathbb{R}^d$ donné, on voit que T_pN contient l'image de $D\Gamma(a)$; puisque $D\Gamma(a)$ est injective, son image est de dimension d, ce qui conclut.

IX.3 Extrema liés

cf poly.

X Annexe : schémas