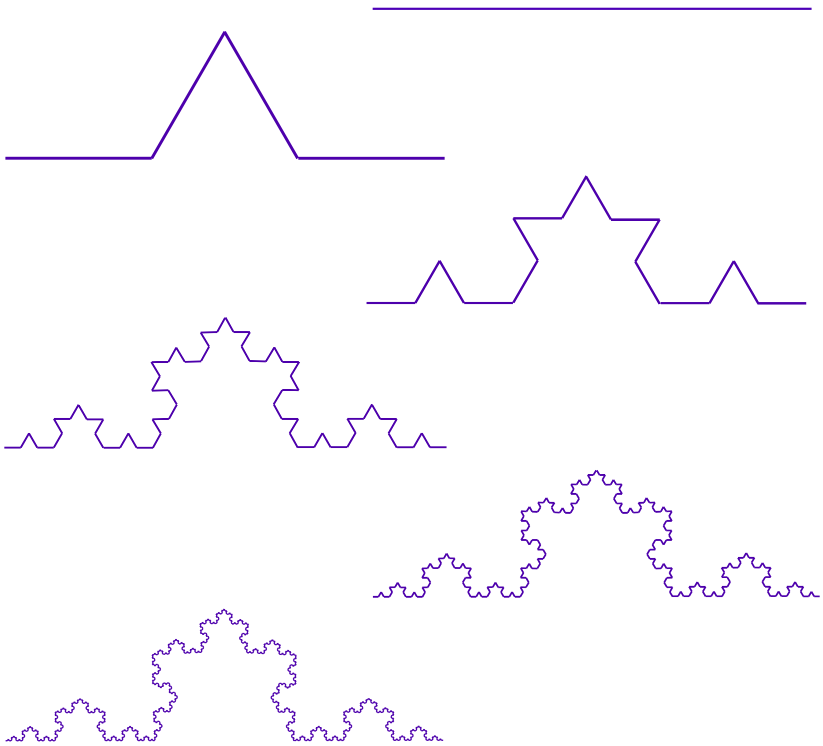
Topologie, Calcul différentiel



M301 – L3 MFA

Université Paris-Sud

D. Hulin 2015–16

Table des matières

Avertissement

Les sections entre *...* et les paragraphes grisés ne sont pas au programme.

Un grand merci à Anne Vaugon et Julien Sabin pour leur relecture
amicale, bienveillante, critique, dévouée, éclairante, féconde, et généreuse,
leurs critiques harassantes (parfois), implacables (toujours)
et leurs judicieux conseils.

Bibliographie

Quéffelec Topologie
Rouvière Petit guide de calcul différentiel
Gonnord-Tosel Topologie et analyse fonctionnelle
Gonnord-Tosel Calcul différentiel
Cartan Cours de calcul différentiel
Rudin Principles of mathematical analysis

0. * Préliminaire : constructions de $\mathbb{R}$ *

Partant du corps $\mathbb{Q}$ des rationnels, on évoque très brièvement deux constructions de $\mathbb{R}$. Ces quelques lignes se veulent simplement être un guide de lecture pour les références mentionnées dans le texte.

A Construire $\mathbb{R}$ par les coupures de Dedekind

Le corps $\mathbb{Q}$ est muni d'une relation d'ordre total, compatible avec sa structure de corps. On dit que $\mathbb{Q}$ est un corps totalement ordonné.

On constate que $\mathbb{Q}$ ne possède pas la propriété de la borne supérieure. Considérer par exemple une suite de rationnels $p_n \geq 0$ tels que $p_n^2 \to 2$ et $p_n^2 \leq 2$ (autrement dit, les p_n sont des approximations rationnelles par défaut de "$\sqrt{2}$" qui n'existe pas encore...).

On va y remédier en plongeant $\mathbb{Q}$ dans un corps plus gros, qui sera $\mathbb{R}$.

Définition 0.1 *Une coupure de $\mathbb{Q}$ est une partie $A \subset \mathbb{Q}$ telle que*

1. *A est non vide et distincte de $\mathbb{Q}$*
2. *si un rationnel p appartient à A, tout rationnel $q \leq p$ appartient également à A*
3. *A n'a pas de plus grand élément : si $p \in A$, il existe un rationnel q tel que $q \in A$ et $q > p$.*

Noter qu'il suit des propriétés (1) et (2) qu'une coupure est une partie majorée de $\mathbb{Q}$.

Si on connaît la fin de l'histoire, c'est-à-dire qu'on a déjà construit $\mathbb{R}$, on associe à une coupure $A \subset \mathbb{Q} \subset \mathbb{R}$ sa borne supérieure $\alpha = \sup A \in \mathbb{R}$. Réciproquement, si α est un réel, on lui associera la coupure définie par $A_\alpha = \{x \in \mathbb{Q} \mid x < \alpha\}$.

On fait maintenant abstraction de ce délit d'initié... et on définit $\mathbb{R}$ comme l'ensemble des coupures de $\mathbb{Q}$[1]. La coupure $A_p = \{x \in \mathbb{Q} \mid x < p\}$ associée à un rationnel $p \in \mathbb{Q}$ injecte $\mathbb{Q}$ dans $\mathbb{R}$.

Cet ensemble $\mathbb{R}$ est muni de la relation d'ordre associée à l'inclusion des coupures, qui prolonge la relation d'ordre de $\mathbb{Q}$. Reste à prolonger à $\mathbb{R}$ les lois de corps de $\mathbb{Q}$, à montrer que cela fait de $\mathbb{R}$ un corps commutatif totalement ordonné, et à constater qu'on a tout fait pour qu'il satisfasse la propriété de la borne supérieure. On montre en 3.4 que $\mathbb{R}$ ainsi construit est complet.

1. Rudin, Principles of mathematical analysis

B Construire $\mathbb{R}$ comme complété de $\mathbb{Q}$

Une suite (q_n) de rationnels est dite de Cauchy si, pour tout rationnel $r > 0$ strictement positif, on a $-r \leq q_n - q_m \leq r$ lorsque n et m sont suffisamment grands. On s'aperçoit (comme ci-dessus) qu'une suite de Cauchy de rationnels n'a pas toujours de limite dans $\mathbb{Q}$.

On va y remédier en construisant $\mathbb{R}$ comme "complété" de $\mathbb{Q}$, c'est-à-dire en rajoutant à $\mathbb{Q}$ une "limite" pour chaque suite de Cauchy de rationnels qui ne converge pas dans $\mathbb{Q}$. On veut bien sûr que deux suites de Cauchy adjacentes (dont la différence tend vers 0) aient même limite.

On introduit l'ensemble $\mathcal{C}(\mathbb{Q})$ des suites de Cauchy de rationnels, qui est naturellement muni d'une structure d'anneau, ainsi que l'idéal $\mathcal{N} \subset \mathcal{C}(\mathbb{Q})$ des suites de rationnels qui convergent vers 0. On vérifie que le quotient $\mathbb{R} = \mathcal{C}(\mathbb{Q})/\mathcal{N}$ est naturellement muni d'une structure de corps commutatif (ou, ce qui revient au même, que l'idéal $\mathcal{N}$ est maximal).

Le corps $\mathbb{Q}$ s'identifie à sous-corps de $\mathbb{R}$, par exemple en associant à un rationnel q la suite constante égale à q. Reste à vérifier que $\mathbb{R}$ est naturellement muni d'une relation d'ordre total, compatible avec celle de $\mathbb{Q}$ et avec sa structure de corps. Enfin, $\mathbb{R}$ ainsi construit est complet : si (x_n) est une suite de Cauchy de réels, on peut en effet choisir une suite (q_n) de rationnels qui soit adjacente à (x_n).

C Unicité

Qu'on se rassure : que l'on ait construit $\mathbb{R}$ par les coupures de Dedekind, ou bien comme complété de $\mathbb{Q}$, on tombe à isomorphisme près sur le même objet.

Définition 0.2 *Un corps $\mathbb{K}$ totalement ordonné est archimédien lorsque pour tout élément $x \in \mathbb{K}$ on peut trouver un entier $n \in \mathbb{N} \subset \mathbb{K}$ tel que $n > x$.*

Théorème 0.3 [2] *Il y a, à isomorphisme près, un unique corps totalement ordonné, archimédien et complet. En particulier, ce corps est commutatif.*
Il s'agit bien sûr de $\mathbb{R}$!

2. Lelong-Ferrand Arnaudiès, tome 2

1. Premiers pas de topologie

Dans ce premier chapitre, on introduit la notion d'espace métrique. Ces espaces – ou plus généralement les espaces topologiques qui ne seront pas étudiés ici – fournissent le bon cadre pour parler de limites et d'applications continues.

Ce premier chapitre est essentiellement constitué de définitions ainsi que d'exemples.

Certaines démonstrations, élémentaires, ne sont pas incluses dans le texte. Le lecteur se doit de les rédiger.

A Espaces métriques

Définition 1.1 *Un espace métrique est un ensemble X muni d'une distance d, c'est-à-dire d'une application $d : X \times X \to \mathbb{R}^+$ satisfaisant, pour tous points x, y, z de X, les conditions suivantes :*

— Séparation : $d(x,y) = 0$ *si et seulement si* $x = y$

— Symétrie : $d(x,y) = d(y,x)$

— Inégalité triangulaire : $d(x,z) \leq d(x,y) + d(y,z)$.

Le premier exemple d'espace métrique est la droite réelle $\mathbb{R}$ munie de la distance définie par $d(x,y) = |x-y|$ pour $x, y \in \mathbb{R}$. De même, on peut munir la droite complexe $\mathbb{C}$ de la distance définie par $d(x,y) = |x-y|$ pour $x, y \in \mathbb{C}$ (où $|\ |$ désigne maintenant le module du nombre complexe). Bien d'autres exemples se déduiront de ces deux là.

Définition 1.2 *Lorsque (X,d) est un espace métrique et A est une partie de X, la restriction d_A de d à $A \times A$ fait de (A, d_A) un espace métrique. On parle alors de "distance induite".*

Par exemple $[0, \infty[$ ou bien encore $\mathbb{Q}$ ou $\mathbb{Z}$, en tant que parties de $\mathbb{R}$, héritent de sa métrique naturelle.

Une autre famille fondamentale d'exemples est donnée par les espaces vectoriels normés.

Sur quels corps définir des normes ?

Soient $\mathbb{K}$ un corps commutatif et E un $\mathbb{K}$-espace vectoriel. Pour pouvoir définir une norme sur E, il faut au préalable que le corps $\mathbb{K}$ soit muni d'une valeur absolue, c'est-à-dire d'une application $t \in \mathbb{K} \to |t| \in [0, \infty[$ vérifiant, pour tous $s, t \in \mathbb{K}$:

— Séparation : $|t| = 0$ si et seulement si $t = 0$
— Inégalité triangulaire : $|t + s| \leq |t| + |s|$
— $|st| = |s|\,|t|$.

La valeur absolue d'un nombre réel et le module d'un nombre complexe sont respectivement des valeurs absolue sur le corps $\mathbb{R}$ des nombres réels, et sur le corps $\mathbb{C}$ des nombres complexes. Nous nous limiterons à ces deux exemples.

Proposition-Définition 1.3 *Soit* $(E, \|\ \|)$ *un espace vectoriel normé, c'est-à-dire un espace vectoriel réel (ou complexe)* E *muni d'une application* $x \in E \to \|x\| \in \mathbb{R}^+$ *vérifiant, pour tous points* x, y *de* E *et tout scalaire* $t \in \mathbb{R}$ *(ou* $t \in \mathbb{C}$*), les conditions suivantes :*

— Séparation : $\|x\| = 0$ *si et seulement si* $x = 0$
— Homogénéité : $\|tx\| = |t|\,\|x\|$
— Inégalité triangulaire : $\|x + y\| \leq \|x\| + \|y\|$.

Alors, l'application $(x, y) \in E \times E \to \|x - y\| \in \mathbb{R}^+$ *est une distance sur* E*.*

Remarquer qu'une distance sur un espace vectoriel qui provient d'une norme est, par construction, invariante par translations. Autrement dit, les translations

$$\tau_a : x \in E \to x + a \in E$$

$(a \in E)$ sont toutes des isométries (voir 1.37).

Exemple 1.4 Les applications

$$\begin{aligned}
\|\ \|_\infty &: (x_1, \cdots, x_n) \in \mathbb{R}^n \to \sup_{i=1..n} |x_i| \in \mathbb{R}^+ \\
\|\ \|_1 &: (x_1, \cdots, x_n) \in \mathbb{R}^n \to \sum_{i=1..n} |x_i| \in \mathbb{R}^+ \\
\|\ \|_2 &: (x_1, \cdots, x_n) \in \mathbb{R}^n \to \Big(\sum_{i=1..n} x_i^2\Big)^{1/2} \in \mathbb{R}^+
\end{aligned}$$

sont des normes sur $\mathbb{R}^n$. La norme $\|\ \|_\infty$ sera appelée “norme sup”.

Preuve Seule l'inégalité triangulaire pour $\|\ \|_2$ n'est pas tout à fait évidente. Elle découle de l'inégalité de Cauchy-Schwarz qui affirme que, pour tous réels $a_1, \cdots, a_n, b_1, \cdots, b_n$ on a

$$\sum_{i=1}^n a_i\, b_i \leq \Big(\sum_{i=1}^n a_i^2\Big)^{1/2} \Big(\sum_{i=1}^n b_i^2\Big)^{1/2}.$$

Cette inégalité se démontre en remarquant que le trinôme

$$t \longrightarrow \sum_{i=1}^n (a_i + tb_i)^2 = \sum_{i=1}^n a_i^2 + 2t \sum_{i=1}^n a_i b_i + t^2 \sum_{i=1}^n b_i^2$$

est toujours positif ou nul, et que son discriminant est donc négatif ou nul. □

Exemple 1.5 L'application $u = (u_n)_{n\in\mathbb{N}} \to \sup_{n\in\mathbb{N}} |u_n|$ définit une norme sur l'espace vectoriel $\ell^\infty(\mathbb{N})$ des suites réelles bornées.

Il ne faudra pas hésiter, pour se forger une intuition, à prendre pour premiers exemples d'espaces métriques $\mathbb{R}$ ou $\mathbb{R}^2$ avec l'une des normes introduites ci-dessus, ou bien encore des parties d'un de ces espaces. Dans ce cas, on peut en effet faire des dessins ! L'intérêt des notions que nous introduirons réside dans le fait que leur champ d'application est infiniment plus large.

On se permettra parfois de parler de "l'espace métrique X" sans faire explicitement référence à la métrique dont il est muni.

B Topologie d'un espace métrique

B.1 Topologie

Dès qu'on a une distance, on introduit les boules.

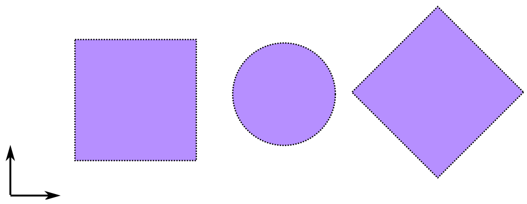
Définition 1.6 *Soit (X, d) un espace métrique, $x \in X$ un point de X et r un réel. La boule ouverte de centre x et de rayon r est, pour $r > 0$,*

$$B(x, r) = \{y \in X \mid d(x, y) < r\}.$$

La boule fermée correspondante est, pour $r \geq 0$,

$$B_f(x, r) = \{y \in X \mid d(x, y) \leq r\}.$$

Exercice 1.7 Dessiner quelques boules ouvertes ou fermées dans $\mathbb{R}$, dans $\mathbb{Z}$, dans $[0, \infty[$, et enfin dans $\mathbb{R}^2$ muni de l'une des normes $\| \|_\infty$, $\| \|_1$ ou $\| \|_2$.

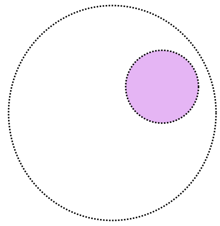


Une fois qu'on a défini les boules, on passe aux ouverts.

Définition 1.8 *Soit (X, d) un espace métrique. Une partie $U \subset X$ est ouverte si, pour tout point $x \in U$, il existe une boule ouverte $B(x, r) \subset U$ centrée en x et de rayon non nul incluse dans U. En particulier, la partie vide est ouverte.*

Proposition 1.9 *Toute boule ouverte est un ouvert.*

Les ouverts sont exactement les réunions de boules ouvertes.



Preuve Il suit de l'inégalité triangulaire qu'une boule ouverte est... un ouvert. Attention, c'est une propriété qui doit être démontrée : ne pas se laisser endormir par la terminologie (dont on appréciera cependant la cohérence !)
Il s'ensuit qu'une réunion de boules ouvertes est un ouvert. Tout ouvert étant, par définition, une réunion de boules ouvertes. $\square$

Proposition 1.10 *L'ensemble des ouverts de (X, d) constitue une sous-famille $\mathcal{T} \subset \mathcal{P}(X)$ de l'ensemble des parties de X. Cette famille $\mathcal{T}$ est une topologie, c'est-à-dire qu'elle vérifie les propriétés suivantes :*
- *le vide et l'ensemble X sont deux ouverts*
- *une union quelconque d'ouverts est ouverte*
- *une intersection finie d'ouverts est ouverte.*

On définit également les fermés.

Définition 1.11 *Soit (X, d) un espace métrique. Une partie $F \subset X$ est fermée si et seulement si son complémentaire ${}^cF \subset X$ est ouvert.*

Comme précédemment, on montrera qu'une boule fermée est un fermé : ouf !

Proposition 1.12 *Les fermés de (X, d) satisfont les propriétés de stabilité suivantes (duales, par passage au complémentaire, de celles satisfaites par les ouverts) :*
- *le vide et l'ensemble X sont deux fermés*
- *une intersection quelconque de fermés est fermée*
- *une union finie de fermés est fermée.*

Une intersection quelconque de parties ouvertes n'est pas toujours ouverte. Par exemple, dans $\mathbb{R}$, on a $\{0\} = \cap_{n \in \mathbb{N}^*}] - 1/n, 1/n[$. De même, une union quelconque de parties fermées peut ne pas être fermée.

La réunion $[0, 1] \cup]2, 3]$ n'est ni ouverte ni fermée dans $\mathbb{R}$.

Attention : Une partie de (X, d) peut à la fois être ouverte et fermée. Nous avons déjà mentionné que c'est le cas de l'ensemble vide et de X tout entier, qui sont tous deux à la fois des parties ouvertes et fermées de X. Mais cela peut également être le cas de parties non triviales de X. Nous reviendrons sur ce point important dans le chapitre 5 consacré à la connexité. Contentons-nous pour l'instant d'un exemple.

Dans $\mathbb{Z}$, un singleton (et donc toute partie de $\mathbb{Z}$) est à la fois ouverte et fermée. En effet, on a par exemple

$$\{0\} = B(0, 1/2) = B(0, 1) = B_f(0, 1/2)\,.$$

Plus généralement :

Exemple 1.13 Espace discret
Soit X un ensemble. L'application $d_0 : X \times X \to \mathbb{R}^+$ définie par $d_0(x,x) = 0$ et $d_0(x,y) = 1$ lorsque $x \neq y$ est une distance sur X. Pour cette distance, toutes les parties sont ouvertes (et donc toutes les parties sont fermées). La topologie correspondante, soit $\mathcal{T} = \mathcal{P}(X)$, est dite topologie discrète sur X.

On dit qu'un espace métrique (X,d) est discret lorsque sa topologie est la topologie discrète. Par exemple, la topologie sur $\mathbb{Z}$ associée à la distance d_0, ou bien à la distance induite par celle de $\mathbb{R}$, est la topologie discrète.

Il faudra prendre soin de distinguer les propriétés qui ne dépendent que de la topologie (continuité, compacité, connexité...) de celles qui dépendent du choix de la métrique (uniforme continuité, complétude...)

Définition 1.14 *On dit que deux distances d_1 et d_2 sur X sont topologiquement équivalentes (on dira parfois simplement "équivalentes") lorsqu'elles ont les mêmes ouverts, c'est-à-dire lorsqu'elles définissent sur X la même topologie.*

Cette notion fournit bien sûr une relation d'équivalence sur l'ensemble des distances sur X.

Par exemple, si d est une distance sur l'ensemble X, $2d$ ou encore $\inf(d,1)$ sont deux distances sur X qui sont équivalentes à d.

Il est facile de voir que les distances sur $\mathbb{R}^n$ associées aux normes $\| \|_\infty$, $\| \|_1$ ou $\| \|_2$ sont toutes équivalentes entre elles. Nous généraliserons ce résultat dans le chapitre consacré à la compacité (voir le théorème 2.25 : "en dimension finie, toutes les normes sont équivalentes").

Il sera également bien commode de parler en termes de voisinages.

Définition 1.15 *Soit x un point de l'espace métrique (X,d). Une partie $V \subset X$ est un voisinage de x si V contient une boule ouverte centrée en x (ou contient un ouvert contenant x).*

Exemple 1.16 Un ouvert est voisinage de tous ses points.

La remarque suivante est une conséquence immédiate de l'inégalité triangulaire.

Remarque 1.17 *Deux points distincts d'un espace métrique possèdent des voisinages disjoints : on dit que l'espace métrique est séparé.*

B.2 Topologie induite

On a déjà évoqué la distance induite sur une partie. Lui est associée la topologie induite. Ce n'est pas une notion compliquée si l'on est soigneux... sinon, elle peut fournir l'occasion de dire pas mal de bêtises : vous êtes prévenus !

Lemme 1.18 Topologie induite

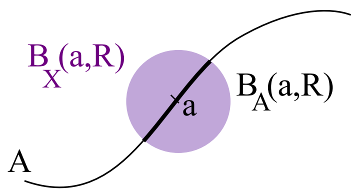
Soit A une partie d'un espace métrique (X, d), que l'on munit de la distance induite d_A.

Une partie $U \subset A$ est ouverte dans (A, d_A) si et seulement si U est la trace sur A d'une partie ouverte de X, c'est-à-dire s'il existe un ouvert $\tilde{U} \subset X$ tel que $U = \tilde{U} \cap A$.

Une partie $F \subset A$ est fermée dans (A, d_A) si et seulement si F est la trace sur A d'une partie fermée de X, c'est-à-dire s'il existe un fermé $\tilde{F} \subset X$ tel que $F = \tilde{F} \cap A$.

Maintenant qu'on a compris de quoi il s'agit, on allègera les notations et on dira désormais "ouvert dans A", et non plus dans (A, d_A). On se permettra également de noter d pour d_A.

Preuve On remarque que, pour tout point $a \in A$, on a l'égalité $B_A(a, r_a) = B_X(a, r_a) \cap A$.



Une partie ouverte $U \subset A$ de A s'écrit comme une réunion $U = \cup_{a \in U} B_A(a, r_a)$ de boules ouvertes de l'espace A centrées en chaque point de U. On a donc bien $U = \tilde{U} \cap A$, où $\tilde{U} := \cup_{a \in U} B_X(a, r_a)$ est un ouvert de X.

Soit maintenant $\tilde{U} \subset X$ un ouvert de X. On suppose que $U := \tilde{U} \cap A$ est non vide et on veut voir que U est un ouvert de A. Soit donc $a \in U$. Puisque $\tilde{U}$ est ouvert, il existe $r_a > 0$ tel que $B_X(a, r_a) \subset \tilde{U}$. Il suit que $B_A(a, r_a) \subset U$.

L'assertion sur les fermés en découle par passage au complémentaire. $\square$

Attention, la notion d'ouvert (ou de fermé) est une notion relative. Tout espace métrique est ouvert dans lui-même. Par exemple, $\mathbb{Q}$ est ouvert dans lui-même ; il n'est pas pour autant ouvert dans $\mathbb{R}$.

L'intervalle $[0, 1]$ est fermé dans $\mathbb{R}$. Il n'est pas ouvert dans $\mathbb{R}$. Par contre il est à la fois fermé et ouvert dans $A = [0, 1] \cup [2, 3]$.

Lorsqu'il y a ambiguïté, on donc prendra bien soin de préciser dans quel espace telle partie est ouverte, ou fermée.

Remarquer cependant que, lorsque $A \subset X$ est ouvert (dans X), les ouverts de A sont les ouverts de X contenus dans A. De même, lorsque $A \subset X$ est fermé (dans X), les fermés de A sont les fermés de X inclus dans A.

Rappelons que toute partie non vide minorée de $\mathbb{R}$ admet une borne inférieure, qui est son plus grand minorant (voir le chapitre préliminaire 0).

Exercice 1.19 Soit $G \subset (\mathbb{R}, +)$ un sous-groupe non réduit à $\{0\}$. Alors :

– soit $G \subset \mathbb{R}$ est dense

– soit G, muni de la topologie induite, est discret (1.13) ; dans ce cas il existe $a > 0$ tel que $G = a\mathbb{Z}$.

En particulier, $\mathbb{Z} + (2\pi)\mathbb{Z} \subset \mathbb{R}$ est dense.
On pourra introduire $a := \inf\{g \in G \mid g > 0\}$ et distinguer selon que a est nul ou pas.

Exercice 1.20 Montrer que les seules parties A de $\mathbb{R}$ qui sont à la fois ouvertes et fermées sont $\mathbb{R}$ lui-même et l'ensemble vide.
On suppose que A n'est pas vide, et on choisit $a \in A$. Remarquer alors que l'ensemble $B = \{b \in]a, \infty[\ / \ [a, b] \subset A\}$ est non vide, et introduire $\sup B \in \mathbb{R} \cup \{+\infty\}$.

B.3 Intérieur, adhérence

Proposition-Définition 1.21 *Soit (X, d) un espace métrique. L'intérieur $\mathring{A}$ d'une partie $A \subset X$ est le plus grand ouvert contenu dans A. C'est également la réunion de toutes les parties ouvertes contenues dans A.*

L'adhérence de A, notée $\overline{A}$, est le plus petit fermé contenant A. C'est également l'intersection de toutes les parties fermées contenant A.

Preuve L'ensemble vide est un ouvert contenu dans A. La réunion de tous les ouverts contenus dans A est un ouvert contenu dans A. Par construction, c'est le plus grand.

L'ensemble X est un fermé contenant A. L'intersection de tous les fermés contenant A est un fermé qui contient A. Par construction, c'est le plus petit. □

Proposition 1.22 *Un point $x \in X$ appartient à l'intérieur $\mathring{A}$ de A si et seulement si il existe une boule ouverte centrée en x et contenue dans A.*

Un point $x \in X$ appartient à l'adhérence $\overline{A}$ de A si et seulement si toute boule ouverte centrée en x rencontre A.

Preuve
Soit $x \in \mathring{A}$. Puisque $\mathring{A}$ est ouvert, il existe une boule ouverte $B(x, r) \subset \mathring{A} \subset A$ centrée en x et incluse dans A. Réciproquement, on suppose que $B(x, r) \subset A$. Alors $\mathring{A}$, comme réunion des ouverts contenus dans A, contient $B(x, r)$.

Soit $x \in X$. On suppose que la boule ouverte $B(x, r)$ (de rayon positif) ne rencontre pas A, i.e. que A est inclus dans le fermé ${}^c(B(x, r))$. On a donc $\overline{A} \subset {}^c(B(x, r))$: le point x n'appartient pas à $\overline{A}$. Réciproquement, supposons que x n'est pas dans l'adhérence de A. Puisque ${}^c(\overline{A})$ est ouvert, cela signifie qu'il existe $r > 0$ tel que $B(x, r) \subset {}^c(\overline{A}) \subset {}^cA$. □

Définition 1.23 *On dit qu'une partie $A \subset X$ est d'intérieur vide lorsqu'elle ne contient pas d'ouvert non vide, c'est-à-dire lorsque $\mathring{A} = \emptyset$.*

On dit qu'une partie $A \subset X$ est dense (dans X) lorsque X est le plus petit fermé contenant A, c'est-à-dire lorsque $\overline{A} = X$.

Par exemple $\mathbb{Q}$ et $\mathbb{R} \setminus \mathbb{Q}$ sont deux parties d'intérieurs vides de $\mathbb{R}$, qui sont denses dans $\mathbb{R}$.

Exercice 1.24 Soit (X, d) un espace métrique.

1. Montrer les identités suivantes entre intérieurs et adhérences d'une partie $A \subset X$ et de son complémentaire :
$$\overline{{}^cA} = {}^c(\mathring{A}) \quad \text{et} \quad {}^c(\overline{A}) = ({}^cA)^\circ.$$
2. On introduit la frontière $\mathrm{Fr}\,(A) = \overline{A} \setminus \mathring{A} = \overline{A} \cap \overline{{}^cA}$ d'une partie $A \subset X$. Montrer que
   (a) la frontière $\mathrm{Fr}\,(A)$ est un fermé de X
   (b) un point $x \in X$ appartient à la frontière de A si et seulement si toute boule ouverte centrée en x rencontre à la fois A et son complémentaire.

Exercice 1.25 On note ℓ^∞ l'espace des suites complexes bornées, c_c le sous-espace des suites nulles à partir d'un certain rang, et c_0 celui des suites qui convergent vers 0. On munit ces espaces de la norme définie par $\|x\|_\infty = \sup_{n\in\mathbb{N}} |x_n|$, et de la distance associée.

1. Montrer que $\| \ \|_\infty$ est bien une norme sur ℓ^∞.
2. Montrer que c_0 est fermé dans ℓ^∞.
3. Montrer que c_c est dense dans c_0.
4. Est-ce que c_c est dense dans ℓ^∞ ? Sinon, quelle est son adhérence ?

C Suites dans un espace métrique

Proposition-Définition 1.26 *Une suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de (X, d) est convergente s'il existe $\ell \in X$ tel que $d(x_n, \ell) \to 0$ lorsque $n \to \infty$, c'est-à-dire si pour tout $\varepsilon > 0$ il existe $N \in \mathbb{N}$ tel que pour tout $n \geq N$ on ait $d(x_n, \ell) \leq \varepsilon$.*

Un tel point $\ell \in X$ est alors unique. On dit que c'est la limite de la suite (x_n) et on note $\ell = \lim_{n\to\infty} u_n$.

En d'autres termes, la suite (x_n) converge vers ℓ si et seulement si son terme général x_n appartient, pour n assez grand, à n'importe quelle boule ouverte centrée en ℓ.

Preuve Soient ℓ_1 et ℓ_2 deux limites pour la suite (u_n). Il suit de l'inégalité triangulaire que, pour tout $n \in \mathbb{N}$, $d(\ell_1, \ell_2) \leq d(\ell_1, u_n) + d(u_n, \ell_2)$. On a donc $d(\ell_1, \ell_2) = 0$, d'où $\ell_1 = \ell_2$. □

Proposition 1.27 Caractérisation séquentielle de l'adhérence

Soit A une partie de l'espace métrique (X, d). Un point $\alpha \in X$ appartient à l'adhérence $\overline{A}$ si et seulement si il existe une suite $(a_n)_{n\in\mathbb{N}}$ d'éléments de A qui converge vers α.

En particulier A est fermée si et seulement si la limite de toute suite convergente (dans X) d'éléments de A appartient encore à A.

Preuve On utilise la proposition 1.22. Si $\alpha \in \bar{A}$, on choisit pour tout $n \geq 1$ un point $a_n \in B(\alpha, 1/n) \cap A$. On a bien $\alpha = \lim a_n$, donc α est limite d'une suite d'éléments de A.

Réciproquement, si $\alpha \in X$ est limite d'une suite (a_n) d'éléments de A, chaque boule ouverte $B(\alpha, r)$ centrée en α contient a_n lorsque n est assez grand. □

La notion de valeur d'adhérence est tout aussi importante que la notion de limite. Voir en particulier le chapitre 2.

Définition 1.28 *Un élément $\alpha \in X$ est une valeur d'adhérence de la suite (x_n) lorsqu'il existe une suite extraite de (x_n) qui converge vers α.*

Lemme 1.29 *Une suite convergente a une unique valeur d'adhérence.*

Par contre, une suite dans un espace métrique quelconque peut avoir une unique valeur d'adhérence sans converger. Considérer par exemple la suite réelle définie par $x_{2n} = 0$ et $x_{2n+1} = n$. Voir cependant la proposition 2.8, lorsque l'espace X sera supposé compact.

Lemme 1.30 *Soit (x_n) une suite de l'espace métrique (X, d). L'ensemble des valeurs d'adhérence de cette suite est*

$$\mathcal{V} = \bigcap_{n\in\mathbb{N}} \overline{\{x_k \mid k \geq n\}}\,.$$

En particulier l'ensemble $\mathcal{V}$ des valeurs d'adhérence de (x_n) est un fermé (éventuellement vide) de X.

Preuve Soit α une valeur d'adhérence de (x_n). Il existe une suite croissante d'indices $n_p < n_{p+1}$, qui tend vers l'infini, telle que $x_{n_p} \to \alpha$ lorsque $p \to \infty$. Le point α appartient donc, pour tout entier n, à $\mathcal{V}_n := \overline{\{x_k \mid k \geq n\}}$.

Soit $\alpha \in X$. On suppose que, pour tout entier n, on a $\alpha \in \mathcal{V}_n$. Nous allons construire une suite (x_{n_p}) extraite de (x_n), et qui converge vers α. Puisque $\alpha \in \mathcal{V}_0$, il existe un indice $n_1 \in \mathbb{N}$ pour lequel $d(x_{n_1}, \alpha) \leq 1$. Supposons construits $(x_{n_1}, \cdots, x_{n_p})$, avec $n_1 < \cdots < n_p$ et $d(x_{n_k}, \alpha) \leq 1/k$ pour $k = 1, \cdots, p$. Puisque $\alpha \in \mathcal{V}_{n_p+1}$, il existe $n_{p+1} \geq n_p + 1$ tel que $d(x_{n_{p+1}}, \alpha) \leq 1/(p + 1)$. □

Noter qu'une suite peut ne pas avoir de valeur d'adhérence. C'est le cas, par exemple, lorsque $X = \mathbb{R}$ pour une suite réelle croissante non majorée.

Exercice 1.31
1. Déterminer l'ensemble des valeurs d'adhérence d'une suite convergente, ou d'une suite périodique.
2. Déterminer l'ensemble des valeurs d'adhérence de la suite réelle de terme général $v_n = \cos(\Sigma_{k=1}^{n} \frac{1}{k})$.
3. Déterminer l'ensemble des valeurs d'adhérence de la suite réelle $u_n = \cos n$. *On pourra utiliser le résultat de l'exercice 1.19.*

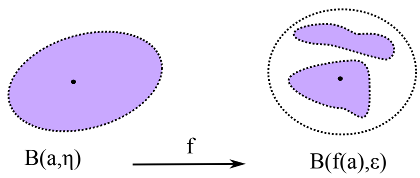
D Applications continues

D.1 Continuité

Définition 1.32 *Une application $f : (X, d_X) \to (Y, d_Y)$ entre deux espaces métriques est continue au point $a \in X$ lorsque, pour $x \in X$ proche de a, son image $f(x)$ est proche de $f(a)$. Autrement dit : pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que*

$$d_X(x, a) < \eta \ \Rightarrow\ d_Y(f(x), f(a)) < \varepsilon\,.$$

Ceci signifie que l'image réciproque de toute boule ouverte de centre $f(a)$ contient une boule ouverte de centre a.



Lemme 1.33 Caractérisation séquentielle de la continuité

L'application f est continue au point a si et seulement si, pour toute suite (x_n) d'éléments de X convergeant vers a, la suite $(f(x_n))$ converge vers $f(a)$.

Preuve Supposons f continue en a. Soit (x_n) convergente vers a. Soit $\varepsilon > 0$ et fixons $\eta > 0$ comme dans la définition précédente. Il existe un rang $N \in \mathbb{N}$ tel que, pour $n \geq N$, on ait $d_X(x_n, a) < \eta$. Il suit que $d_Y(f(x_n), f(a)) < \varepsilon$ pour $n \geq N$.

Supposons f non continue en a. Cela signifie qu'il existe $\varepsilon > 0$ tel que, pour tout entier $n \geq 1$, on puisse trouver un point $x_n \in X$ tel que $d_X(x_n, a) < 1/n$ mais avec $d_Y(f(x_n), f(a)) \geq \varepsilon$. La suite (x_n) converge donc vers a, mais la suite des images $(f(x_n))$ ne converge pas vers $f(a)$. □

Définition 1.34 *Une application $f : (X, d_X) \to (Y, d_Y)$ entre deux espaces métriques est continue si et seulement si elle est continue en chaque point.*

Proposition 1.35 *Soit $f : (X, d_X) \to (Y, d_Y)$ une application entre deux espaces métriques. Les conditions suivantes sont équivalentes :*

1. *l'application f est continue*
2. *l'image réciproque $f^{-1}(U)$ de tout ouvert $U \subset Y$ est un ouvert de X*
3. *l'image réciproque $f^{-1}(F)$ de tout fermé $F \subset Y$ est un fermé de X.*

Noter que l'image directe par $f : X \to Y$ continue d'une partie ouverte de X n'est pas toujours ouverte dans Y. Considérer par exemple une application constante $f : \mathbb{R} \to \mathbb{R}$ pour laquelle $f(\mathbb{R})$ est un singleton. De même l'image $f : X \to Y$ continue d'une partie fermée de X n'est pas toujours fermée dans Y. Penser cette fois-ci à l'application arctg $: \mathbb{R} \to \mathbb{R}$, pour laquelle $f(\mathbb{R})$ est un intervalle ouvert.

Preuve Une partie $U \subset Y$ est ouverte si et seulement si son complémentaire $F = {}^cU \subset Y$ est fermé. L'équivalence $2 \Leftrightarrow 3$ résulte donc de ce que, pour toute application $\varphi : E \to F$ entre deux ensembles, et toute partie $A \subset F$, on a l'égalité $\varphi^{-1}({}^c(A)) = {}^c(\varphi^{-1}(A))$.

Supposons f continue. Soient $U \subset Y$ une partie ouverte et $a \in f^{-1}(U)$. La partie U contient une boule ouverte $B(f(a), r)$ centrée en $f(a)$. Il suit de la continuité de f que $f^{-1}(U)$ contient une boule ouverte centrée en a. Ainsi, $f^{-1}(U)$ est bien ouvert dans X.

Supposons la condition 2 satisfaite. Soient $a \in X$ et $B(f(a), \varepsilon) \subset Y$ une boule ouverte. L'image réciproque $f^{-1}(B(f(a), \varepsilon))$ par f de cette boule est un ouvert de X contenant a, donc contient une boule ouverte $B(a, \eta)$. □

Exemple 1.36 Une application $f : X \to Y$ entre deux espaces métriques, avec X discret, est toujours continue.

Exemple 1.37 Quelques catégories d'applications continues

1. Une application $f : (X, d_X) \to (Y, d_Y)$ entre deux espaces métriques est **isométrique** lorsqu'elle préserve les distances : pour tous $x_1, x_2 \in X$ on a
$$d_Y(f(x_1), f(x_2)) = d_X(x_1, x_2)\,.$$
   Elle est alors continue.

   Une **isométrie** est une application $f : (X, d_X) \to (Y, d_Y)$ bijective et isométrique.

   Exemple : Les translations $\tau_a : x \in E \to x + a \in E$ d'un espace vectoriel normé (E, N) sont des isométries.

2. Soit $k \geq 0$. Une application $f : (X, d_X) \to (Y, d_Y)$ est **k-lipschitzienne** lorsque, pour tous pour tous $x_1, x_2 \in X$ on a
$$d_Y(f(x_1), f(x_2)) \leq k\ d_X(x_1, x_2)\,.$$
   Une application lipschitzienne est continue.

   Exemple : Soit (E, N) un espace vectoriel normé, muni de la distance associée d_N. Il suit de l'inégalité triangulaire que la fonction norme $N : (E, d_N) \to \mathbb{R}$ est 1-lipschitzienne, donc continue.

3. Soit $0 < \alpha \leq 1$. Une application $f : (X, d_X) \to (Y, d_Y)$ est **α-hölderienne** lorsqu'il existe une constante $c \geq 0$ telle que, pour tous pour tous $x_1, x_2 \in X$, on ait

$$d_Y(f(x_1), f(x_2)) \leq c\, d_X(x_1, x_2)^\alpha \,.$$

Une application hölderienne est continue.

De telles applications (lipschitziennes, ou hölderiennes) sont plus précisément uniformément continues (voir la définition 2.16).

Exemple 1.38 Soit (X, d) un espace métrique. Soient $a \in X$ et $A \subset X$ une partie non vide de X.

La fonction $x \in X \to d(x, a) \in [0, \infty[$ (distance à un point) est continue.

La fonction $x \in X \to d(x, A) := \inf_{a \in A} d(x, a) \in [0, \infty[$ (distance à une partie) est continue.

Il résulte en fait de l'inégalité triangulaire que ces fonctions sont toutes deux 1-lipschitziennes.

Exercice 1.39 Soit $F \subset X$ une partie fermée non vide d'un espace métrique.

1. Soit $x \in X$. Montrer que $d(x, F) = 0$ si et seulement si $x \in F$. Trouver un contre-exemple à cette assertion lorsque F n'est pas fermée.
2. Montrer que le fermé F est une intersection dénombrable d'ouverts.

D.2 Opérations sur les applications continues

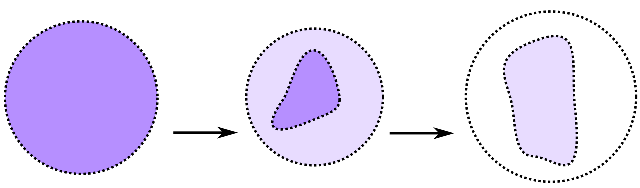
Proposition 1.40 Restriction d'une application continue

Soient $f : X \to Y$ une application continue entre espaces métriques et $A \subset X$ une partie de A. La restriction $f_{|A} : A \to Y$ est alors continue.

Preuve Laissée au lecteur. □

Proposition 1.41 Composée d'applications continues

Soient $f : X \to Y$ et $g : Y \to Z$ des applications entre espaces métriques. On suppose que f est continue au point $x_0 \in X$ et que g est continue au point $y_0 := f(x_0) \in Y$. Alors la composée $g \circ f : X \to Z$ est continue en x_0.



Preuve Notons $z_0 = g(f(x_0))$, et soit $B(z_0, r) \subset Z$ une boule ouverte centrée en z_0. Puisque g est continue au point y_0, $g^{-1}\big(B(z_0,r)\big)$ contient une boule ouverte $B(y_0, r_1)$. Puisque f est continue en x_0, $f^{-1}\big(B(y_0,r_1)\big)$ contient une boule ouverte $B(x_0, r_2)$. L'image réciproque $(g \circ f)^{-1}(B(z_0,r)) = f^{-1} \circ g^{-1}(B(z_0,r))$ contient donc la boule $B(x_0, r_2)$. □

Lemme 1.42 *Soient $f, g : X \to \mathbb{R}$ deux fonctions continues (au point a). La somme $f + g : X \to \mathbb{R}$ et le produit $fg : X \to \mathbb{R}$ sont encore continues (au point a).*

Preuve Laissée au lecteur. □

Corollaire 1.43 *Une fonction polynomiale $\mathcal{P} : \mathbb{R}^n \to \mathbb{R}$, associée à $P \in \mathbb{R}[X_1, \cdot, X_n]$, est continue sur $(\mathbb{R}^n, \| \, \|_\infty)$.*

Preuve D'après le lemme précédent, il suffit de le montrer pour un monôme. On raisonne alors par récurrence sur le degré du monôme sachant que les fonctions coordonnées sont continues. □

Exemple 1.44 Munissons $\mathrm{M}_n\mathbb{R} \simeq \mathbb{R}^{n^2}$ de la norme sup. L'application déterminant $\det : \mathrm{M}_n\mathbb{R} \to \mathbb{R}$ est continue. Le groupe linéaire $\mathrm{Gl}_n\mathbb{R} \subset \mathrm{M}_n\mathbb{R}$ est donc ouvert dans $\mathrm{M}_n\mathbb{R}$. Le groupe spécial linéaire

$$\mathrm{Sl}_n\mathbb{R} = \{M \in \mathrm{M}_n\mathbb{R} \mid \det M = 1\}$$

est une partie fermée de $\mathrm{M}_n\mathbb{R}$, et un sous-groupe fermé de $\mathrm{Gl}_n\mathbb{R}$.

Enfin, nous évoquons les suites d'applications continues.

Définition 1.45 *Soient X, Y deux espaces métriques, $f_n : X \to Y$ ($n \in \mathbb{N}$) et $f : X \to Y$ des applications.*

On dit que la suite (f_n) converge simplement vers f lorsque, pour tout point $x \in X$ et tout réel $\varepsilon > 0$, il existe un entier $N = N(x, \varepsilon)$ pour lequel, pour tout $n \geq N$, on ait $d(f_n(x), f(x)) \leq \varepsilon$.

On dit que la suite (f_n) converge uniformément vers f lorsque, pour tout réel $\varepsilon > 0$, il existe un entier $N = N(\varepsilon)$ pour lequel, pour tout $n \geq N$ et tout point $x \in X$, on ait $d(f_n(x), f(x)) \leq \varepsilon$.

On rappelle que la continuité ne se conserve pas par convergence simple. Considérer par exemple la suite définie sur $X = [0, 1]$ par $f(x) = x^n$. Par contre, la continuité se conserve par convergence uniforme.

Théorème 1.46 Limite uniforme d'applications continues

Soient $f_n : X \to Y$ ($n \in \mathbb{N}$) des applications continues (au point $a \in X$). On suppose que la suite (f_n) converge uniformément vers une application $f : X \to Y$. Alors f est également continue (au point $a \in X$).

Preuve Montrons la continuité de f au point a. Soit $\varepsilon > 0$. Il existe un rang $N \in \mathbb{N}$ pour lequel on a $d(f_n(x), f(x)) \leq \varepsilon$ pour tout point $x \in X$ et tout entier $n \geq N$. L'application f_N est continue (au point a). Il existe donc $\eta > 0$ tel que, lorsque $d(a, y) \leq \eta$, on a $d(f_N(a), y) \leq \varepsilon$. L'inégalité triangulaire montre alors que, pour $d(a, y) \leq \eta$, on a

$$d(f(a), f(y)) \leq d(f(a), f_N(a)) + d(f_N(a), f_N(y)) + d(f_N(y), f(y)) \leq 3\varepsilon\,.$$

□

D.3 Homéomorphismes

Définition 1.47 Homéomorphismes *Une application $f : X \to Y$ entre deux espaces métriques est un homéomorphisme si elle est bijective, continue ainsi que sa réciproque.*

Deux espaces métriques X et Y sont dits homéomorphes lorsqu'il existe un homéomorphisme $f : X \to Y$.

L'image directe d'un ouvert (ou d'un fermé) par un homéomorphisme est donc ouverte (ou fermée).

Exercice 1.48 On munit le cercle unité $\mathbb{S}^1 \subset \mathbb{C} \simeq \mathbb{R}^2$ de la topologie induite. L'application $t \in [0, 1[\to e^{2i\pi t} \in \mathbb{S}^1$ est continue et bijective, mais n'est pas bicontinue (la réciproque n'est pas continue).

Lorsque deux espaces métriques sont homéomorphes, ils sont indistinguables d'un point de vue topologique. Ce sont deux avatars d'un même objet. Pour montrer que deux espaces métriques sont homéomorphes, on a rarement d'autre option que d'exhiber un homéomorphisme entre eux.

Exercice 1.49

1. Montrer que deux intervalles fermés et bornés de $\mathbb{R}$ sont homéomorphes.
2. Montrer que deux intervalles ouverts de $\mathbb{R}$ sont homéomorphes.
3. Montrer que, pour $n \geq 1$, l'espace $(\mathbb{R}^n, \| \|_\infty)$ est homéomorphe à sa boule unité ouverte.

Pour montrer que deux espaces ne sont pas homéomorphes, on pourra chercher une "propriété topologique" (c'est-à-dire invariante par homéo–morphismes) qui soit satisfaite par l'un des espaces mais pas par l'autre : compacité ou connexité par exemple, voir les chapitres 2 et 5.

Exercice 1.50 L'espace métrique $\mathbb{R}$ est-il homéomorphe à $\mathbb{R} \setminus \{0\}$ muni de la topologie induite ? *On pourra utiliser la description des parties de $\mathbb{R}$ qui sont à la fois ouvertes et fermées, voir l'exercice 1.20.*

Revenons maintenant, pour la préciser, sur la définition 1.14.

Définition 1.51 Distances équivalentes *Soient d_1 et d_2 deux distances sur le même ensemble X.*

On dit que ces distances sont topologiquement équivalentes (on dira parfois simplement "équivalentes") lorsque ces distances induisent la même topologie sur X, ce qui équivant à demander que l'application identité $\mathrm{Id} : (X, d_1) \to (X, d_2)$ *soit un homéomorphisme.*

On dit que ces distances sont Lipschitz-équivalentes lorsque l'application $\mathrm{Id} : (X, d_1) \to (X, d_2)$ *est bi-lipschitzienne, c'est-à-dire lorsque cette application ainsi que sa réciproque* $\mathrm{Id} : (X, d_2) \to (X, d_1)$ *sont toutes deux lipschitziennes. Cela revient à dire qu'il existe une constante $c \geq 1$ telle que, pour tous $x_1, x_2 \in X$, on ait l'encadrement*

$$(1/c)\, d_1(x_1, x_2) \leq d_2(x_1, x_2) \leq c\, d_1(x_1, x_2)\,.$$

Bien entendu, lorsque deux distances sont Lipschitz-équivalentes, elles sont *a fortiori* topologiquement équivalentes.

Exemple 1.52 Soit $d : X \times X \to [0, \infty[$ une distance sur X. Alors $\delta = \inf(1, d)$ est une autre distance sur X, topologiquement équivalente à la première. En particulier, tout espace métrique est donc homéomorphe à un espace métrique borné. Par contre, si l'espace (E, d) n'est pas borné, les distances d et δ ne sont pas Lipschitz-équivalentes.

Exemple 1.53 Les distances induites sur $\mathbb{R}^n$ par les normes $\|\ \|_1$, $\|\ \|_2$ et $\|\ \|_\infty$ sont toutes Lipschitz-équivalentes. (Voir également le chapitre 4.)

Une dernière petite remarque, qu'il est plus sage d'écrire explicitement.

Lemme 1.54 *Soient d_1 et d_2 deux distances équivalentes sur X et (Y, d) un espace métrique. Une application $f : X \to Y$, ou bien $g : Y \to X$ est continue lorsqu'on munit X de la distance d_1 si et seulement si elle l'est lorsqu'on munit X de la distance d_2.*

E Produits d'espaces métriques

Commençons par étudier rapidement le cas d'un produit fini. Soit $(X_i, d_i)_{1 \leq i \leq n}$ une famille finie d'espaces métriques. On note $X = \prod_{i=1}^n X_i$ l'espace produit.

Lemme 1.55 *Les expressions*

$$\delta_1((x_1, \cdots, x_n), (y_1, \cdots, y_n)) = \sum_{i=1}^n d_i(x_i, y_i)$$

ou

$$\delta_\infty((x_1, \cdots, x_n), (y_1, \cdots, y_n)) = \sup_{i=1 \dots n} d_i(x_i, y_i)$$

définissent deux distances toopologiquement équivalentes sur le produit X.

Définition 1.56 *La topologie induite sur le produit $X = \prod_{i=1}^{n} X_i$ par l'une des distances précédentes est dite topologie produit sur X.*

La propriété fondamentale de la topologie produit (et qui la caractérise, voir la caractérisation séquentielle de la continuité) est la suivante.

Lemme 1.57 *Pour la topologie produit, une suite $(x(k) \,|\, k \in \mathbb{N})$ d'éléments du produit $X = \prod_{i=1}^{n} X_i$ converge vers ℓ si et seulement si chaque suite de coordonnées converge : $x_i(k) \to \ell_i$ lorsque $k \to \infty$ pour tout $1 \leq i \leq n$.*

Une application $f = (f_1, \cdots, f_n) : Y \to X$ à valeurs dans le produit est continue si et seulement si chaque application coordonnée f_i l'est.

Les preuves sont laissées au lecteur.

Exemple 1.58 L'espace $\mathbb{R}^n$ muni de la norme sup est l'espace métrique produit $((\mathbb{R})^n, \delta_\infty)$.

Exercice 1.59 Soit d une distance sur un espace X. L'application

$$d : (x, y) \in X \times X \to d(x, y) \in \mathbb{R}$$

est continue sur l'espace produit $X \times X$.

Passons maintenant à un produit dénombrable. Dans la définition suivante, nous munissons un produit dénombrable d'espaces métriques d'une topologie naturelle. Cette topologie est métrisable, c'est-à-dire associée à un choix de métrique - en fait, à tout plein de choix de métriques (comme dans le cas d'un produit fini, que l'on pouvait munir notamment de la distance δ_1 ou bien δ_∞).

Il faut retenir que cette topologie est caractérisée par ses suites convergentes : une suite dans l'espace produit converge si et seulement si chacune de ses composantes converge (proposition 1.63).

Définition 1.60 *La topologie produit sur l'espace produit $X = \prod X_i$ d'une famille dénombrable d'espaces métriques $(X_i, d_i)_{i \in \mathbb{N}}$ est la topologie associée à la distance*

$$d(x, y) = \sum_{i=0}^{\infty} 2^{-i} \inf(1, d_i(x_i, y_i))\,.$$

Expliquons la construction de d. On a remplacé chaque distance d_i par $\delta_i = \inf(1, d_i)$ qui est une autre distance sur X_i, équivalente à la première, et qui a l'avantage d'être majorée par 1 (exercice 1.52). Ensuite on somme les δ_i mais on prend soin de les pondérer par le terme général d'une série convergente (ici 2^{-i}) pour assurer la convergence de la série. Le fait que d ainsi construite soit une métrique sur X est immédiat.

Remarque 1.61 *Chaque projection* $p_j : x = (x_i) \in X \to x_j \in X_j$ $(j \in \mathbb{N})$ *est une application continue.*

Exemple 1.62 On obtient ainsi une topologie naturelle sur le produit infini d'intervalles $[0,1]^{\mathbb{N}}$, ou sur le produit infini d'espaces discrets $\mathcal{C} = \{0,1\}^{\mathbb{N}}$. Nous reviendrons plus tard sur ce dernier exemple : il s'agit de l'ensemble de Cantor dont nous construirons de jolies réalisations comme parties de $\mathbb{R}^2$ ou de $\mathbb{R}^3$ comme fractals.

Comme dit plus haut, ce n'est pas l'expression explicite de la métrique produit d qui est importante, mais bien la topologie qu'elle définit ou, ce qui revient ici au même (topologie métrisable), ses suites convergentes que nous décrivons maintenant.

Proposition 1.63 Suites dans un produit

Soit $(x(k) \,|\, k \in \mathbb{N})$ *une suite d'éléments du produit* X *avec, pour chaque entier* k, $x(k) = (x_i(k))_{i \in \mathbb{N}}$. *La suite est convergente, de limite* $\ell \in X$, *si et seulement si chaque composante converge, soit* $x_i(k) \to \ell_i$ *pour tout* $i \in \mathbb{N}$, *lorsque* $k \to \infty$.

Autrement dit, la convergence d'une suite dans le produit infini X équivaut à la convergence simple de chacune des suites de coordonnées.

Preuve Supposons que la suite $(x(k))$ converge vers $\ell = (\ell_i)_{i \in \mathbb{N}}$ dans X. Cela signifie que la distance

$$d(x(k), \ell) = \sum_{i=0}^{\infty} 2^{-i} \inf(1, d_i((x_i(k)), \ell_i))$$

tend vers 0 lorsque $k \to \infty$. En particulier, chaque terme de la somme, soit $\inf(1, d_i((x_i(k)), \ell_i))$ tend vers 0 lorsque $k \to \infty$. On en déduit que, pour tout $i \in \mathbb{N}$, $d_i((x_i(k)), \ell_i)$ tend vers 0 lorsque $k \to \infty$.

Supposons maintenant que chacune des suites de coordonnées $(x_i(k) \,|\, k \in \mathbb{N})$ converge vers ℓ_i dans X_i. On veut montrer que la suite $(x(k))$ converge vers ℓ dans X. Fixons donc $\varepsilon > 0$. On choisit $N \in \mathbb{N}$ tel que $\sum_{i \geq N} 2^{-i} \leq \varepsilon$. On a donc, pour tout k dans $\mathbb{N}$

$$d(x(k), \ell) \leq \sum_{i=0}^{N} 2^{-i} \inf(1, d_i((x_i(k)), \ell_i)) + \varepsilon\,.$$

Maintenant que N est fixé, la convergence simple des suites de coordonnées assure qu'il existe un rang K au-delà duquel chaque coordonnée $x_i(k)$ est arbitrairement proche de ℓ_i pour i entre 1 et N. Précisément, il existe K tel que, pour tout $1 \leq i \leq N$ et tout $k \geq K$, on ait $d_i((x_i(k), \ell_i) \leq \varepsilon$. Il suit que

$$d(x(k), \ell) \leq \varepsilon \sum_{i=0}^{N} 2^{-i} + \varepsilon \leq 3\varepsilon\,,$$

ce qu'on voulait. □

Proposition 1.64 Applications à valeurs dans un produit
Soit Y un espace métrique, et $f = (f_i) : Y \to X = \prod X_i$ une application. Alors f est continue (ou continue au point $y \in Y$) si et seulement si chaque facteur f_i est continu (ou continu au point $y \in Y$).

En d'autres termes, une application à valeurs dans un produit est continue si et seulement si chaque facteur l'est.
Preuve On utilise la caractérisation séquentielle de la continuité 1.33 et la description des suites convergentes dans un produit 1.63. □

Passons aux voisinages dans l'espace produit.

Proposition 1.65 Voisinages dans un produit
Soit $x = (x_i)$ un point de X. Une partie V de X est un voisinage de x si et seulement si il existe un entier N et un rayon $r > 0$ pour lequel le produit

$$\prod_{i=0}^{N} B_{X_i}(x_i, r) \times \prod_{i=N+1}^{\infty} X_i$$

est inclus dans V.

En d'autres termes, se donner un voisinage de x, c'est se donner une contrainte sur un nombre fini de facteurs.
Preuve Même démonstration que pour les suites convergentes, et laissée en exercice. On aurait d'ailleurs pu déduire la proposition 1.63 de celle-ci. □

Dans l'exercice suivant on construit une autre distance naturelle sur un produit dénombrable d'espaces métriques, mais qui n'induit en général pas la topologie produit que nous venons de construire.

Exercice 1.66 Soit $(X_i, d_i)_{i\in\mathbb{N}}$ une famille dénombrable d'espaces métriques.

1. Montrer que l'expression

   $$\delta(x, y) = \sup_{i\in\mathbb{N}} \{\inf(1, d_i(x_i, y_i))\}$$

   définit une distance sur le produit $X = \prod_{i\in\mathbb{N}} X_i$.
2. Décrire les suites convergentes dans (X, δ). Comparer aux suites convergentes dans (X, d), c'est-à-dire pour la topologie produit.
3. On suppose que tous les facteurs (X_i, d_i) sont égaux à $\{0, 1\} \subset \mathbb{R}$. Montrer que la distance δ induit sur X la topologie discrète.

F * Topologies non métrisables *

On peut aussi définir une topologie sur un ensemble X sans passer par une distance. C'est une famille $\mathcal{T} \subset \mathcal{P}(X)$ de parties de X qui vérifie les trois axiomes de la définition 1.10. Une topologie est dite métrisable lorsqu'elle est associée à une distance par la construction du paragraphe B.1. Ce n'est pas le cas de toutes les topologies (voir les exemples ci-dessous).

Il n'y a donc plus de notion de boule. On sait ce que sont les ouverts, les fermés, et les voisinages d'un point.

Une suite (x_n) de l'espace topologique X converge vers $\ell \in X$ lorsque, pour tout voisinage V de ℓ, il existe un rang à partir duquel tous les termes de la suite sont dans V. Une application $f : X \to Y$ entre deux espaces topologiques est continue lorsque l'image réciproque par f de tout ouvert de Y est un ouvert de X.

Dans ce cadre, l'image d'une suite convergente par une application continue est encore convergente, mais la caractérisation séquentielle de la continuité d'une application (1.33) peut tomber en défaut. Même chose pour la caractérisation séquentielle de l'adhérence d'une partie (1.27). La raison en est qu'un point n'a plus toujours une base dénombrable de voisinages.

A notre niveau, le champ d'utilisation de ces topologies non métrisables est restreint, voire inexistant, et nous nous limiterons donc aux espaces métriques.

Exemple 1.67 La topologie grossière sur un ensemble X, pour laquelle seuls l'ensemble vide et X lui-même sont ouverts, n'est pas métrisable dès que X a plus d'un point.

Et pour terminer cet aparté, complètement hors programme, un exemple pour vous montrer qu'algèbre et topologie peuvent faire bon ménage. On peut penser $\mathbb{K} = \mathbb{R}$, c'est déjà très intéressant !

Exemple 1.68 ** Topologie de Zariski **

Soient $\mathbb{K}$ un corps commutatif et $n \geq 1$. On désigne par $\mathbb{K}[X_1, \cdots, X_n]$ le $\mathbb{K}$-espace vectoriel des polynômes en n variables à coefficients dans $\mathbb{K}$. Pour chaque partie $S \subset \mathbb{K}[X_1, \cdots, X_n]$, on définit l'ensemble des zéros communs aux fonctions de S, soit

$$Z(S) = \{x \in \mathbb{K}^n \mid f(x) = 0\,,\ \forall f \in S\}\,.$$

1. Soit $S \subset \mathbb{K}[X_1, \cdots, X_n]$ et $< S >$ l'idéal engendré par S. Montrer qu'on a l'égalité $Z(S) = Z(< S >)$.
2. Soient I et J deux idéaux de $\mathbb{K}[X_1, \cdots, X_n]$. Montrer que $Z(I) \cup Z(J) = Z(I \cap J)$.
3. Soit $(I_k)_{k \in K}$ une famille d'idéaux de $\mathbb{K}[X_1, \cdots, X_n]$.
   (a) Montrer que $< \cup_{k\in K} I_k >= \sum_{k\in K} I_k$ (sommes finies d'éléments des I_k).
   (b) Montrer que $\cap_{k\in K} Z(I_k) = Z(\sum_{k\in K} I_k)$.
4. Montrer qu'il existe une topologie sur $\mathbb{K}^n$ dont les fermés sont les ensembles $Z(I)$, où I est un idéal de $\mathbb{K}[X_1, \cdots, X_n]$. C'est la topologie de Zariski.

5. Décrire les fermés de Zariski de la droite $\mathbb{K}$.
6. Si $f \in \mathbb{K}[X_1, \cdots, X_n]$, on note $D(f) = \{x \in \mathbb{K}^n \mid f(x) \neq 0\}$. Soient U un ouvert de Zariski non vide et $x \in U$. Montrer qu'il existe un polynôme f tel que $x \in D(f) \subset U$.
7. Montrer que, lorsque $\mathbb{K}$ est infini, deux ouverts non vides de $\mathbb{K}^n$ s'intersectent. La topologie de Zariski n'est pas séparée.
8. Un anneau $\mathbb{A}$ est dit noetherien [1] lorsque ses idéaux sont de type fini. On admet que, lorsque $\mathbb{A}$ est noetherien, l'anneau de polynômes $\mathbb{A}[X]$ est également noetherien.

   En déduire que toute partie de $\mathbb{K}^n$ possède la propriété de Borel-Lebesgue (2.40).
9. Montrer que tout ouvert de Zariski de $\mathbb{R}^n$ est dense pour la topologie usuelle.

2. Espaces compacts

La première notion topologique que nous allons aborder est la compacité. Cette notion joue un rôle crucial, en particulier dans l'étude des extrema d'une fonction à valeurs réelles.

A La compacité par les suites

Définition 2.1 Compacité : la propriété de Bolzano-Weierstrass
Un espace métrique X est compact si, de toute suite (x_n) à valeurs dans X, on peut extraire une sous-suite convergente.

Autrement dit, un espace métrique X est compact si et seulement si toute suite de points de X admet une valeur d'adhérence.

Définition 2.2 *Une partie $A \subset X$ d'un espace métrique (X, d) est compacte si elle l'est pour la distance induite d_A.*

Remarque 2.3 Il est important de noter que, si le fait pour une partie $A \subset (X, d)$ d'être ouverte ou fermée était une propriété relative, le fait que (A, d_A) soit compact est une propriété intrinsèque de cet espace métrique. On est ouvert ou fermé **dans** quelque chose. On est compact, ou on ne l'est pas !

Donnons tout de suite un premier exemple d'espace compact.

Exemple 2.4 Un intervalle fermé borné $[a, b] \subset \mathbb{R}$, muni de la distance induite par celle de $\mathbb{R}$, est compact.

Pour la preuve, nous supposons que $\mathbb{R}$ a été construit de sorte qu'il vérifie la propriété de la borne supérieure : toute partie non vide $A \subset \mathbb{R}$ majorée possède une borne supérieure, qui est son plus petit majorant (se reporter au chapitre préliminaire 0).

Preuve Soit (x_n) une suite à valeurs dans $[a, b]$. Cette suite peut avoir plusieurs valeurs d'adhérence. Nous allons construire la plus grande d'entre elles, qu'on appelle limite supérieure de x_n et que l'on note $\overline{\lim}\, x_n$.

Pour tout $n \in \mathbb{N}$, notons $v_n := \sup\{u_k \mid k \geq n\} \leq b$. La suite (v_n) est décroissante (au sens large) et minorée par a. Nous allons voir que $\alpha := \lim v_n = \inf v_n \in [a, b]$ est une valeur d'adhérence de (u_n).

On veut construire une suite (u_{n_k}) extraite de (u_n) qui converge vers α. On pose $n_0 = 0$. Supposons construits $n_0 < \cdots < n_k$ de sorte que, pour tout $1 \leq j \leq k$, on ait $\alpha - 2^{-j} \leq u_{n_j} \leq \alpha + 2^{-j}$. Choisissons $N > n_k$ tel que $\alpha \leq v_N \leq \alpha + 2^{-(k+1)}$. Par définition de v_N, il existe $n_{k+1} \geq N$ tel que $v_N - 2^{-(k+1)} \leq u_{n_{k+1}} \leq v_N$. On a donc $\alpha - 2^{-(k+1)} \leq u_{n_{k+1}} \leq \alpha + 2^{-(k+1)}$. La suite extraite ainsi construite converge vers α. □

Partant de cet exemple, il va être facile de décrire toutes les parties compactes de $\mathbb{R}$. Commençons par évoquer quelques propriétés générales des espaces compacts.

Lemme 2.5 *Soient (X, d) un espace métrique, et $A \subset X$ une partie de X.*

1. *Si A est compacte, alors A est une partie fermée de X.*
2. *Si X est compact et A est fermée dans X, alors A est compacte.*

Preuve

1. Supposons A compacte. Soient $\alpha \in \overline{A}$ et une suite (a_n) d'éléments de A qui converge vers $\alpha \in X$. Cette suite admet α comme seule valeur d'adhérence. Puisque A est compacte, (a_n) a une valeur d'adhérence dans A, donc $\alpha \in A$.
2. Supposons maintenant X compact et $A \subset X$ fermée. Soit (a_n) une suite d'éléments de $A \subset X$. Puisque X est compact, on peut en extraire une sous-suite qui converge vers un point $\alpha \in X$. Puisque $A \subset X$ est fermée, il suit que $\alpha \in A$. La conclusion suit. □

On en déduit immédiatement une description des parties compactes de $\mathbb{R}$, qui sera généralisée à $\mathbb{R}^n$ dans le paragraphe C. Commençons par une remarque.

Lemme 2.6 *Un espace métrique compact X est borné : si $x \in X$, il existe un rayon $R > 0$ tel que $X = B(x, R)$ (et donc $X \subset B(y, 2R)$ pour tout $y \in X$).*

Preuve Soit $x \in X$. Si X n'est pas borné, on peut construire une suite (x_n) d'éléments de X avec $d(x, x_n) \to \infty$. L'inégalité triangulaire assure alors que cette suite ne peut avoir de valeur d'adhérence. □

Corollaire 2.7 Les parties compactes de $\mathbb{R}$

Une partie $A \subset \mathbb{R}$ est compacte si et seulement si elle est fermée et bornée.

Preuve Supposons $A \subset \mathbb{R}$ compacte. On vient de voir que A est bornée (lemme 2.6), et que c'est une partie fermée de $\mathbb{R}$ (lemme 2.5).

Soit $A \subset \mathbb{R}$ une partie fermée et bornée. Elle est donc incluse dans un intervalle $[-R, R]$, qui est compact. Puisque A est fermée dans $\mathbb{R}$, elle est *a fortiori* fermée dans $[-R, R]$. Le lemme 2.5 assure donc que A est compacte. □

La proposition suivante est parfois bien utile pour montrer qu'une suite est convergente.

Proposition 2.8 *Soit (x_n) une suite à valeurs dans un espace compact X. Alors cette suite est convergente si et seulement si elle a une unique valeur d'adhérence.*

Preuve Une suite convergente dans un espace métrique quelconque a une unique valeur d'adhérence.

Supposons maintenant que X est compact, et que la suite (x_n) a une unique valeur d'adhérence α. Procédons par l'absurde et supposons que (x_n) ne converge pas vers α. Il existe alors une boule $B(\alpha, \varepsilon)$ de rayon non nul telle que $F := X \setminus B(\alpha, \varepsilon)$ contienne une infinité de termes de la suite. Cette partie F est fermée dans X, donc compacte. On obtient donc une seconde valeur d'adhérence $\beta \in F$ (et donc avec $\beta \neq \alpha$) pour (x_n). □

Exercice 2.9

1. Soit (x_n) une suite réelle à valeurs dans l'intervalle $[a, b]$.
   (a) Montrer que cette suite admet une plus grande valeur d'adhérence notée $\overline{\lim}\, x_n$ ainsi qu'une plus petite valeur d'adhérence notée $\underline{\lim}\, x_n$, et que
   - $\overline{\lim}\, x_n = \inf v_n = \lim v_n$, où $v_n := \sup\{u_k \mid k \geq n\}$
   - $\underline{\lim}\, x_n = \sup w_n = \lim w_n$, où $w_n := \inf\{u_k \mid k \geq n\}$.

   (b) En déduire que la suite (x_n) converge si et seulement si $\overline{\lim}\, x_n = \underline{\lim}\, x_n$.
2. Soit (x_n) une suite réelle positive. On suppose la suite sous-additive, i.e. telle que $x_{n+m} \leq x_n + x_m$ pour tous $n, m \in \mathbb{N}$.
   (a) Montrer que, pour tout $k \in \mathbb{N}^*$, $\overline{\lim}\left(\frac{x_n}{n}\right) \leq \frac{x_k}{k}$.
   (b) En déduire que la suite de terme général $\frac{x_n}{n}$ est convergente, et qu'on a $\lim_{n\to\infty} \frac{x_n}{n} = \inf_n \frac{x_n}{n}$.

B Compacité et applications continues

Théorème 2.10 *Soit $f : X \to Y$ une application continue entre deux espaces métriques. Supposons X compact. Alors l'image $f(X)$ est compacte.*

Preuve Montrons que $f(X)$ satisfait la propriété de Bolzano-Weierstrass. Soit (y_n) une suite dans l'image $f(X)$. On choisit, pour chaque indice n, un point $x_n \in X$ tel que $y_n = f(x_n)$. L'espace X étant compact, on peut extraire de la suite (x_n) une sous-suite (x_{n_k}) qui converge vers $\alpha \in X$. Par continuité de f au point α, il suit que $(y_{n_k}) = (f(x_{n_k}))$ converge vers $f(\alpha)$. □

Il s'ensuit en particulier que la compacité est une propriété topologique, c'est-à-dire dire invariante par homéomorphisme.

Corollaire 2.11 *Soit $f : X \to Y$ un homéomorphisme entre deux espaces métriques. Alors l'espace X est compact si et seulement si l'espace Y l'est.*

Une autre application fondamentale concerne les extrema des fonctions à valeurs réelles.

Corollaire 2.12 *Soit $f : X \to \mathbb{R}$ une application continue à valeurs réelles, définie sur un espace compact non vide. Alors f est bornée et atteint ses bornes.*

Preuve L'image $f(X) \subset \mathbb{R}$ est une partie compacte de $\mathbb{R}$. Elle est donc bornée et admet ainsi une borne inférieure et une borne supérieure. Puisque $f(X) \subset \mathbb{R}$ est fermée, on a $\sup\{f(x) \mid x \in X\} \in f(X)$. De même pour l'inf.
□

Rappelons la conséquence suivante de 2.12, dont la preuve est laissée au lecteur.

Corollaire 2.13 Théorème de Rolle

Soit $f : [a, b] \to \mathbb{R}$ une fonction continue, et dérivable en chaque point de l'intervalle ouvert $]a, b[$. On suppose que $f(a) = f(b)$.

Alors, il existe un point $c \in]a, b[$ où la dérivée $f'(c)$ s'annule.

On a vu qu'une bijection continue entre deux espaces métriques n'est pas toujours un homéomorphisme (1.48). Par contre, lorsque la source est compacte, on a le résultat suivant.

Proposition 2.14 *Soit $f : X \to Y$ une bijection continue entre deux espaces métriques compacts. L'application f est un homéomorphisme.*

Preuve On veut voir que l'application réciproque f^{-1} est continue. Soit donc $F \subset X$ une partie fermée de X. Cette partie F est donc un compact, et son image par f, soit $f(F) = (f^{-1})^{-1}(F) \subset Y$ est compacte, donc fermée dans Y. Le critère 1.35 permet de conclure. □

Exercice 2.15 Donner une autre preuve de la proposition 2.14, qui utilise cette fois le critère séquentiel de continuité (lemme 1.33).

Soit $f : X \to Y$ une application entre deux espaces métriques. L'application f est continue si, pour tout $x \in X$ et pour tout $\varepsilon > 0$, il existe $\eta > 0$ (qui dépend de ε **et** de x) et tel que $d(x, y) \leq \eta$ implique $d(f(x), f(x)) \leq \varepsilon$.

Considérons la fonction $x \in]0, 1] \to 1/x \in \mathbb{R}$. A ε fixé, plus x est proche de 0, plus on doit prendre η petit. Le théorème de Heine assure que ce phénomène ne se produit pas lorsque X est un espace compact.

Définition 2.16 Application uniformément continue
Soit $f : X \to Y$ une application entre deux espaces métriques. Elle est uniformément continue si, pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que pour tous $x, y \in X$ tels que $d(x, y) \leq \eta$, on a $d(f(x), f(y)) \leq \varepsilon$.

Autrement dit, à ε fixé, on peut choisir le même η pour tous les points $x \in X$.

Proposition 2.17 Théorème de Heine
Une application continue $f : X \to Y$ définie sur un espace compact est uniformément continue.

Preuve Supposons f non uniformément continue. Il existe donc un réel $\varepsilon > 0$ et deux suites (x_n) et (y_n) dans X avec $d(x_n, y_n) \to 0$ et telles que $d(f(x_n), f(y_n)) \geq \varepsilon$. Quitte à extraire une sous-suite, on peut supposer que la suite (x_n) est convergente, de limite α. Puisque les suites (x_n) et (y_n) sont adjacentes, on a aussi $y_n \to \alpha$. La continuité de f au point α assure alors que les suites $(f(x_n))$ et $(f(y_n))$ convergent toutes deux vers $f(\alpha)$, une contradiction. □

Exercice 2.18
1. Montrer qu'une fontion $f : \mathbb{R} \to \mathbb{R}$ dérivable et à dérivée bornée est lipschitzienne, donc uniformément continue.
2. Les applications $x \in \mathbb{R}_*^+ \to \sin(1/x) \in \mathbb{R}$, $x \in \mathbb{R} \to x^2 \in \mathbb{R}$ et $x \in \mathbb{R} \to |x|^{1/2} \in \mathbb{R}$ sont-elles uniformément continues ?

La notion de module de continuité que nous introduisons ci-dessous se révèle bien utile, par exemple pour quantifier les résultats d'approximation.

Définition 2.19 Module de continuité uniforme
Soit $f : X \to Y$ une application entre deux espaces métriques.
Son module de continuité uniforme $\omega_f : [0, \infty[\to [0, \infty]$ est défini pour tout réel $t \geq 0$ par

$$\omega_f(t) = \sup\{d(f(x), f(y)) \mid d(x, y) \leq t\} .$$

On a le résultat immédiat suivant :

Lemme 2.20 *L'application f est uniformément continue si et seulement si son module de continuité uniforme est continu en 0, i.e. si $\omega_f(t) \to 0$ lorsque $t \to 0$.*

Exemple 2.21 Une application f est k-lipschitzienne lorsque $\omega_f(t) \leq kt$. Elle est hölderienne d'exposant α lorsqu'il existe $c > 0$ avec $\omega_f(t) \leq ct^\alpha$.

Remarque 2.22 *Soit $f : [a,b] \to \mathbb{R}$ une fonction continue. Il suit du théorème de Heine que f peut être approchée uniformément par des fonctions en escaliers (on dit alors que la fonction f est réglée). C'est ce qui permet de démontrer que f est Riemann-intégrable.*

Le théorème de Heine permet également de montrer que l'on peut approcher uniformément f sur l'intervalle $[a,b]$ par une suite de fonctions polynomiales. C'est le théorème d'approximation de Weierstrass. Supposons pour simplifier que $[a,b] = [0,1]$; on peut alors montrer par exemple que la suite des polynômes de Bernstein

$$B_n^f(x) = \sum_{k=0}^{n} f(\frac{k}{n})\, x^k (1-x)^{n-k}$$

convient.

On peut même quantifier la vitesse d'approximation de f par B_n^f. Il existe une constante $c > 0$ telle qu'on ait, pour toute fonction continue f sur $[0,1]$, l'estimation

$$\sup_{t \in [0,1]} |f(x) - B_n^f(x)| \leq c\, \omega_f(1/\sqrt{n})\,,$$

ω_f désignant le module de continuité uniforme de f.

C Compacité dans un espace vectoriel normé de dimension finie

L'objectif de ce paragraphe est de décrire les parties compactes d'un espace vectoriel normé de dimension finie. Nous commençons par étudier la compacité d'un produit fini d'espaces métriques, résultat important en soi et sur lequel nous reviendrons au paragraphe D.

Proposition 2.23 *Un produit fini d'espaces métriques est compact si et seulement si chaque facteur est compact.*

Preuve Soient, pour $1 \leq i \leq n$, des espaces métriques (X_i, d_i). On note $X = \prod_{i=1}^n X_i$ leur produit que l'on munit de la topologie produit (voir 1.56), associée à la distance définie par

$$d(x,y) = \sup_{i=1}^{n} d_i(x_i, y_i)\,.$$

Chaque projection $p_i : X \to X_i$ est 1-lipschitzienne donc continue, et surjective. Il suit que si X est compact, chaque facteur X_i l'est également.

Supposons maintenant chaque facteur compact. Soit

$$x(k) = (x_1(k), \cdots, x_n(k))$$

($k \in \mathbb{N}$) une suite d'éléments de X. Le premier facteur X_1 étant compact, on peut extraire une sous-suite de sorte que la suite des premières coordonnées $x_1(k)$ converge. Après n extractions, on obtient une suite extraite de $(x(k))$ dont chacune des n suites coordonnées converge : cette sous-suite est donc convergente dans X (proposition 1.63). □

Nous généraliserons cet énoncé à un produit dénombrable dans le paragraphe D. Pour l'instant, nous en tirons des conséquences sur la topologie de $\mathbb{R}^n$. En un premier temps, $\mathbb{R}^n$ sera muni de la norme sup.

Lemme 2.24 Enoncé préliminaire, voir l'énoncé définitif en 2.27.
Les parties compactes de $(\mathbb{R}^n, \| \|_\infty)$ *sont les parties fermées et bornées. En particulier, la sphère unité*

$$S_\infty = \{x \in \mathbb{R}^n \mid \|x\|_\infty = 1\}$$

est un compact de $(\mathbb{R}^n, \| \|_\infty)$.

Preuve On raisonne comme dans le lemme 2.7 (description des compacts de $\mathbb{R}$) en remarquant qu'un produit $([-R, R]^n, \| \|_\infty)$ est compact, puisque chacun des facteurs l'est.

La sphère unité $S_\infty \subset (\mathbb{R}^n, \| \|_\infty)$ est bornée, et fermée puisque la norme $\| \|_\infty$ est une application continue sur $(\mathbb{R}^n, \| \|_\infty)$ (exemple 1.37). □

Nous déduisons maintenant de la description des parties compactes de $(\mathbb{R}^n, \| \|_\infty)$ le théorème fondamental suivant qui nous dit que, sur un espace vectoriel de dimension finie, toutes les normes sont équivalentes.

Théorème 2.25 *Soient* N_1 *et* N_2 *deux normes sur* $\mathbb{R}^n$. *Ces normes sont équivalentes, c'est-à-dire qu'il existe une constante* $C \geq 1$ *telle qu'on ait, pour tout* $x \in \mathbb{R}^n$, *l'encadrement*

$$C^{-1} N_1(x) \leq N_2(x) \leq C\, N_1(x)\,.$$

On retiendra, en particulier, le corollaire fondamental suivant.

Corollaire 2.26 *Toutes les normes sur* $\mathbb{R}^n$ *définissent la même topologie. C'est la topologie "naturelle" de l'espace vectoriel* $\mathbb{R}^n$.

Preuve L'équivalence des normes étant... une relation d'équivalence, il nous suffit de montrer que toute norme N sur $\mathbb{R}^n$ est équivalente à la norme sup, noté $\| \|_\infty$. Nous noterons $(e_1, \cdots, e_n)$ la base canonique de $\mathbb{R}^n$. Soit $x = \sum_{i=1}^n x_i e_i$ un vecteur. L'inégalité triangulaire et l'homogénéité de N donnent

$$\begin{aligned} N(x) = N(\sum_{i=1}^n x_i e_i) &\leq \sum_{i=1}^n |x_i|\, N(e_i) \\ &\leq (\sum_{i=1}^n N(e_i))\ \sup_{i=1}^n |x_i| = (\sum_{i=1}^n N(e_i))\ \|x\|_\infty\,. \end{aligned}$$

Il nous reste maintenant à minorer la norme N par la norme sup. Notons $C = (\sum_{i=1}^n N(e_i))$. On déduit de ce qui précède et de l'inégalité triangulaire que, pour tous $x, y \in \mathbb{R}^n$, on a

$$|N(x) - N(y)| \leq N(x-y) \leq C\|x-y\|_\infty\,.$$

L'application $N : (\mathbb{R}^n, \| \|_\infty) \to \mathbb{R}$ est donc continue. Restreignons la à la sphère unité

$$S_\infty = \{x \in \mathbb{R}^n \mid \|x\|_\infty = 1\}$$

qui est compacte (lemme 2.24) L'application N ne s'annule pas sur S_∞. Elle y atteint son inf, soit c, qui est donc non nul. L'homogénéité des normes assure alors que, pour tout $x \in \mathbb{R}^n$ non nul, on a

$$N(x) \geq c\,\|x\|_\infty\,,$$

ce qu'on voulait. □

On peut maintenant énoncer le lemme 2.24 de façon plus intrinsèque. Rappelons que dans un espace vectoriel normé, comme dans tout espace métrique, les parties compactes sont toujours fermées et bornées. Dans un espace vectoriel de dimension finie (qui s'identifie à $\mathbb{R}^n$ par le choix d'une base) on a équivalence entre ces conditions.

Corollaire 2.27 *Les compacts de $\mathbb{R}^n$, pour sa topologie naturelle d'espace vectoriel, sont les parties fermées et bornées.*

Exemple 2.28 La sphère unité $S_N = \{x \in \mathbb{R}^n \,|\, N(x) = 1\}$ associée à une norme sur $\mathbb{R}^n$ est compacte.

Exemple 2.29 Le groupe orthogonal $\mathrm{O}_n\mathbb{R} = \{M \in \mathrm{M}_n\mathbb{R} \mid {}^tM\ M = \mathrm{Id}\}$ est un sous-groupe compact de $\mathrm{Gl}_n\mathbb{R}$.

Le groupe spécial linéaire $\mathrm{Sl}_n\mathbb{R} = \{M \in \mathrm{M}_n\mathbb{R} \mid \det M = 1\}$ n'est pas compact lorsque $n > 1$.

Exercice 2.30 Les ensembles de matrices suivants sont-ils compacts ?

1. L'ensemble $\{P \in \mathrm{M}_n\mathbb{R} \mid P^2 = P\}$ des projecteurs.
2. L'ensemble $\{P \in \mathrm{M}_n\mathbb{R} \mid P^2 = P\,,\ \mathrm{Ker} P \perp \mathrm{Im}\, P\}$ des projecteurs orthogonaux, $\mathbb{R}^n$ étant muni d'une structure euclidienne.

Mentionnons en exercice une application fondamentale de la compacité du groupe orthogonal.

Exercice 2.31 La décomposition polaire.

1. Soit $M \in \mathrm{Gl}_n\mathbb{R}$. Montrer qu'il existe une unique matrice symétrique définie positive $S \in \mathrm{Sym}_n^+$ pour laquelle ${}^tM\, M = S^2$.
2. En déduire que l'application continue $\mathcal{P} : \mathrm{O}(n) \times \mathrm{Sym}_n^+ \to \mathrm{Gl}_n\mathbb{R}$ définie par $\mathcal{P}(O, S) = O\, S$ est bijective, puis que c'est un homéomorphisme.

C.1 Evn de dimension infinie

En dimension infinie, toutes les normes ne sont pas équivalentes.

Exercice 2.32 Soit ℓ^1 l'espace vectoriel des suites réelles (u_n) sommables, c'est-à-dire telles que $\sum |u_n| < \infty$. Une suite sommable est bornée. On peut donc munir ℓ^1 des normes

$$\|(u_n)\|_1 = \sum |u_n| \quad \text{et} \quad \|(u_n)\|_\infty = \sup |u_n|\,.$$

Ces deux normes ne sont pas équivalentes.

Dans un espace vectoriel normé de dimension infinie, il n'y a plus identité entre les parties compactes, et les parties fermées et bornées.

On démontrera le résultat suivant une fois qu'on aura dégagé la notion de complétude (chapitre 4.A).

Théorème 2.33 de compacité de Riesz *Soit (E, N) un espace vectoriel normé. Sa boule unité fermée est compacte si et seulement si l'espace E est de dimension finie.*

D Produits dénombrables d'espaces compacts

Nous généralisons ce qui a été vu plus haut (un produit fini d'espaces compacts est compact, proposition 2.23) à un produit dénombrable de compacts.

Proposition 2.34 Produit de compacts

Soit $(X_i, d_i)_{i\in\mathbb{N}^*}$ une famille dénombrable d'espaces métriques. L'espace métrique produit $X = \prod_{i=1}^\infty X_i$ (définition 1.60) est compact si et seulement si chacun des facteurs X_i est compact.

La partie directe suit, comme plus haut, de la continuité des projections $p_i : X \to X_i$ (remarque 1.61). Intéressons nous donc à la réciproque.

Essayons de faire la même preuve que pour un produit fini, en utilisant le critère de Bolzano-Weierstrass. On est amenés, partant d'une suite dans l'espace produit X, à en extraire successivement des sous-suites pour faire converger chaque facteur. Le hic est qu'après avoir extrait une infinité de fois des sous-suites, il risque cette fois-ci de ne plus rien rester ! On va donc être un peu soigneux dans nos extractions, et appliquer ce qu'on appelle le procédé diagonal de Cantor. Ce procédé constitue un "tour de main" tout aussi important que le théorème que nous démontrons ici.

Preuve : **Le procédé diagonal**
Soit $S = (x(k))_{k\in\mathbb{N}}$ une suite dans l'espace produit X. On cherche à en extraire une sous-suite convergente Σ. On note, pour $k \in \mathbb{N}$,

$$x(k) = (x_i(k) \mid i \in \mathbb{N}^*)\,.$$

Considérons la première suite coordonnées $(x_1(k))_{k\in\mathbb{N}}$. C'est une suite à valeurs dans l'espace compact X_1. On peut donc extraire de S une sous-suite $S_1 \prec S$ dont la suite des premières coordoonnées converge. Notons $x(k_1)$ le premier terme de S_1 (on le met bien au chaud) et introduisons $S_1' \prec S_1$ qui est la suite S_1 privée de son premier terme.

On extrait alors de S_1' une suite $S_2 \prec S_1'$ pour laquelle la suite des coordonnées dans le second facteur X_2, qui est compact, converge. Comme S_2 est extraite de S_1, la suite de ses premières coordonnées converge encore. On note $x(k_2)$ le premier terme de S_2. On observe que $k_2 > k_1$ puisque $S_2 \prec S_1'$.

En réitérant le procédé on construit successivement une famille de suites extraites $S_{n+1} \prec S_n' \prec S_n \prec \cdots \prec S_1' \prec S_1 \prec S$, où S_n' est S_n privée de son premier terme $x(k_n)$, et de sorte que les suites des $n+1$ premières coordonnées de S_{n+1} convergent. On note $x(k_{n+1})$ le premier terme de S_{n+1} et l'on observe que $k_{n+1} > k_n$.

Introduisons la suite $\Sigma = (x(k_1), x(k_2), \cdots, x(k_n) \cdots)$ (premier terme de S_1, premier terme de $S_2 \prec S_1$, premier terme de $S_3 \prec S_2 \prec S_1$ ...). D'une part, Σ est une suite extraite de S puisque, par construction, la suite des rangs (k_n) est strictement croissante.

Il nous reste à constater que la suite Σ est bien convergente dans X, c'est-à-dire que chacune de ses suites coordonnées converge. Fixons donc $i \in \mathbb{N}$. La suite Σ privée de ses premiers termes, soit $(x(k_i), x(k_{i+1}), x(k_{i+2}), \cdots)$ est une suite extraite de S_i. Il s'ensuit que la i-ème composante de Σ (à valeurs dans l'espace X_i) converge, ce qu'on voulait. $\square$

Exemple 2.35 L'espace de Cantor $\mathcal{C} = \{0,1\}^{\mathbb{N}}$ est un espace compact.

Le "cube infini" $[0,1]^{\mathbb{N}}$ est également un espace métrique compact.

Le procédé diagonal de Cantor est un outil fondamental. Vous le retrouverez, sans surprise, dans la preuve du théorème d'Ascoli (compacité dans les espaces de fonctions, voir le cours magistère du second semestre).

E Compacité par les recouvrements

Nous avons introduit les espaces métriques compacts en utilisant la propriété de Bolzano-Weierstrass. Nous en donnons maintenant une définition équivalente, en termes de recouvrements ouverts (propriété de Borel-Lebesgue).

E.1 Recouvrements ouverts dans les compacts

Commençons par quelques remarques concernant les recouvrements ouverts d'espaces compacts.

Définition 2.36 *Un recouvrement ouvert de l'espace métrique X est une famille $(U_i)_{i\in I}$ d'ouverts de X telle que $X = \cup_{i\in I} U_i$.*

Lemme 2.37 d'uniformité de Lebesgue

Soient X un espace métrique compact, et $(U_i)_{i\in I}$ un recouvrement ouvert de X. Il existe alors un rayon $r > 0$ tel que toute boule $B(x, r) \subset X$ de rayon r soit incluse dans un (au moins) des ouverts U_i du recouvrement.

Preuve Procédons par l'absurde. Sinon il existe une suite (x_n) de X telle que, pour chaque $n \in \mathbb{N}^*$, la boule $B(x_n, 1/n)$ ne soit contenue dans aucun des ouverts U_i du recouvrement. L'espace X étant compact, on peut supposer – quitte à passer à une suite extraite – que la suite (x_n) est convergente de limite α. Puisque la famille d'ouverts U_i recouvre X, il existe un indice i_0 ainsi qu'un rayon $r_0 > 0$ tels que $B(\alpha, r_0) \subset U_{i_0}$. Pour n grand, on aura $B(x_n, 1/n) \subset B(\alpha, r_0)$: contradiction. □

Lemme 2.38 Approximation d'un compact par un ensemble fini

Soient X un espace métrique compact. Soit $r > 0$. Il existe un recouvrement fini de X par des boules ouvertes de rayon r.

Preuve Si ce n'est pas le cas, on peut construire récursivement une suite (x_n) de points de X telle que, pour tout $n \geq 1$, on ait $x_n \notin \cup_{i=1}^{n-1} B(x_i, r)$. Pour $i \neq j$, on a donc $d(x_i, x_j) \geq r$, et cette suite n'admet pas de valeur d'adhérence : une contradiction. □

Ce lemme nous dit que, pour tout $r > 0$, on peut trouver une famille finie $\mathcal{F}$ de points de X (les centres de nos boules de rayon r) telle que tout point $x \in X$ se trouve à distance au plus r d'un point de $\mathcal{F}$.

La notion d'espace séparable ne nous servira pas cette année. Nous lui consacrons cependant un exercice.

Exercice 2.39

1. Soit (X, d) un espace métrique compact. Montrer qu'il existe une partie $D \subset X$ dénombrable, et dense dans X. On dit que (X, d) est séparable.
2. Montrer que la boule unité fermée de l'espace $(\ell^1(\mathbb{N}), \| \|_1)$ des suites réelles sommables est séparable.
3. Montrer que la boule unité fermée de l'espace $(\ell^\infty(\mathbb{N}), \| \|_\infty)$ des suites réelles bornées n'est pas séparable.

E.2 Borel-Lebesgue

Définition 2.40 La propriété de Borel-Lebesgue

Un espace métrique X satisfait la propriété de Borel-Lebesgue si, de tout recouvrement ouvert $(U_i)_{i\in I}$ de X, on peut extraire un sous-recouvrement fini.

Cela signifie que, pour toute famille d'ouverts $(U_i)_{i\in I}$ de X pour laquelle $X = \cup_{i\in I} U_i$, il existe une partie finie $J \subset I$ telle que $X = \cup_{i\in J} U_i$.

Exemple 2.41 L'intervalle $]0, 1[$ ne vérifie pas la propriété de Borel-Lebesgue. Considérer par exemple la famille d'ouverts $U_i =]1/i, 1 - 1/i[$ $(i \geq 2)$.

Par passage au complémentaire on obtient une autre formulation, équivalente, de la propriété de Borel-Lebesgue.

Remarque 2.42 Propriété duale de Borel-Lebesgue

Un espace métrique X satisfait la propriété de Borel-Lebesgue si et seulement si, pour toute famille $(F_i)_{i\in I}$ de fermés de X dont l'intersection $\cap_{i\in I} F_i$ est vide, on peut extraire une sous-famille finie $(F_i)_{i\in J}$ (avec $J \subset I$ fini) d'intersection $\cap_{i\in J} F_i$ vide.

Comme annoncé plus haut, nous donnons maintenant un nouveau critère de compacité, cette fois-ci en termes de recouvrements.

Théorème 2.43 Compacité : la propriété de Borel-Lebesgue

Un espace métrique X est compact si et seulement si il vérifie la propriété de Borel-Lebesgue.

Preuve Supposons l'espace métrique X compact. Soit $(U_i)_{i\in I}$ un recouvrement ouvert de X. Soit $r > 0$ tel que toute boule de rayon r dans X soit incluse dans l'un des ouverts U_i du recouvrement (lemme d'uniformité 2.37). Soit $X = \cup_{j=1}^k B(x_j, r)$ un recouvrement ouvert fini de X par des boules de rayon r (lemme d'approximation finie 2.38). Chacune des boules $B(x_j, r)$

est incluse dans l'un des ouverts $U_{i(j)}$ du recouvrement ($1 \leq j \leq k$), et le résultat suit.

Supposons maintenant que l'espace métrique X vérifie la propriété de Borel-Lebesgue. Soit (x_n) une suite de points de X. Rappelons que l'ensemble des valeurs d'adhérence de cette suite est

$$\mathcal{V} = \bigcap_{n \in \mathbb{N}} \overline{\{u_k \mid k \geq n\}}$$

(lemme 1.30). Chacun des fermés $F_n = \overline{\{u_k \mid k \geq n\}}$ est non vide. La famille (F_n) est décroissante. Il suit donc de la propriété duale de Borel-Lebesgue (remarque 2.42) que $\mathcal{V}$ est non vide. □

Ce dernier argument nous permet plus généralement d'énoncer la proposition suivante.

Proposition 2.44 *Soit $(K_n)_{n\in\mathbb{N}}$ une famille décroissante de parties compactes non vides d'un espace métrique X. Alors l'intersection $\cap_{n\in\mathbb{N}} K_n$ est compacte et non vide.*

Preuve Pour $n \geq 1$, les K_n sont des parties compactes, donc fermées, du compact K_0. Si l'intersection de cette famille de fermés est vide, une sous-famille finie est d'intersection vide (2.42). La conclusion suit de ce que, la famille étant décroissante, toute intersection finie des K_n est non vide. Enfin, $\cap_{n\in\mathbb{N}} K_n$ est fermée dans K_0 donc compacte. □

La remarque suivante sera bien utile dans la pratique, lorsqu'il s'agira de montrer qu'une partie Y d'un espace métrique X est compacte.

Corollaire 2.45 *Soient X un espace métrique, et $Y \subset X$ une partie de X. Alors l'espace métrique Y est compact (pour la métrique induite, bien sûr) si et seulement si, de tout recouvrement de Y par des ouverts de X, on peut extraire un sous-recouvrement fini.*

Preuve Conséquence immédiate du critère de compacité de Borel-Lebesgue, et de la description des ouverts de la topologie induite. □

Nous avons donc maintenant à notre disposition deux critères de compacité de nature bien différente. Pour montrer qu'un espace métrique est compact, on pourra donc utiliser l'un ou l'autre critère, selon ce qui nous semble plus commode. L'exemple suivant peut se traiter avec Bolzano-Weierstrass, mais la rédaction est un peu délicate à mettre en place. Il est immédiat lorsqu'on utilise Borel-Lebesgue.

Exemple 2.46 Une suite et sa limite
Soient X un espace métrique, et (x_n) une suite convergente d'éléments de X de limite α. Alors $Y = \{x_n \mid n \in \mathbb{N}\} \cup \{\alpha\} \subset X$ est compact (pour la topologie induite).

Exercice 2.47 Donner de nouvelles démonstrations du théorème 2.10 et de la proposition 2.17 en utilisant le critère de Borel-Lebesgue.

Un espace topologique est dit compact lorsqu'il vérifie la propriété de Borel-Lebesgue. Hors du cadre métrique aucune des propriétés Borel-Lebesgue ou Bolzano-Weierstrass n'implique l'autre.

F * Fractales *

En bonus, une jolie application de la compacité : la construction de fractales, c'est-à-dire d'objets auto-similaires dans $\mathbb{R}^p$.

Dans tout le paragraphe, $\mathbb{R}^p$ sera muni d'une norme (par exemple de la norme euclidienne). On se donne une famille finie

$$\mathcal{F} = \{f_1, \cdots, f_k\}$$

d'applications $f_i : \mathbb{R}^p \to \mathbb{R}^p$. On suppose que chaque f_i est strictement contractante, c'est-à-dire lipschitzienne de constante de lipschitz $k_i < 1$ (définition 3.18).

F.1 Existence et unicité

Théorème 2.48 *Il existe un unique compact non vide $K_{\mathcal{F}} \subset \mathbb{R}^p$ qui vérifie la propriété*

$$K_{\mathcal{F}} = \bigcup_{i=1}^{k} f_i(K_{\mathcal{F}})\,. \tag{2.1}$$

Lorsque $\mathcal{F} = \{f_1\}$, ce résultat préfigure le théorème du point fixe de Picard 3.20. Nous verrons que, dans ce cas, le compact $K_{\mathcal{F}}$ est un singleton.

Dans les exemples ci-dessous (paragraphe F.2), les applications f_i seront des applications affines, et plus précisément des similitudes de $\mathbb{R}^p$ euclidien.

Le compact $K_{\mathcal{F}}$ que nous allons construire sera inclus toute boule $B_f(0, R)$ vérifiant la propriété de stabilité ci-dessous.

Lemme 2.49 *Pour R suffisament grand, la boule fermée $B_f(0,R)$ vérifie*

$$\cup_{i=1}^k f_i(B_f(0,R)) \subset B_f(0,R)\,.$$

Preuve On choisit une constante $c<1$ pour laquelle chacune des applications strictement contractantes f_i est c-lipschitzienne, et on pose $M = \sup_{i=1}^k \|f_i(0)\|$. Il suit que, pour tous $x\in\mathbb{R}^p$ et $1\le i\le k$, on a

$$\|f_i(x)\| \le c\|x\| + M\,.$$

Prendre alors R assez grand, de sorte que $M \le (1-c)\,R$. □

Preuve du théorème 2.48
On introduit l'espace métrique produit $\mathcal{C}_k := \{1,\cdots,k\}^{\mathbb{N}}$. C'est un espace compact. On va construire le compact $K_{\mathcal{F}}$ comme image de $\mathcal{C}_k$ par une application continue $\ell : \mathcal{C}_k \to \mathbb{R}^p$. La construction de cette application ℓ dépendra, *a priori*, du choix d'un point $x\in\mathbb{R}^p$.

Soit R comme dans le lemme 2.49. On fixe un point $x\in\mathbb{R}^p$ tel que $\|x\|\le R$. A chaque élément $I=(i_1,\cdots,i_n,\cdots)\in\mathcal{C}_k$ on associe la suite de points $x_n(I)\in B_f(0,R)$ définie par $x_0(I)=x$ et, pour $n\ge 1$, par

$$x_n(I) = f_{i_1}\circ f_{i_2}\cdots\circ f_{i_{n-1}}\circ f_{i_n}(x)\,.$$

On vérifie facilement par récurrence que $d(x_n(I),x_{n+1}(I))\le 2c^n\,R$, et donc que la suite $(x_n(I))$ est de Cauchy. On note $\ell_x(I)=\lim_{n\to\infty} x_n(I)\in B_f(0,R)$ sa limite. L'application

$$\ell_x : I\in\mathcal{C}_k \to \ell_x(I)\in\mathbb{R}^p$$

est continue. En effet rappelons que dire que les deux suites $I,J\in\mathcal{C}_k$ sont proches, c'est dire qu'elles ont les mêmes premiers termes, autrement dit que $(i_1,\cdots,i_n)=(j_1,\cdots,j_n)$ avec n d'autant plus grand que I et J sont proches (proposition 1.65). Puisque l'application $f_{i_1}\circ\cdots\circ f_{i_n}$ est continue, il suit de la discussion précédente qu'on a

$$\ell_x(I) = f_{i_1}\circ\cdots\circ f_{i_n}(y) \text{ et } \ell_x(J) = f_{i_1}\circ\cdots\circ f_{i_n}(z)$$

pour deux points $y,z\in B_f(0,R)$, d'où $d(\ell_x(I),\ell_x(J))\le 2c^n\,R$. L'application ℓ_x est donc bien continue.

Existence Définissons $K_x=\ell_x(\mathcal{C}_k)$. C'est une partie compacte de $\mathbb{R}^n$ comme image du compact $\mathcal{C}_k$ par l'application continue ℓ_x. Montrons qu'il satisfait la propriété (2.1).

Soit $y\in K_x$. Par construction il existe $I\in\mathcal{C}_k$ tel que $y=\ell_x(I)$.

Soit $1\le j\le k$. Par continuité de f_j, on a $f_j(y)=\ell_x(j,I)$ où $(j,I)=(j,i_1,\cdots,i_n,\cdots)$ désigne la suite concaténée de j et de I. On a donc bien $f_j(K_x)\subset K_x$.

On écrit ensuite $I = (i_1, I')$ où $I' = (i_2, \cdots, i_n, \cdots) \in \mathcal{C}_k$. On a donc, par continuité de l'application f_{i_1}, $y = \ell_x(I) = f_{i_1}(\ell_x(I')) \in f_{i_1}(K_x)$. Ceci montre que $K_x \subset \bigcup_{i=1}^k f_i(K_x)$.

Unicité Observons pour commencer que deux choix de points $x, y \in \mathbb{R}^p$ induisent le même compact $K_x = K_y$, puisque

$$d(f_{i_1} \circ \cdots \circ f_{i_n}(x), f_{i_1} \circ \cdots \circ f_{i_n}(y)) \leq c^n \, d(x, y) \,. \tag{2.2}$$

Soit maintenant K un compact non vide vérifiant la propriété (2.1). Il suit que, pour chaque $x \in K$, on a $K_x \subset K$.

Réciproquement, soit $y \in K = \cup_{i=1}^k f_i(K)$. Il existe donc un indice i_1 et un point $x_1 \in K$ avec $y = f_{i_1}(x_1)$. On réitère le procédé et on obtient une suite $I = (i_1, \cdots, i_n, \cdots) \in \mathcal{C}_k$ et une suite de points $x_n \in K \subset B_f(0, R)$ tels que, pour tout entier $n \geq 1$, on ait $y = f_{i_1} \circ \cdots \circ f_{i_n}(x_n)$. Soit $x \in \mathbb{R}^n$ un point quelconque. Il suit alors de (2.2) que $y = \ell_I(x)$ et donc que $K \subset K_x$. $\square$

F.2 Comment le construire ?

Soit $K \subset \mathbb{R}^p$ un compact non vide tel que $\cup_{i=1}^k f_i(K) \subset K$. On peut prendre par exemple $K = B_f(0, R)$, où $B_f(0, R)$ est l'une des boules données par le lemme 2.49. On pose, pour chaque $n \geq 1$,

$$K_n = \bigcup_{(i_1, \cdots, i_n) \in \{1, \cdots, k\}^n} f_{i_1} \circ \cdots \circ f_{i_n}(K) \,.$$

Puisque $\cup_{i=1}^k f_i(K) \subset K$, la famille de compacts K_n est décroissante. Comme K est non vide, chaque K_n est également non vide. Il suit donc de la proposition 2.44 que l'intersection des K_n est non vide. La démonstration précédente assure que

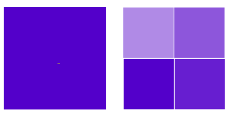
$$K_{\mathcal{F}} = \bigcap_{n \geq 1} K_n \,.$$

Poursuivons par quelques exemples.

- Un premier exemple, pas palpitant mais destiné à vous rassurer, est fourni par le carré $K = [0, 1]^2$ qui est la fractale associé à la famille des quatre applications affines définies sur $\mathbb{R}^2 \simeq \mathbb{C}$ par

$$f_1(z) = z/2,\ f_2(z) = z/2 + 1/2,\ f_3(z) = z/2 + i/2 \text{ et } f_4(z) = z/2 + (1+i)/2 \,.$$

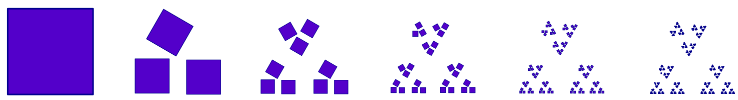
Dans cet exemple, on a $K_n = K$ pour tout entier $n \geq 1$.



• On continue par une variante de ce premier exemple, mais qui donne cette fois-ci une fractale digne de ce nom ! On a choisi d'illustrer notre construction par la famille d'applications affines définies sur $\mathbb{R}^2$ euclidien, que l'on identifie à la droite complexe $\mathbb{C}$, par

$$f_1(z) = 0.4\,z,\ f_2(z) = 0.4\,z + 0.6 \text{ et } f_3(z) = 0.4\,e^{i\pi/3}(z - i) + (i + 1/3)$$

avec pour compact initial $K = [0,1]^2$.

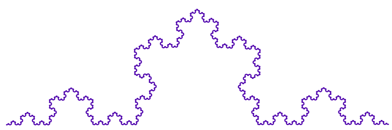


Construction d'une fractale : $K_{\mathcal{F}} = \cap K_n$

• Le Cantor-1/3 est le compact de $\mathbb{R}$ associé aux deux transformations affines définies pour $t \in \mathbb{R}$ par $f_1(t) = t/3$ et $f_2(t) = 2/3 + t/3$. On le notera $\mathcal{C}_{1/3} \subset \mathbb{R}$. On dessine ci-dessous les itérées associées au compact initial $[0,1]$.

Le Cantor $\mathcal{C}_{1/3}$

• On peut aussi mentionner la courbe de von Koch qui sera construite, par une autre méthode, dans le paragraphe 3.F. On peut vérifier que cette courbe est le compact associé aux applications affines définies sur $\mathbb{R}^2 \simeq \mathbb{C}$ par $f_1(z) = z/3$, $f_2(z) = e^{i\pi/3}z/3 + 1/3$, $f_3(z) = e^{-i\pi/3}z/3 + 1/2 + i\sqrt{3}/2$ et $f_4(z) = z/3 + 2/3$.



La courbe de von Koch

F.3 Fractales homéomorphes à l'espace de Cantor

On reprend les notations et les hypothèses du théorème 2.48. Dans certains cas, par exemple dans les deux exemples médians ci-dessus, le compact $K_{\mathcal{F}}$ est homéomorphe à l'espace de Cantor $\mathcal{C} = \{0,1\}^{\mathbb{N}}$.

Théorème 2.50 *On suppose de plus que*

1. *chaque application* $f_i : \mathbb{R}^n \to \mathbb{R}^n$ *est injective*
2. *il existe un compact non vide* $K_0 \subset \mathbb{R}^n$ *tel que*
   — $\cup_{i=1}^k f_i(K_0) \subset K_0$
   — *pour* $1 \leq i < j \leq k$*, on a* $f_i(K_0) \cap f_j(K_0) = \emptyset$.

Alors $K_{\mathcal{F}}$ *est homéomorphe à l'espace de Cantor* $\mathcal{C}$.

Ce résultat sera une conséquence immédiate du lemme suivant, et de ce qui a déjà été dit.

Lemme 2.51 *Soit* $k \geq 2$ *un entier. L'espace produit* $\mathcal{C}_k = \{1, \cdots, k\}^{\mathbb{N}}$ *est homéomorphe à l'espace de Cantor* $\mathcal{C} = \{0,1\}^{\mathbb{N}}$.

Preuve Démontrons que $\mathcal{C}$ est homéomorphe à $\mathcal{C}_3$. Ecrivons plutôt, pour plus de clarté, $\mathcal{C}_3 = \{a,b,c\}^{\mathbb{N}}$ et considérons l'application

$$\varphi : \{0,1\}^{\mathbb{N}} \to \{a,b,c\}^{\mathbb{N}}$$

définie comme suit. Un élément x de $\mathcal{C}$ est une suite infinie de 0 et de 1 ; on le considère comme une chaîne de trois caractères 0, 10 et 11 et on lui associe par φ l'élément de $\mathcal{C}$ obtenu en remplaçant le caractère 0 par a, le caractère 10 par b et le caractère 11 par c. Par exemple,

$$\varphi(0,0,1,0,0,1,1,1,1,0,1,0,1,\cdots) = (a,a,b,a,c,c,a,b,b \text{ ou } c,\cdots)\,.$$

L'application $\varphi : \{0,1\}^{\mathbb{N}} \to \{a,b,c\}^{\mathbb{N}}$ est bijective et continue. C'est donc un homémorphisme entre ces deux espaces compacts.

Le même procédé permet de construire un homéomorphisme entre $\mathcal{C}_k$ et $\mathcal{C}_{k+1}$ pour tout entier $k \geq 2$. □

Preuve du théorème 2.50

On reprend la preuve du théorème 2.48 et on fixe un point $x \in K_0$. Sous l'hypothèse faite, l'application continue

$$I \in \mathcal{C}_k = \{1, \cdots, k\}^{\mathbb{N}} \to \ell(I) \in \mathbb{R}^n$$

est désormais injective, donc réalise un homéomorphisme sur son image $K_{\mathcal{F}}$.
□

3. Espaces métriques complets

Contrairement à la compacité ou à la connexité, qui sont des notions topologiques, la complétude est une notion métrique.

A Suites de Cauchy

Définition 3.1 *Soit (X, d) un espace métrique. Une suite (x_n) à valeurs dans X est de Cauchy lorsque, pour tout $\varepsilon > 0$, il existe un rang N tel que l'on ait*

$$d(x_n, x_p) \leq \varepsilon$$

lorsque $n, p \geq N$.

Noter qu'une suite de Cauchy est bornée.

Exemple 3.2 Une suite convergente est de Cauchy.

La réciproque est fausse en général, comme le montre l'exemple de la suite de terme général $x_n = 1/n$ ($n \geq 1$) dans l'espace $X =]0, 1[$. Ceci nous amène à la notion d'espace métrique complet.

Définition 3.3 *Un espace métrique X est complet lorsque toute suite de Cauchy dans X est convergente.*

Donnons immédiatement un premier exemple.

Exemple 3.4 L'espace métrique $\mathbb{R}$ est complet.

Si $\mathbb{R}$ a été construit comme "complété" de $\mathbb{Q}$, la complétude de $\mathbb{R}$ est quasi immédiate. Si $\mathbb{R}$ a été construit par les coupures, il vérifie la propriété de la borne supérieure, et on en déduit la complétude, voir le corollaire 3.9.

Remarque 3.5 *Les espaces $\mathbb{R}$ et $]0, 1[$ sont homéomorphes. Le premier est complet, mais pas le second. La complétude n'est donc pas une propriété topologique, mais une propriété métrique.*

L'exercice suivant complète cette remarque : la notion de suite convergente est une notion topologique, tandis que la notion de suite de Cauchy est une notion métrique.

Exercice 3.6 Soient d_1 et d_2 deux distances sur l'ensemble X. Soit (x_n) une suite d'éléments de X.

1. On suppose que les distances d_1 et d_2 sont topologiquement équivalentes. Supposons (x_n) de Cauchy dans (X, d_1). Est-elle toujours de Cauchy dans (X, d_2) ?
   On pourra prendre pour exemple la droite réelle $\mathbb{R}$, munie de la distance usuelle d, et de la distance définie par $\delta(s,t) = |\arctan s - \arctan t|$.
2. On suppose maintenant que d_1 et d_2 sont Lipschitz équivalentes.
   (a) Montrer que (x_n) est de Cauchy dans (X, d_1) si et seulement si elle est de Cauchy dans (X, d_2).
   (b) Montrer que (X, d_1) est complet si et seulement si (X, d_2) l'est.
3. Soient N_1 et N_2 deux normes équivalentes sur l'espace vectoriel E. Montrer que (E, N_1) est complet si et seulement si (E, N_2) l'est.

B Premiers exemples d'espaces complets

Nous donnons quelques exemples d'espaces métriques complets, ainsi que des méthodes pour en construire de nouveaux.

Commençons par une remarque immédiate.

Remarque 3.7 *Une suite de Cauchy à valeurs dans un espace métrique quelconque converge dès qu'elle possède au moins une valeur d'adhérence.*

On en déduit immédiatement la proposition suivante.

Proposition 3.8 *Un espace métrique compact est complet.*

Venons-en à la complétude de $\mathbb{R}$, annoncée plus haut.

Corollaire 3.9 *Soient $k \geq 1$ et N une norme sur $\mathbb{R}^k$. L'espace métrique $(\mathbb{R}^k, N)$ est complet.*

Preuve Soit (x_n) une suite de Cauchy dans $(\mathbb{R}^k, N)$. Cette suite est bornée. Il existe donc $R > 0$ telle que la suite prenne ses valeurs dans la boule fermée $B_f(0, R)$, qui est compacte. Cette suite converge donc. $\square$

On étudie maintenant les parties complètes d'un espace métrique.

Lemme 3.10 *Soient X un espace métrique et $Y \subset X$ une partie de X.*

Si Y est complet (muni de la distance induite), alors $Y \subset X$ est une partie fermée de X.

Si X est complet, toute partie fermée $Y \subset X$ est encore complète.

Preuve Soit (y_n) une suite d'éléments de Y qui converge vers une limite $\alpha \in X$. C'est donc une suite de Cauchy. Puisque Y est supposé complet, cette suite converge dans Y, donc $\alpha \in Y$ (unicité de la limite). Ceci prouve que Y est fermé dans X.

Soit (y_n) une suite de Cauchy d'éléments de Y. Cette suite est *a fortiori* une suite de Cauchy dans X. Puisque X est complet, la suite (y_n) y converge vers $\alpha \in X$. Puisque Y est fermé dans X, on a $\alpha \in Y$. Il suit que (y_n) converge dans Y. □

On en déduit en particulier le résultat suivant, qui sera notamment utilisé dans la preuve du théorème de Riesz 2.33, voir le chapitre 4.

Proposition 3.11 *Soient (E, N) un espace vectoriel normé, et $F \subset E$ un sous-espace vectoriel de dimension finie. Alors F est fermé dans E.*

Preuve La norme N fait de l'espace vectoriel de dimension finie F un espace complet (corollaire 3.9) et F est donc fermé dans E (lemme 3.10). □

Mentionnons pour conclure le fait qu'un produit d'espaces complets est complet. Attention, comme on l'a dit plus haut, il faut être prudent et préciser pour quelle distance (remarque 3.5).

Lemme 3.12 *Soit $(X_1, d_1), \cdots, (X_n, d_n)$ des espaces métriques complets. L'espace produit $X = \prod_{i=1}^n X_i$ muni de la distance*

$$\delta_\infty((x_1, \cdots, x_n), (y_1, \cdots, y_n)) = \sup_{i=1}^{n} d_i(x_i, y_i)$$

est encore complet.

Preuve On vérifie facilement qu'une suite $(x(k))$ dans (X, δ_∞) est de Cauchy si et seulement si chacune de ses composantes $(x_i(k))$ est une suite de Cauchy dans (X_i, d_i). Le résultat suit car une suite du produit converge si et seulement si chacune de ses composantes converge. □

Remarque 3.13 Noter que résultat persiste, avec la même démonstration, si l'on munit le produit X de la distance $\delta_1((x_1, \cdots, x_n), (y_1, \cdots, y_n)) = \sum_{i=1}^n d_i(x_i, y_i)$. D'ailleurs les distances δ_1 et δ_∞ sont Lipschitz équivalentes, voir l'exercice 3.6.

C Espaces de fonctions

Nous examinons maintenant la complétude de certains espaces fonctionnels.

Définition 3.14 *Soient X et Y deux espaces métriques. On désignera par $\mathcal{B}(X,Y)$ l'espace des applications $f : X \to Y$ qui sont bornées, c'est-à-dire dont l'image $f(X) \subset Y$ est bornée. On notera $C_b(X,Y)$ l'espace des applications de X dans Y qui sont continues et bornées. Sauf mention expresse du contraire, ces espaces de fonctions seront munis de la distance*

$$d(f,g) = \sup_{x\in X} d(f(x), g(x))$$

(dite distance de la convergence uniforme).

N.B. On se permettra de noter par la même lettre d la distance sur chacun des espaces X, Y, $\mathcal{B}(X,Y)$ et $C_b(X,Y) \subset \mathcal{B}(X,Y)$.

Remarque 3.15 Une suite d'applications (f_n) converge vers f dans $\mathcal{B}(X,Y)$ si et seulement si $f_n \to f$ uniformément sur X.

Lorsque X est compact, toute application continue est bornée. L'espace $C(X,Y)$ des applications continues de X dans Y coïncide alors avec $C_b(X,Y)$.

Théorème 3.16 *Soient X et Y deux espaces métriques. On suppose que Y est complet. Les espaces $\mathcal{B}(X,Y)$ et $C_b(X,Y)$, munis de la distance de la convergence uniforme, sont deux espaces métriques complets.*

Corollaire 3.17 *Lorsque X est compact et Y est complet, l'espace $C(X,Y)$ est un espace métrique complet.*

Preuve Démontrons que $\mathcal{B}(X,Y)$ est complet. Soit donc (f_n) une suite de Cauchy dans $\mathcal{B}(X,Y)$.

On observe pour commencer que la suite (f_n) converge simplement. En effet, pour tout point $x \in X$, la suite des images $(f_n(x))$ est de Cauchy dans Y complet, donc converge. On note $f(x) = \lim_{n\to\infty} f_n(x)$.

Montrons que la limite $f : X \to Y$ est une application bornée. Fixons un point de référence $y_0 \in Y$. Pour tout entier p, soit r_p un rayon tel que $f_p(X) \subset B(y_0, r_p)$. Il existe $N \in \mathbb{N}$ tel que pour tout $n \geq N$, on ait $d(f_N, f_n) \leq 1$. L'image $f(X)$ est donc incluse dans la boule fermée de centre y_0 et de rayon $R = r_N + 1$.

Il nous reste à constater que la suite (f_n) converge uniformément vers f. Soit $\varepsilon > 0$. Il existe un rang $N \in \mathbb{N}$ tel que, pour tous $p \geq N$ et tout point $x \in X$, on ait $d(f_N(x), f_p(x)) \leq d(f_N, f_p) \leq \varepsilon$. Faisons tendre p vers l'infini. Il vient, pour tout $x \in X$, $d(f_N(x), f(x)) \leq \varepsilon$ et donc $d(f_N, f) \leq \varepsilon$: ce qu'on voulait.

Montrons maintenant que $C_b(X,Y) \subset \mathcal{B}(X,Y)$ est complet. Il suffit pour cela de voir que $C_b(X,Y)$ est une partie fermée de $\mathcal{B}(X,Y)$. C'est bien le cas, puisqu'une limite uniforme d'applications continues et encore continue (théorème 1.46). $\square$

D Le théorème de point fixe de Picard

Un point fixe pour une application $f : X \to X$ d'un ensemble dans lui-même est un point $x \in X$ tel que $f(x) = x$.

Un "théorème de point fixe" assure, sous certaines conditions portant sur l'application f et/ou sur l'ensemble X, l'existence d'un point fixe pour $f : X \to X$. Ces théorèmes de point fixe sont utilisés notamment en "analyse fonctionnelle", c'est-à-dire dans un contexte où l'espace X est un espace de fonctions, et permettent de résoudre des équations fonctionnelles – typiquement, mais pas exclusivement, des équations différentielles.

La démonstration du théorème de Picard ci-dessous est très élémentaire. Vous en verrez néanmoins cette année plusieurs applications frappantes : le théorème d'inversion locale (théorème 8.4) ou des fonctions implicites et le théorème de Cauchy-Lipschitz (existence et unicité de solutions pour les équations différentielles).

Définition 3.18 *Soit (X, d) un espace métrique. Une application $f : X \to X$ est strictement contractante lorsqu'elle est k-lipschitzienne, où k est une constante telle que $k < 1$: on a, pour tous points $x, y \in X$, l'inégalité*

$$d(f(x), f(y)) \leq k\, d(x, y)\,.$$

Exercice 3.19 Il ne suffit pas, pour assurer que f est strictement contractante, de demander que $d(f(x), f(y)) < d(x, y)$ pour x et y distincts. Trouver un exemple.

Théorème 3.20 du point fixe de Picard

Soient X un espace métrique complet et $f : X \to X$ une application strictement contractante. Alors f possède un unique point fixe $a \in X$.

De plus, pour tout point $x_0 \in X$, le point fixe est obtenu comme limite des itérés de x_0 sous f, c'est-à-dire que $a = \lim_{n\to\infty} f^n(x_0)$. On peut même préciser la vitesse de convergence :

$$d(f^n(x_0), a) \leq \frac{k^n}{1-k}\, d(x_0, f(x_0))\,. \tag{3.1}$$

Preuve L'unicité du point fixe suit de ce qu'on a, pour tous $a \neq b$, l'inégalité

$$d(f(a), f(b)) \leq k\, d(a, b) < d(a, b)\,.$$

Démontrons l'existence. Soit donc $x_0 \in X$ et définissons récursivement, pour tout entier $n \geq 1$, $x_n = f^n(x_0) = f(x_{n-1})$. On a, pour tout $n \geq 1$,

$$d(x_n, x_{n+1}) \leq k\, d(x_{n-1}, x_n) \leq k^n\, d(x_0, x_1)\,.$$

Il suit alors de l'inégalité triangulaire que l'on a, pour tous entiers $0 \le n \le m$,

$$d(x_n, x_m) \le \Big(\sum_{j=n}^{m-1} k^j\Big)\, d(x_0, x_1) \le k^n\, \frac{d(x_0, x_1)}{1-k} \tag{3.2}$$

et donc que la suite (x_n) est de Cauchy. Puisque X est complet, cette suite converge vers un point $a \in X$. Il nous reste à montrer que a est point fixe de f. Pour cela, on observe que f étant lipschitzienne, elle est *a fortiori* continue. On a donc

$$f(a) = f(\lim_{n\to\infty} x_n) = \lim_{n\to\infty} x_{n+1} = a\,.$$

L'estimation (3.1) suit de (3.2) en fixant n et en faisant tendre m vers ∞. □

E * Complété d'un espace métrique *

Lorsqu'un espace métrique X n'est pas complet, on peut le “compléter” en lui adjoignant des points pour le rendre complet. En d'autres termes, on réalise $X \subset \tilde{X}$ comme sous-espace métrique d'un espace métrique complet $\tilde{X}$. On peut même se débrouiller pour que l'espace complété soit “minimal”, c'est-à-dire pour que $X \subset \tilde{X}$ soit dense. Dans ce cas, il y a unicité de $\tilde{X}$ (aux isométries près), et on parle alors “du” complété de l'espace métrique X.

Exemple 3.21 Le complété de $\mathbb{Q}$ est $\mathbb{R}$.
Le complété de l'intervalle ouvert $]0,1[$ est l'intervalle fermé $[0,1]$.

Exemple 3.22 Munissons l'espace vectoriel $\mathbb{R}[X]$ des polynômes en une variable de la norme

$$\|P\| = \sup_{t\in[0,1]} |P(t)|\,.$$

Son complété est l'espace $(C([0,1],\mathbb{R}), \|\ \|_\infty)$ muni de la norme de la convergence uniforme (cela suit du théorème d'approximation de Weierstrass, voir la remarque 2.22).

Exemple 3.23 Soit $1 \le p < \infty$. Munissons l'espace vectoriel $C([0,1],\mathbb{R})$ des fonctions continues sur l'intervalle $[0,1]$ de la norme

$$\|f\|_p = \Big(\int_0^1 |f(t)|^p\, dt\Big)^{1/p}.$$

Son complété est l'espace L^p muni de la norme $\|\ \|_p$.

Exercice 3.24 **Complétion d'un espace métrique.**
Soient (X, d) un espace métrique, et $C_b(X, \mathbb{R})$ l'espace des fonctions réelles continues bornées sur X, muni de la norme définie par $\| f \| = \sup_{x \in X} |f(x)|$.

1. On fixe $x_0 \in X$. Montrer que l'application $\phi : X \to C_b(X, \mathbb{R})$ définie par

$$\phi(x) = [z \in X \to d(z, x) - d(z, x_0)] \in C_b(X, \mathbb{R})$$

est isométrique (elle préserve les distances).

2. En déduire que (X, d) est isométrique à une partie dense de l'espace métrique complet $\tilde{X} := \overline{\phi(X)} \subset C_b(X, \mathbb{R})$.

Exercice 3.25 **Prolongement des applications uniformément continues**
Soient (X, d) et (Y, δ) deux espaces métriques, $A \subset X$ une partie de X et $f : A \to Y$ une application continue (pour la distance induite).

1. Peut-on toujours prolonger f en une application continue $\tilde{f} : X \to Y$? Démonstration, ou contre-exemple.
2. Montrer que si A est dense dans X, si Y est complet, et si l'application f est uniformément continue sur A (définition 2.16), il existe une unique application continue $\tilde{f} : X \to Y$ qui prolonge f.
3. Exemples.
   (a) Définition de l'intégrale de Riemann des fonctions réglées.
   (b) Définition de la transformée de Fourier sur $L^2(\mathbb{R})$ (théorème de Plancherel).
4. Existence et unicité du complété d'un espace métrique.
   On a vu dans l'exercice précédent que, pour tout espace métrique (X, d), il existe un espace métrique complet $(\tilde{X}, \tilde{d})$ et une application isométrique $i : (X, d) \mapsto (\tilde{X}, \tilde{d})$ d'image dense.
   Montrer que l'espace $\tilde{X}$ et l'application i sont uniques aux isométries près.

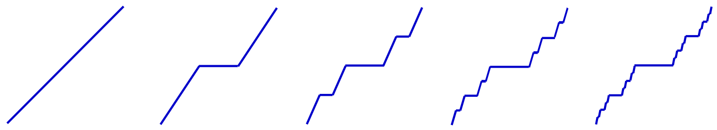
F *Quelques illustrations du théorème de Picard*

Exemple 3.26 **La fonction de Lebesgue, ou escalier du diable**
On travaille dans l'espace

$$X = \{u \in C([0, 1], [0, 1]) \mid u(0) = 0 \ , \ u(1) = 1\}$$

muni de la distance de la convergence uniforme. C'est un espace métrique complet. On considère l'application $F : X \to X$ définie pour toute fonction continue $u : [0, 1] \to [0, 1]$ par

$$\begin{aligned}(Fu)(t) =& u(3t)/2 \text{ lorsque } 0 \leq t \leq 1/3 \\ =& 1/2 \text{ lorsque } 1/3 \leq t \leq 2/3 \\ =& 1/2 + u(3t - 2)/2 \text{ lorsque } 2/3 \leq t \leq 1.\end{aligned}$$



L'escalier du diable : les premières itérées $F^n(u)$ pour $u(t) \equiv t$

L'application F est strictement contractante de rapport $1/2$, et admet donc un unique point fixe $u_0 \in X$, qui est une fonction continue et croissante de l'intervalle $[0,1]$ dans lui-même. Elle est dérivable presque partout (c'est-à-dire sur une réunion dénombrable d'intervalles ouverts disjoints dont la somme des longueurs vaut 1) et a, aux points où elle est dérivable, une dérivée nulle. Pourtant $u_0(0) = 0$ et $u_0(1) = 1$.

Exemple 3.27 La courbe de von Koch (snowflake)

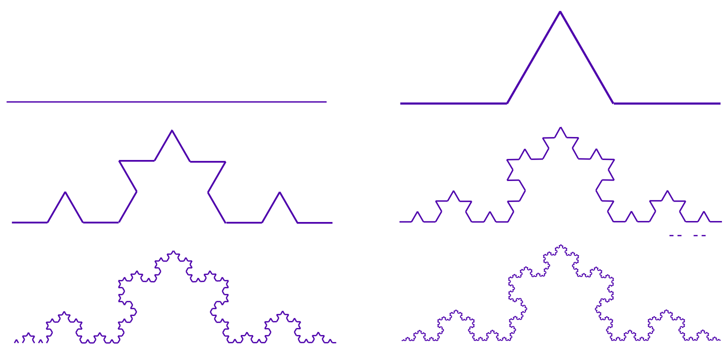
On travaille dans l'espace

$$X = \{u \in C([0,1],[0,1]^2) \mid u(0) = (0,0)\ ,\ u(1) = (1,0)\}$$

muni de la distance de la convergence uniforme. C'est un espace métrique complet. On identifie $\mathbb{R}^2$ à $\mathbb{C}$ et on considère l'application $F : X \to X$ définie pour toute fonction continue $u : [0,1] \to [0,1]$ par

$$\begin{aligned}(Fu)(t) =& u(4t)/3 \text{ lorsque } 0 \le t \le 1/4\\ =& 1/3 + e^{i\pi/3}u(4t-1)/3 \text{ lorsque } 1/4 \le t \le 1/2\\ =& 1/2 + i\sqrt{3}/6 + e^{-i\pi/3}u(4t-2)/3 \text{ lorsque } 1/2 \le t \le 3/4\\ =& 2/3 + u(4t-3)/3 \text{ lorsque } 3/4 \le t \le 1.\end{aligned}$$

L'application F est strictement contractante de rapport $1/3$, et admet donc un unique point fixe $u_0 \in X$, qui est une fonction continue de l'intervalle $[0,1]$ dans le carré $[0,1]^2$. Son image est la courbe de von Koch. C'est une fractale, associée aux quatre transformations affines $f_i : \mathbb{C} \to \mathbb{C}$ définies pour $z \in \mathbb{C}$ par $f_1(z) = z/3$, $f_2(z) = 1/3 + e^{i\pi/3}z/3$, $f_3(z) = 1/2 + i\sqrt{3}/6 + e^{-i\pi/3}z/3$ et $f_4(z) = 2/3 + z/3$.



La courbe de von Koch : les premières itérées $F^n(u)$ pour $u(t) \equiv (t,0)$

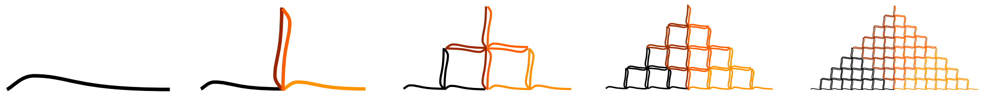
Exemple 3.28 La courbe de Péano

C'est une application continue $u_0 : [0,1] \to \mathbb{R}^2 \simeq \mathbb{C}$ dont l'image est le triangle plein

$$\mathcal{T} = \{x + iy\,,\ 0 \leq y \leq x\,,\ y \leq 1 - x\}\,.$$

Elle est obtenue comme point fixe de la transformation $F : X \to X$, où $X = \{u \in C^0([0,1],\mathbb{C}) \mid u(0) = 0\,,\ u(1) = 1\}$ et $F : X \to X$ est définie par $Fu(t) = s_j(f(4t - j))$ lorsque $\frac{j}{4} \leq t \leq \frac{j+1}{4}$ $(0 \leq j \leq 3)$, où les $s_j : \mathbb{C} \to \mathbb{C}$ sont les similitudes

$$s_0(z) = \frac{z}{2}\,,\ s_1(z) = \frac{1+iz}{2}\,,\ s_2(z) = \frac{1+i-iz}{2}\,,\ s_3(z) = \frac{1+z}{2}\,.$$



La courbe de Péano, dont l'image remplit tout le triangle

4. Espaces vectoriels normés, applications linéaires continues

A Le théorème de compacité de Riesz

Nous sommes maintenant en mesure de démontrer le théorème suivant, qui avait été énoncé en 2.33.

Théorème de compacité de Riesz

Soit $(E, \| \ \|)$ un espace vectoriel normé. La boule unité fermée de E est compacte si et seulement si l'espace E est de dimension finie.

Preuve Supposons que la boule unité fermée $B_f(0,1)$ de $(E, \| \ \|)$ soit compacte. On peut alors la recouvrir par un nombre fini de boules ouvertes $B(x_i, 1/2)$ de rayons $1/2$ et de centres $x_1, \cdots, x_n$.

Notons $F = \mathrm{vect}(x_1, \cdots, x_n) \subset E$ le sous-espace vectoriel de E engendré par les centres de ces boules. Nous allons voir que $E = F$, et est donc bien de dimension finie. Notons $F + B(0,r) = \{f + u \mid f \in F\, ,\ \|u\| < r\}$ pour tout $r > 0$. On a

$$B(0,1) \subset B_f(0,1) \subset \bigcup_{i=1}^{n} B(x_i, 1/2) \subset F + B(0, 1/2)$$

Il suit que $B(0,1/2) \subset F + B(0, 1/2^2)$ (appliquer une homothétie de rapport $1/2$) d'où, par une récurrence élémentaire,

$$B(0,1) \subset \bigcap_{p \in \mathbb{N}} (F + B(0, 1/2^p))\, .$$

Tout élément de la boule unité $B(0,1) \subset E$ est donc limite d'une suite d'éléments de F. Puisque F est un espace vectoriel de dimension finie, la proposition 3.11 assure que $F \subset E$ est fermé. Il suit que la boule unité de E est incluse dans F. Par homogénéité, on conclut que $E = F$ est bien de dimension finie.

La réciproque a été démontrée en 2.27. $\square$

B Applications linéaires continues sur un evn

On commence par donner diverses caractérisations des applications linéaires continues.

Proposition 4.1 *Soient E et F deux espaces vectoriels normés. Soit $u : E \to F$ une application linéaire. Les conditions suivantes sont équivalentes :*

1. *l'application linéaire u est continue*
2. *l'application linéaire u est continue en l'origine*
3. *l'image par u de la boule unité ouverte de E est bornée*
4. *l'image par u de toute partie bornée de E est bornée*
5. *il existe une constante $C > 0$ telle que, pour tout $x \in E$ on ait la majoration*
$$\|u(x)\| \leq C \|x\|$$
6. *l'application linéaire u est lipschitzienne.*

Certaines implications sont vraies pour toute application de E dans F. D'autres dépendent de façon essentielle de la linéarité de u (additivité, homogénéité).

Dans la pratique, on utilisera souvent la condition 5. Voir, en particulier, la définition de la norme d'une application linéaire 4.6 et la remarque 4.7.

Preuve $1 \Rightarrow 2$ est immédiat.
$2 \Rightarrow 3$ Si u est continue en l'origine, il existe un rayon $r > 0$ tel que $u(B_E(0, r)) \subset B_F(0, 1)$. Par homogénéité, il suit que l'image par u de la boule unité ouverte de E est incluse dans la boule $B_F(0, 1/r)$ et donc bornée.
$3 \Rightarrow 4$ Supposons que l'image par u de la boule unité ouverte de E (ou de tout voisinage de l'origine) soit bornée dans F. L'homogénéité de u assure alors que l'image par u de toute partie bornée de E est bornée dans F.
$4 \Rightarrow 5$ Supposons u bornée sur la sphère unité, i.e. qu'il existe $C > 0$ telle que $\|u(x)\| \leq C$ lorsque $\|x\| = 1$. L'homogénéité de u assure alors que, pour tout $x \in E$ non nul, on a

$$\|u(x)\| = \|x\| \, u\Big(\frac{x}{\|x\|}\Big) \leq C \, \|x\| \, .$$

$5 \Rightarrow 6$ Puisque u est linéaire on a alors en effet, pour tous $x, y \in E$, l'estimation
$$\|u(x) - u(y)\| = \|u(x - y)\| \leq C \, \|x - y\| \, .$$

$6 \Rightarrow 1$ Toute application lipschitzienne est continue. $\square$

Proposition 4.2 *Une application linéaire définie sur un espace vectoriel normé de dimension finie, à valeurs dans un espace vectoriel normé quelconque, est toujours continue.*

Preuve Soit $u : \mathbb{R}^n \to F$ une application linéaire à valeurs dans un espace vectoriel normé F. Puisque, en dimension finie, toutes les normes sont équivalentes, il suffit de démontrer que u est continue lorsqu'on munit $\mathbb{R}^n$ de la norme sup, soit $\| \|_\infty$. Désignons par $(e_1, \cdots, e_n)$ la base canonique de $\mathbb{R}^n$. On a alors, pour $x = \sum_{i=1}^n x_i e_i$,

$$\|u(x)\| \leq \sum |x_i| \, \|u(e_i)\| \leq C \, \|x\|_\infty \,,$$

où $C = \sup_{i=1}^n \|u(e_i)\|$. C'est la continuité. □

En dimension infinie, on peut par contre construire des applications linéaires non continues.

Exercice 4.3 On munit l'espace vectoriel $E = C([0,1], \mathbb{R})$ des normes définies par $\|f\|_\infty = \sup_{t\in[0,1]} |f(t)|$ et $\|f\|_1 = \int_0^1 |f(t)| \, dt$.

Soit $c \in [0,1]$ et $u_c : f \in E \to f(c) \in \mathbb{R}$ l'application d'évaluation au point c.

1. L'application u_c est-elle continue sur E muni de $\| \|_\infty$?
2. L'application u_c est-elle continue sur E muni de $\| \|_1$?

Exercice 4.4 Soient E un espace vectoriel normé de dimension infinie, et $(e_a)_{a\in A}$ une base de E. L'existence de cette base (admise) résulte du lemme de Zorn. Construire une forme linéaire $\Lambda : E \to \mathbb{R}$ non continue.

Notation 4.5 *Soient E, F deux espaces vectoriels normés. On désigne par $\mathcal{L}_c(E,F)$ l'espace vectoriel des applications linéaires continues de E dans F.*

Proposition-Définition 4.6 Norme subordonnée

On définit une norme sur l'espace vectoriel $\mathcal{L}_c(E,F)$ des applications linéaires continues de E dans F en posant, pour toute $u \in \mathcal{L}_c(E,F)$,

$$\|u\| = \sup_{x\in E\,,\, \|x\|\leq 1} \|u(x)\| = \sup_{x\in E\,,\, \|x\|=1} \|u(x)\| = \sup_{x\in E\,,\, x\neq 0} \frac{\|u(x)\|}{\|x\|} \,.$$

Preuve Immédiate. □

Remarque 4.7 *Soit $u : E \to F$ linéaire continue. La norme de u est la plus petite constante $C > 0$ pour laquelle $u : E \to F$ est C-lipschitzienne.*

Propriété 4.8 *Soient E, F et G trois espaces vectoriels normés. On munit les espaces $\mathcal{L}_c(E,F)$, $\mathcal{L}_c(F,G)$ et $\mathcal{L}_c(E,G)$ des normes subordonnées. On a alors, pour toutes $u \in \mathcal{L}_c(E,F)$ et $v \in \mathcal{L}_c(F,G)$,*

$$\|v \circ u\| \leq \|v\| \, \|u\| \,.$$

Rappel 4.9 *Soit $\mathbb{K}$ un corps.*

Une $\mathbb{K}$-algèbre $\mathcal{A}$ est un $\mathbb{K}$-espace vectoriel muni d'une application bilinéaire $(a,b) \in \mathcal{A} \to a\,b \in \mathcal{A}$ (loi de "multiplication interne").

Soit $\mathcal{A}$ une $\mathbb{K}$-algèbre sur le corps $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Une norme sur l'espace vectoriel $\mathcal{A}$ est une norme d'algèbre lorsqu'on a $\|ab\| \le \|a\|\,\|b\|$ pour tous $a, b \in \mathcal{A}$.

Corollaire 4.10 *L'algèbre $\mathcal{L}_c(E,E)$ des applications continues de E dans lui-même est ainsi munie d'une norme d'algèbre (la norme de la composée est au plus égale au produit des normes). En particulier, tout choix de norme sur $\mathbb{R}^n$ induit une norme subordonnée sur $\mathrm{M}_n\mathbb{R}$, qui est une norme d'algèbre.*

Il faudra être capable, pour une application linéaire $u : E \to F$ entre deux espaces vectoriels normés, de décider si elle est continue et (si elle l'est) de calculer sa norme.

Observons que, en dimension finie, la norme d'une application continue (définie comme un sup) est toujours atteinte.

Proposition 4.11 *Soient $(\mathbb{R}^n, \|\ \|)$ normé, $(F, \|\ \|)$ un espace vectoriel normé et $u : \mathbb{R}^n \to F$ une application linéaire. Alors il existe un vecteur unitaire $x_0 \in \mathbb{R}^n$ (c'est-à-dire tel que $\|x_0\| = 1$) et tel que*

$$\|u\| = \sup_{x \in E\,,\ \|x\|=1} \|u(x)\| = \|u(x_0)\|\,.$$

Preuve La sphère unité $S = \{x \in \mathbb{R}^n\,,\ \|x\| = 1\}$ est compacte (exemple 2.28). L'application $x \in S \to \|u(x)\| \in [0, \infty[$ est continue, comme composée d'applications continues. Elle atteint donc sa borne supérieure. $\square$

En dimension quelconque, la norme n'est pas toujours atteinte.

Exercice 4.12 On travaille dans l'espace $E = C([0,1], \mathbb{R})$ que l'on munit de la norme de la convergence uniforme.

1. Montrer que l'application linéaire $u : E \to \mathbb{R}$ définie par $u(f) = \int_0^1 f(t)\,dt$ est continue. Déterminer sa norme. Existe-t-il $f_0 \in E$ de norme 1 pour laquelle $|u(f_0)| = \|u\|$?
2. Mêmes questions pour $v : E \to \mathbb{R}$ définie par $v(f) = \int_0^{1/2} f(t)\,dt - \int_{1/2}^1 f(t)\,dt$.

Exercice 4.13 Exemples de normes subordonnées

1. Déterminer la norme N_∞ sur $\mathrm{M}_n\mathbb{R}$ qui est subordonnée à la norme définie sur $\mathbb{R}^n$ par $\|x\|_\infty = \sup_{i=1}^n |x_i|$.
2. Déterminer la norme N_1 sur $\mathrm{M}_n\mathbb{R}$ qui est subordonnée à la norme définie sur $\mathbb{R}^n$ par $\|x\|_1 = \sum_{i=1}^n |x_i|$.

3. On rappelle que le rayon spectral d'une matrice réelle ou complexe est $\rho := \sup |\lambda_i|$, où les λ_i sont ses valeurs propres sur $\mathbb{C}$.
   Montrer que la norme N_2 sur $\mathrm{M}_n\mathbb{R}$, subordonnée à la norme euclidienne canonique, est
   $$N_2(M) = \sqrt{\rho({}^tM\,M)}$$
   où $\rho({}^tM\,M)$ désigne le rayon spectral de la matrice symétrique positive ${}^tM\,M$.

Exercice 4.14 Applications multilinéaires continues
Soient $E_1, \cdots, E_k$ et F des espaces vectoriels normés.

1. Montrer qu'une application k-linéaire
   $$u : E_1 \times \cdots \times E_k \to F$$
   est continue si et seulement si il existe une constante $C > 0$ telle que, pour tout vecteur $(x_1, \cdots, x_k) \in E_1 \times \cdots \times E_k$, on ait
   $$\|u(x_1, \cdots, x_k)\|_F \leq C \prod_{i=1}^{k} \|x_i\|_{E_i} .$$
2. Montrer qu'une application multilinéaire définie sur un produit d'espaces vectoriels de dimensions finies est continue.
3. Montrer enfin que l'expression
   $$\|u\| = \sup\{\|u(x_1, \cdots, x_k)\|_F \ , \ \|x_i\|_{E_i} = 1 \text{ pour tout } i = 1 \cdots k\}$$
   définit une norme sur l'espace vectoriel des applications k-linéaires continues de $E_1 \times \cdots \times E_k$ vers F.

C Espaces de Banach

Définition 4.15 *Un espace de Banach est un espace vectoriel normé complet.*

Exemple 4.16 Tout espace vectoriel normé de dimension finie est un espace de Banach (corollaire 3.9).

Exercice 4.17 La norme sup, définie par $\|f\|_\infty = \sup_{x\in[0,1]} |f(x)|$, fait-elle de l'espace vectoriel $C^1([0,1],\mathbb{R})$ un espace de Banach ?
On pourra se demander si $C^1([0,1],\mathbb{R})$ est une partie fermée de $(C^0([0,1],\mathbb{R}), \|\ \|_\infty)$.

Connaissant un espace de Banach, on peut en construire d'autres.

Soient F un espace vectoriel, et X un espace métrique. L'ensemble $F^X = \{f : X \to F\}$ des applications de X dans F hérite d'une structure d'espace vectoriel. La distance de la convergence uniforme (3.14) sur l'espace vectoriel des applications bornées $\mathcal{B}(X,F)$ provient maintenant d'une norme. En conséquence immédiate du théorème 3.16, on obtient les exemples suivants.

Proposition 4.18 *Soit F un espace de Banach. Soit X un espace métrique.*

1. *L'espace vectoriel $\mathcal{B}(X, F)$ des applications bornées de X dans F, muni de la norme "sup" définie pour $f \in \mathcal{B}(X, F)$ par*
$$\|f\|_\infty = \sup_{x\in X} \|f(x)\|\,,$$
est un espace de Banach.
2. *L'espace vectoriel $C_b(X, F)$ des applications continues bornées sur X à valeurs dans F, muni de la norme sup, est un espace de Banach.*
3. *On suppose X compact. L'espace vectoriel $C(X, F)$ des applications continues sur X à valeurs dans F, muni de la norme sup, est un espace de Banach.*

Exemple 4.19 L'espace vectoriel $\ell^\infty = \mathcal{B}(\mathbb{N}, \mathbb{R})$ des suites réelles bornées, muni de la norme sup, est un espace de Banach.

Un autre exemple extrêmement important.

Proposition 4.20 *Soient E un espace vectoriel normé, et F un espace de Banach. L'espace $\mathcal{L}_c(E, F)$ des applications linéaires continues de E dans F, muni de la norme subordonnée, est un espace de Banach.*

Preuve Soit (u_n) une suite de Cauchy dans $\mathcal{L}_c(E, F)$. Fixons un rayon $R > 0$, et considérons la suite $u_{n,R} : B(0, R) \to F$ des restrictions des u_n à la boule ouverte de rayon R. Pour tous $n, p \in \mathbb{N}$, on a
$$\sup_{x\in B(0,R)} |u_n(x) - u_p(x)| \leq R\,\|u_n - u_p\|\,.$$
Ainsi, la suite $(u_{n,R})$ est de Cauchy dans l'espace de Banach $C_b(B(0, R), F)$ muni de la norme sup, donc converge. On note $u_R : B(0, R) \to F$ sa limite.

On observe que, pour $R_1 < R_2$, l'application u_{R_2} prolonge u_{R_1}. Les applications u_R ($R > 0$) sont donc toutes des restrictions d'une même application $u : E \to F$, et la suite (u_n) converge simplement vers u sur E. La linéarité se conservant par convergence simple, $u : E \to F$ est linéaire. Puisque u est bornée sur la boule unité, il suit de la proposition 4.1 que l'application linéaire u est continue. La convergence uniforme de u_n vers u sur la sphère unité de E assure que u_n converge vers u pour la norme subordonnée. □

En particulier, le dual d'un espace vectoriel normé est toujours un espace de Banach.

Corollaire 4.21 *Soient E un espace vectoriel normé et $E' = \mathcal{L}_c(E, \mathbb{R})$ son dual, c'est-à-dire l'ensemble des formes linéaires continues sur E. L'espace E', muni de la norme subordonnée à celle de E, est un espace de Banach.*

Pour montrer qu'un espace vectoriel normé est de Banach, on utilisera volontiers le critère suivant 4.23. Voir par exemple la preuve du théorème de Riesz-Fischer en Intégration (complétude des espaces L^p pour $1 \leq p \leq \infty$), qui étendra le résultat proposé dans l'exercice 4.24.

On commence par une définition.

Définition 4.22 *Soit E un espace vectoriel normé. Soit (x_n) une suite d'éléments de E. On dit que la série de terme général x_n, soit $\sum x_n$, est :*

- *convergente si la suite $s_p = \sum_{n=0}^{p} x_n$ de ses sommes partielles converge. Dans ce cas, la somme de la série est*

$$\sum_{n\in\mathbb{N}} x_n := \lim_{p\to\infty} \sum_{n=0}^{p} x_n\,;$$

- *absolument convergente si la série à termes positifs $\sum_{n\in\mathbb{N}} \|x_n\|$ converge dans $\mathbb{R}$.*

Théorème 4.23 *Un espace vectoriel normé E est de Banach si et seulement si toute série d'éléments de E absolument convergente est convergente.*

Preuve Supposons que chaque série absolument convergente de l'espace vectoriel normé E soit convergente. Soit (y_n) une suite de Cauchy dans E. On veut montrer que la suite (y_n) converge. Pour cela, il suffit de montrer qu'elle admet une valeur d'adhérence. On veut donc construire une suite extraite de (y_n) qui soit convergente.

La suite (y_n) étant de Cauchy, on peut construire récursivement une suite croissante d'entiers $n_1 < n_2 < \cdots < n_p < \cdots$ tels que, pour tout $n > n_p$, on ait $\|y_n - y_{n_p}\| \leq 1/2^p$. En particulier, on a $\|y_{n_{p+1}} - y_{n_p}\| \leq 1/2^p$, et la série de terme général $x_0 = y_{n_1}$ et $x_p = y_{n_{p+1}} - y_{n_p}$ pour $p \geq 1$ est absolument convergente, donc convergente. Puisque les sommes partielles de la série $\sum x_k$ vérifient

$$\sum_{k=0}^{p} x_k = y_{n_{p+1}}\,,$$

la suite (y_n) admet pour valeur d'adhérence la limite $\sum_{k\in\mathbb{N}} x_k$. Puisque (y_n) est de Cauchy, elle converge donc.

Supposons réciproquement que E est un espace de Banach. Soit $\sum x_n$ une série absolument convergente. Pour $p \in \mathbb{N}$, la somme partielle d'ordre p est $s_p = \sum_{n=0}^{p} x_n$. Pour $0 \leq p < q$, on a

$$\|s_q - s_p\| \leq \sum_{p+1}^{q} \|x_n\|\,.$$

La suite des sommes partielles de la série est donc de Cauchy. Elle converge donc puisque E est complet. □

Exercice 4.24 Utiliser le critère du théorème 4.23 pour montrer que :

1. l'espace ℓ^∞ des suites réelles bornées, muni de la norme $\|(u_n)\|_\infty = \sup |u_n|$, est un espace de Banach. On retrouve ainsi le résultat de l'exemple 4.19.
2. l'espace ℓ^1 des suites réelles sommables, muni de la norme $\|(u_n)\|_1 = \sum |u_n|$, est un espace de Banach.
3. plus généralement, l'espace ℓ^p des suites réelles de puissance p-ième sommable ($1 \leq p < \infty$) muni de la norme $\|(u_n)\|_p = (\sum |u_n|^p)^{1/p}$ est un espace de Banach.

D Exponentielle matricielle

Nous allons, pour conclure ce chapitre, donner deux applications de la partie directe du critère 4.23. La première concerne l'exponentielle matricielle qui joue un rôle central notamment pour l'étude des équations différentielles linéaires autonomes.

Proposition 4.25 Exponentielle de matrices
Soit $M \in \mathrm{M}_n\mathbb{R}$. La série de matrices

$$\sum_{k=0}^{\infty} \frac{M^k}{k!}$$

est convergente. Elle définit une application continue

$$\exp : M \in \mathrm{M}_n\mathbb{R} \to \mathrm{e}^M \in \mathrm{M}_n\mathbb{R}\,.$$

On verra plus tard que l'application exp est de classe C^1, et même C^∞ (voir les théorèmes 7.20 et 9.22).

Preuve • On se préoccupe d'abord de la convergence de la série qui définit l'exponentielle matricielle. On travaille dans $\mathrm{M}_n\mathbb{R}$ que l'on munit d'une norme d'algèbre, par exemple de la norme subordonnée à une norme sur $\mathbb{R}^n$. On alors, pour toute matrice $M \in \mathrm{M}_n\mathbb{R}$ et tout entier $N \in \mathbb{N}$:

$$\sum_{k=0}^{N} \frac{\|M^k\|}{k!} \leq \sum_{k=0}^{N} \frac{\|M\|^k}{k!} = \exp(\|M\|)\,.$$

La série définissant e^M est donc absolument convergente dans l'espace $\mathrm{M}_n\mathbb{R}$, de dimension finie donc complet. Le théorème 4.23 s'applique pour montrer que cette série converge : e^M est donc bien définie.

- On veut maintenant montrer que l'application exp est continue. Notons

$$\mathrm{e}_k : M \in \mathrm{M}_n\mathbb{R} \to M^k \in \mathrm{M}_n\mathbb{R}\,,$$

de sorte que $\exp = \sum_{k=0}^{\infty} \mathrm{e}_k/(k!)$. Cette série d'applications $\sum_{k=0}^{\infty} \mathrm{e}_k/(k!)$ converge uniformément sur les parties bornées de $\mathrm{M}_n\mathbb{R}$: observer en effet qu'on a pour $R > 0$

$$\sup_{M\in\mathrm{M}_n\mathbb{R}\,,\, \|M\|\leq R} \|\mathrm{e}_k(M)\| \leq R^k\,,$$

avec $\sum_{k=0}^{\infty} R^k/(k!) < \infty$. La restriction de l'application exponentielle à chaque boule $B(0,R) \subset \mathrm{M}_n\mathbb{R}$ est donc continue (théorème 1.46). La continuité étant une propriété locale, il suit que l'application exponentielle est continue sur $\mathrm{M}_n\mathbb{R}$. $\square$

Remarque 4.26 *La fin de l'argument utilise le fait que "la continuité est une propriété locale". Soit $f : X \to Y$ une application entre deux espaces métriques. Si f est continue au voisinage de chaque point (c'est-à-dire tout point $x \in X$ possède un voisinage V_x tel que la restriction $f : V_x \to Y$ soit continue), alors f est continue.*

Exercice 4.27 Soit $f : X \to Y$ une application entre deux espaces métriques. Soit $X = \cup_{i\in I} A_i$ un recouvrement de X par des parties quelconques. On suppose que les restrictions $f_{|A_i} : A_i \to Y$ sont toutes continues. Est-ce que cela suffit à assurer que f est continue ?

E Groupe des endomorphismes d'un Banach

Une seconde illustration du théorème 4.23 concerne l'étude de l'ensemble $\mathcal{GL}_c(E)$ des isomorphismes bicontinus d'un espace de Banach.

Définition 4.28 *Soit E un espace vectoriel normé. Soit $\mathcal{GL}_c(E)$ le groupe des isomorphismes bicontinus de E, c'est-à-dire le groupe des applications linéaires continues $u \in \mathcal{L}_c(E,E)$, bijectives et dont la réciproque est une application (linéaire) continue.*

Exemple 4.29 Soit E un espace vectoriel normé de dimension finie. Toute application linéaire bijective $u : E \to E$ est continue ainsi que sa réciproque. Autrement dit, on a dans ce cas l'égalité $GL(E) = \mathcal{GL}_c(E)$.

De plus une application linéaire $u : E \to E$ est injective si et seulement si elle est surjective, si et seulement si elle est bijective.

On commence par se convaincre que, en dimension quelconque, les choses ne sont pas si simples.

Exercice 4.30 1. On définit deux applications linéaires $S : \ell^\infty \to \ell^\infty$ (le shift) et $T : \ell^\infty \to \ell^\infty$ sur l'espace de Banach $(\ell^\infty, \| \|_\infty)$ des suites réelles bornées en posant

$$S((x_1, x_2, \cdots, x_n, \cdots)) = (x_2, x_3, \cdots, x_n, \cdots)$$
$$T((x_1, x_2, \cdots, x_n, \cdots)) = (0, x_1, x_2, \cdots, x_n, \cdots)\,.$$

Montrer que S et T sont deux applications linéaires continues dont on déterminera la norme. Montrer que S est surjective non injective, que T est injective non injective, et que $S \circ T = \mathrm{Id}$. Déterminer $T \circ S$.

2. On munit le sous-espace vectoriel $\mathcal{F} \subset \ell^\infty$ des suites à support fini (c'est-à-dire nulles à partir d'un certain rang) de la norme $\| \|_\infty$. Soit

$$R : (u_n)_{n\in\mathbb{N}^*} \in \mathcal{F} \to (u_n/n)_{n\in\mathbb{N}^*} \in \mathcal{F}\,.$$

(a) Montrer que R est une application linéaire continue. Déterminer sa norme.

(b) Montrer que l'application R est bijective. L'application linéaire réciproque $R^{-1} : \mathcal{F} \to \mathcal{F}$ est-elle continue ?

Remarque 4.31 *Le théorème de l'isomorphisme de Banach (conséquence du théorème de Baire) affirme néanmoins que, si E est un espace de Banach, toute application linéaire continue et bijective $u : E \to E$ est bicontinue.*

L'application identité $\mathrm{Id} : E \to E$ est un endomorphisme continu de l'espace vectoriel normé E. On va voir ci-dessous que, lorsque E est de Banach, une petite perturbation de l'identité est encore un endomorphisme continu de E.

Exercice 4.32 Soit E un espace de Banach.

1. Soit $u \in \mathcal{L}_c(E, E)$. On suppose que $\|u\| < 1$ et on pose $v := \mathrm{Id} - u$.

   (a) Montrer que la série $\sum_{n=0}^{\infty} u^n$ est convergente dans $\mathcal{L}_c(E, E)$. On note w sa somme.

   (b) Montrer que $(\mathrm{Id} - u) \circ w = w \circ (\mathrm{Id} - u) = \mathrm{Id}$. En déduire que $\mathrm{Id} - u$ appartient à $\mathcal{GL}_c(E)$.

2. Montrer alors que $\mathcal{GL}_c(E) \subset \mathcal{L}_c(E, E)$ est ouvert.

5. Connexité

Les raisonnements “par connexité” permettent de montrer qu’une propriété est vraie sur tout un espace métrique connexe, sachant qu’elle est vraie en un point de cet espace et que l’ensemble des points où elle est satisfaite est à la fois ouvert et fermé (argument de “passage du local au global”).

Vous verrez des illustrations de ce type d’arguments notamment dans la proposition 5.15, lors de la démonstration du théorème des accroissements finis 6.13, de son corollaire 6.17 ou encore du théorème de Cauchy-Lipschitz. Une autre application importante, le “principe du prolongement analytique”, sera vue dans le cours de fonctions holomorphes .

A Espaces connexes

Rappel 5.1 *Une partition $X = \sqcup_{i \in I} A_i$ d’un ensemble X est une collection de parties $A_i \subset X$ non vides de X, deux à deux disjointes, et dont la réunion est X.*

Définition 5.2 *Un espace métrique (X, d) est connexe si les seules parties de X qui sont à la fois ouvertes et fermées sont triviales, c’est-à-dire si ce sont l’ensemble vide et X tout entier.*

Il est équivalent de dire qu’il n’existe pas de partition de X en deux ouverts. Ou encore qu’il n’existe pas de partition de X en deux fermés.

Heuristiquement, cela signifie qu’on ne peut pas décomposer X en deux parties “qui s’ignorent”.

Définition 5.3 *On dit qu’une partie d’un espace métrique est connexe si elle l’est pour la topologie induite.*

Tout comme la compacité, la connexité d’une partie $Y \subset X$ est une propriété intrinsèque de cette partie.

Proposition 5.4 *L’image continue d’un connexe est connexe.*

Preuve Soit $f : X \to Y$ une application continue entre deux espaces métriques. On suppose que X est connexe. On écrit $f(X) = V_0 \sqcup V_1$ comme réunion de deux ouverts disjoints de $f(X)$. On a donc $V_i = U_i \cap f(X)$, où U_i est un ouvert de Y $(i = 0, 1)$.
Il suit de la continuité de f que $X = f^{-1}(V_0) \sqcup f^{-1}(V_1)$ est réunion de deux ouverts disjoints de X. Puisque X est supposé connexe, l'un de ces ouverts est vide. On peut supposer que c'est $f^{-1}(V_1)$. Il suit que $V_1 = \emptyset$. □

Corollaire 5.5 *La connexité est une propriété topologique : tout espace métrique homéomorphe à un espace métrique connexe est lui-même connexe.*

Exemple 5.6 Les parties connexes de $X = \{0, 1\}$ sont l'ensemble vide et les singletons.

Plus généralement, les parties connexes non vides d'un ensemble discret (1.13) sont les singletons.

Passons à la description des parties connexes de $\mathbb{R}$. Cet exemple est fondamental (voir par exemple le paragraphe B, où l'on en déduira une condition suffisante de connexité).

Proposition 5.7 *Les parties connexes de $\mathbb{R}$ sont les intervalles. En particulier, $\mathbb{R}$ est connexe.*

Preuve Soit $A \subset \mathbb{R}$ une partie non vide. On suppose que A n'est pas un intervalle. Il existe donc deux points a_1 et a_2 dans A et un point $\alpha \notin A$ avec $a_1 < \alpha < a_2$. On obtient donc une partition de A en deux ouverts de A (chacun est non vide), soit

$$A = (A \cap]-\infty, \alpha[) \ \cup \ (A \cap]\alpha, \infty[)\,,$$

et A n'est donc pas connexe.

Soient I un intervalle non vide. On procède par l'absurde et on suppose qu'il existe une partition $I = U_0 \sqcup U_1$ de I en deux ouverts de I (qui sont donc également des fermés de I). On choisit $x_0 \in U_0$ et $x_1 \in U_1$. Sans perte de généralité, on supposera que $x_0 < x_1$. On introduit

$$y = \sup\{x \geq x_0 \mid [x_0, x] \subset U_0\}\,.$$

On a, par définition, $x_0 \leq y \leq x_1$. Puisque I est un intervalle, on a donc $y \in I$. Puisque $U_0 \subset I$ est une partie fermée de I il suit que $y \in U_0$. En particulier, $y < x_1$. Puisque U_0 est une partie ouverte de I, avec $[y, x_1] \subset I$, il suit qu'il existe $\varepsilon > 0$ pour lequel $[y, y + \varepsilon[\subset U_0$, en contradiction avec la définition de y. □

Exercice 5.8 Les espaces $\mathbb{R}$ et $\mathbb{R}^*$ sont-ils homéomorphes ?

Le théorème des valeurs intermédiaires est conséquence immédiate de la description des connexes de $\mathbb{R}$ et de la proposition 5.4.

Corollaire 5.9 Théorème des valeurs intermédiaires
Soient X un espace métrique connexe et $f : X \to \mathbb{R}$ continue. Alors l'image $f(X)$ est un intervalle.

Notamment, si $f : I \subset \mathbb{R} \to \mathbb{R}$ est une application continue définie sur un intervalle I, son image $f(I)$ est elle aussi un intervalle.

Les critères de compacité de Borel-Lebesgue et de Bolzano-Weierstrass étaient de natures très différentes. Par contre, le critère de connexité ci-dessous est une reformulation immédiate de la définition donnée plus haut.

Lemme 5.10 *Un espace métrique (X, d) est connexe si et seulement si toute application continue $h : X \to \{0, 1\}$ est constante.*

Preuve Supposons X connexe. Soit $h : X \to \{0, 1\}$ une application continue. Alors $h(X)$ est une partie connexe de $\{0, 1\}$, donc un singleton, c'est-à-dire que h est constante.

Si X n'est pas connexe, on choisit une partition $X = U_0 \sqcup U_1$ de X en deux parties ouvertes et fermées. L'application $h : X \to \{0, 1\}$ définie par $h(x) = 0$ lorsque $x \in U_0$ et $h(x) = 1$ lorsque $x \in U_1$ est localement constante (chaque point possède un voisinage sur lequel h est constante), donc continue. □

B Connexité par arcs

Dans ce paragraphe, nous introduisons une condition suffisante de connexité qui se révèle bien commode à utiliser. Les espaces connexes que vous rencontrerez cette année seront essentiellement tous connexes par arcs. Il faut se fatiguer un peu pour construire un espace connexe qui ne soit pas connexe par arcs (exercice 5.23) !

Définition 5.11 *Soit (X, d) un espace métrique. On dit que X est connexe par arcs lorsque, pour tout couple de points $x, y \in X$, on peut trouver un chemin (ou arc) continu qui joint x à y, c'est-à-dire une application continue $\gamma : [0, 1] \to X$ telle que $\gamma(0) = x$ et $\gamma(1) = y$.*

Exemple 5.12 Une partie convexe d'un espace vectoriel normé est connexe par arcs. Par exemple, un espace vectoriel normé, ou encore une boule ouverte ou une boule fermée de cet espace, sont connexes par arcs.

La connexité par arcs entraîne la connexité.

Proposition 5.13 *Un espace métrique connexe par arcs est connexe.*

Preuve C'est une conséquence de la connexité de l'intervalle. Supposons que l'espace métrique (X, d) soit connexe par arcs, et soit $h : X \to \{0,1\}$ continue. Nous voulons montrer que h est constante. Soient deux points $x_0, x_1 \in X$. Puisque X est supposé connexe par arcs, il existe un chemin continu $\gamma : [0,1] \to X$ qui joint x_0 à x_1. La composée $h \circ \gamma : [0,1] \to \{0,1\}$ est une application continue, donc constante. Il suit que $h(x_0) = h(x_1)$. $\square$

Corollaire 5.14 *Une partie convexe d'un espace vectoriel normé est connexe.*

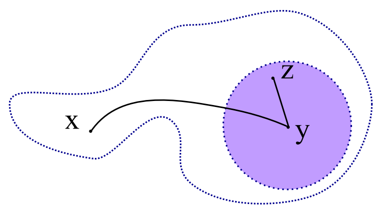
La réciproque de la proposition 5.13 n'est pas toujours vraie (voir l'exercice 5.23). Cependant, on a le résultat suivant qui nous fournit l'occasion d'un premier raisonnement "par connexité".

Proposition 5.15 *Soit $n \geq 1$. Un ouvert connexe $U \subset \mathbb{R}^n$ est connexe par arcs .*

Preuve Soit $x \in U$. On va montrer que l'ensemble

$$\lfloor x \rfloor = \{y \in U \text{ qu'on peut joindre à } x \text{ par un chemin continu tracé dans } U\}$$

est une partie non vide, ouverte et fermée de U. On concluera par connexité de U que $\lfloor x \rfloor = U$.



Un ouvert connexe de $\mathbb{R}^n$ est connexe par arcs

La relation définie sur U par $y\mathcal{R}z$ si et seulement si il existe un chemin continu de y à z tracé dans U est une relation d'équivalence sur U. La reflexivité et la symétrie sont immédiates (considérer un chemin constant, ou bien renverser le sens de parcours du chemin en introduisant $\gamma^\wedge(t) = \gamma(1-t)$). La transitivité suit quant à elle de ce que, si $\gamma_1 : [0,1] \to U$ est un chemin continu de y à z et $\gamma_2 : [0,1] \to U$ est un chemin continu de z à w, le concaténé $\gamma_2 * \gamma_1 : [0,1] \to U$ défini par $\gamma_2 * \gamma_1(t) = \gamma_1(2t)$ pour $0 \leq t \leq 1/2$ et $\gamma_2 * \gamma_1(t) = \gamma_2(2t-1)$ pour $1/2 \leq t \leq 1$ est un chemin continu de y à w.

Les boules étant connexes par arcs (exemple 5.12), il suit que la classe $\lfloor x \rfloor$ d'un point $x \in U$ pour cette relation d'équivalence est ouverte (dans U, ou dans $\mathbb{R}^n$, c'est la même chose puisque $U \subset \mathbb{R}^n$ est ouvert). Les classes pour la relation d'équivalence $\mathcal{R}$ constituant une partition de U, chaque classe

$$\lfloor x \rfloor = U \setminus \bigcup_{y \in U \setminus \lfloor x \rfloor} \lfloor y \rfloor$$

est donc également fermée dans U comme complémentaire d'un ouvert. $\square$

Exercice 5.16 1. Soit $n \geq 2$. Montrer que $\mathbb{R}^n$ privé d'un nombre fini de points est connexe par arcs.

2. Montrer que, pour $n \geq 2$, $\mathbb{R}^n$ n'est pas homéomorphe à $\mathbb{R}$.
3. Montrer que le groupe $\mathrm{Gl}_n\mathbb{C}$ est connexe par arcs.

Remarque 5.17 Théorème d'invariance du domaine
On peut montrer plus généralement que, pour $n \neq m$, $\mathbb{R}^n$ n'est pas homéomorphe à $\mathbb{R}^m$. Pour $n = 2 < m$, cela résulte de la notion de simple connexité. Pour $3 \leq n < m$, c'est nettement plus cher ![1]

Exercice 5.18 Soit $n \geq 1$.

1. Montrer que le groupe des rotations
$$\mathrm{SO}(2) = \left\{ \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \mid t \in \mathbb{R} \right\}$$
est connexe par arcs.
2. En déduire que le groupe spécial orthogonal $\mathrm{SO}(n)$ est connexe par arcs.
On pourra utiliser la réduction par blocs des matrices de $\mathrm{SO}(n)$.

C Opérations sur les connexes

Proposition 5.19 *Soit (X, d) un espace métrique.*

L'adhérence $\overline{A} \subset X$ d'une partie connexe A est encore connexe. Mieux, si A est connexe et si $A \subset B \subset \overline{A}$, alors B est connexe.

La réunion $\cup_{i\in I} A_i$ d'une famille de parties connexes de X d'intersection $\cap_{i\in I} A_i$ non vide est encore connexe.

Preuve Soit $h : B \to \{0, 1\}$ une application continue. La restriction $h_{|A}$ est constante. Tout point de B étant adhérent à A, h est constante.

Soit $h : \cup_{i\in I} A_i \to \{0, 1\}$ une application continue. En restriction à chaque A_i, cette application est constante et vaut $\alpha_i \in \{0, 1\}$. Tous les α_i sont égaux puisque les A_i ont un même point en commun. □

Exercice 5.20 Par contre l'intérieur d'une partie connexe, une réunion de parties connexes, ou une intersection de parties connexes ne sont pas toujours connexes (faire des dessins).

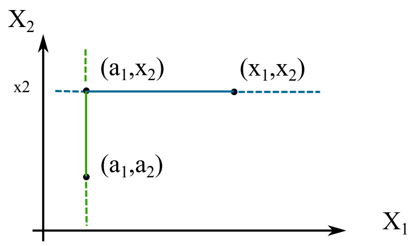
Proposition 5.21 *Un produit fini ou dénombrable $X = \prod_i X_i$ d'espaces métriques (X_i, d_i) est connexe si et seulement si chaque facteur est connexe.*

1. Dugundji, Topology

Preuve Rappelons que l'espace produit X est implicitement muni de la topologie produit (définition 1.60) pour laquelle chacune des projections $p_i : X \to X_i$ est continue.

Si X est connexe, il suit donc de la proposition 5.4 que chaque facteur $X_i = p_i(X)$ est connexe comme image continue d'un connexe.

Passons à la réciproque, et traitons pour commencer le cas d'un produit $X = X_1 \times X_2$ de deux espaces métriques connexes. Soit $h : X \to \{0,1\}$ une application continue. Nous allons voir que h est constante, et la connexité de X suivra alors du critère 5.10. Soient donc (x_1, x_2) et (a_1, a_2) deux points de X. L'application $h_1 : u \in X_1 \to h(u, x_2) \in \{0,1\}$ est continue comme composée de $i_1 : u \in X_1 \to (u, x_2) \in X_1 \times X_2$ et de h, donc constante puisque X_1 est connexe. On a donc $h(x_1, x_2) = h(a_1, x_2)$. De même, l'application $h_2 : v \in X_2 \to h(a_1, v) \in \{0,1\}$ est continue donc constante par connexité de X_2. Il suit que $h(x_1, x_2) = h(a_1, x_2) = h(a_1, a_2)$ comme annoncé.



Un produit de deux espaces connexes est connexe

Traitons maintenant le cas d'un produit dénombrable $X = \prod_{i \in \mathbb{N}} X_i$ d'espaces métriques connexes. Soit $h : X \to \{0,1\}$ une application continue.

Soient $x = (x_i)_{i \in \mathbb{N}}$ et $a = (a_i)_{i \in \mathbb{N}}$ deux points de X. Notons $\alpha := h(x)$. On veut montrer que $h(a) = \alpha$. Puisque $\{\alpha\} \subset \{0,1\}$ est ouvert, son image réciproque $h^{-1}(\alpha)$ est un voisinage ouvert de x dans X. Il suit de la proposition 1.65 qu'il existe un entier $n \in \mathbb{N}$ et un rayon $r > 0$ tels que

$$\prod_{i=0}^{n} B_{X_i}(x_i, r) \times \prod_{i \geq n+1} X_i \subset h^{-1}(\alpha)\,.$$

En particulier, on a $h(x_0, x_1, \ldots, x_n, a_{n+1}, a_{n+2}, \ldots) = \alpha$. Introduisons les points $\xi_k = (x_0, x_1, , \ldots, x_{k-1}, a_k, a_{k+1}, \ldots) \in X$ pour $k = 0$ à $n+1$. On vient de voir que $h(\xi_{n+1}) = \alpha$. On procède comme précédemment (produit de deux espaces connexes) pour montrer récursivement que

$$h(\xi_{n+1}) = f(\xi_n) = \cdots = h(\xi_0)\,.$$

Le résultat suit puisque $\xi_0 = a$. $\square$

Exercice 5.22 On dira qu'une application $f : I \to \mathbb{R}$ définie sur un intervalle $I \subset \mathbb{R}$ vérifie la propriété des valeurs intermédiaires lorsque l'image de tout sous-intervalle $J \subset I$ est un intervalle.

1. Une application continue $f : I \to \mathbb{R}$ vérifie la propriété des valeurs intermédiaires.
2. Une application $f : I \to \mathbb{R}$ qui vérifie la propriété des valeurs intermédiaires n'est pas forcément continue. Exemple ?
3. **Théorème de Darboux** Soit $f : \mathbb{R} \to \mathbb{R}$ dérivable en chaque point. Montrer que sa dérivée $f' : \mathbb{R} \to \mathbb{R}$ vérifie la propriété des valeurs intermédiaires.

   Soit $[a, b] \subset \mathbb{R}$ un intervalle non réduit à un point. On remarquera que l'ensemble des valeurs prises par f' sur cet intervalle vérifie

   $$\{\frac{f(s)-f(t)}{s-t} \mid a \leq s < t \leq b\} \subset \{f'(c) \mid a \leq c \leq b\} \subset \overline{\{\frac{f(s)-f(t)}{s-t} \mid a \leq s < t \leq b\}}.$$

4. Exhiber une fonction $f : \mathbb{R} \to \mathbb{R}$ dérivable en chaque point dont la dérivée ne soit pas continue.

Exercice 5.23 Un espace connexe non connexe par arcs
Soient $f : t \in]0, 1] \to \sin(1/t) \in \mathbb{R}$, et $\Gamma = \{(t, f(t))\,,\ t \in]0, 1]\} \subset \mathbb{R}^2$ son graphe.

1. Montrer que Γ est connexe par arcs.
2. Montrer que l'adhérence $\overline{\Gamma}$ du graphe de f est connexe. La dessiner.
3. Montrer que $\overline{\Gamma}$ n'est pas connexe par arcs.

   On pourra procéder par l'absurde, introduire un chemin continu $\gamma : [0, 1] \to \overline{\Gamma}$ qui joint les points $m_0 = (1, f(1))$ et $m_1 = (0, 0)$, et utiliser l'uniforme continuité de l'application γ.

D Composantes connexes

Proposition-Définition 5.24 *Soient (X, d) un espace métrique et $x \in X$ un point de cet espace. Il existe une plus grande partie connexe $C_x \subset X$ contenant x. C'est la composante connexe du point x.*

Chaque composante connexe est fermée dans X.

La famille des composantes connexes de X constitue une partition de X.

Preuve Il existe une partie connexe de X contenant x, à savoir le singleton $\{x\}$. La réunion des parties connexes contenant x est, d'après la proposition 5.19, une partie connexe de X contenant x. Par construction, c'est la plus grande. Les autres assertions résultent également de la proposition 5.19. □

Remarque 5.25 *La relation "$x\mathcal{R}y$ si et seulement si x et y appartiennent à la même composante connexe de X" est une relation d'équivalence sur X.*

Exercice 5.26 Déterminer les composantes connexes de l'ensemble de Cantor $\mathcal{C} = \{0,1\}^{\mathbb{N}}$.

Exemple 5.27 Les composantes connexes de $\mathbb{Q}$ et de $\mathbb{R} \setminus \mathbb{Q}$ sont les singletons.

En particulier, la composante connexe $C_x \subset X$ d'un point dans un espace métrique quelconque peut ne pas être ouverte. Cependant, pour les ouverts de $\mathbb{R}^n$, on fait l'observation suivante.

Lemme 5.28 *Soit $U \subset \mathbb{R}^n$ un ouvert. Ses composantes connexes sont ouvertes (dans U ou dans $\mathbb{R}^n$, c'est la même chose).*

Preuve Soit $x \in U$ et soit $y \in C_x$ un point de U qui est dans la même composante connexe que x. On a donc $C_x = C_y$. Puisque U est ouvert, il existe une boule ouverte $B(y,r) \subset U$. Cette boule est connexe, donc $B(y,r) \subset C_y = C_x$. $\square$

Connaissant les parties connexes de $\mathbb{R}$, on en déduit immédiatement une description des ouverts de $\mathbb{R}$.

Corollaire 5.29 *Tout ouvert de $\mathbb{R}$ s'écrit comme réunion d'une famille finie ou dénombrable d'intervalles ouverts deux à deux disjoints.*

Preuve On écrit l'ouvert $U \subset \mathbb{R}$ comme réunion disjointe de ses composantes connexes, qui sont des ouverts connexes de $\mathbb{R}$, et donc des intervalles ouverts disjoints. Ces intervalles contenant chacun un rationnel, il y en a au plus une infinité dénombrable. $\square$

Exercice 5.30

1. On rappelle que le groupe spécial orthogonal $\mathrm{SO}(n)$ est connexe (exercice 5.18). Déterminer les composantes connexes du groupe orthogonal $\mathrm{O}(n)$.
2. En déduire les composantes connexes de $\mathrm{Gl}_n\mathbb{R}$ (on utilisera la décomposition polaire, exercice 2.31).

Groupes topologiques

Un “groupe topologique” est un groupe $(G, .)$ muni d’une topologie compatible avec sa structure de groupe, c’est-à-dire pour laquelle les applications

$$(x, y) \in G \times G \to x\,y \in G \text{ et } x \in G \to x^{-1} \in G$$

sont continues.

Exercice 5.31 Les groupes $(\mathbb{R}, +)$, $(\mathbb{R}^n, +)$, $(\mathbb{R}^*, .)$, $(\mathbb{R}^*_+, .)$ ou encore $\mathrm{Gl}_n\mathbb{R}$ munis de leur topologie usuelle sont des groupes topologiques.
L’ensemble de Cantor $\{0, 1\}^{\mathbb{N}} = (\mathbb{Z}/2\mathbb{Z})^{\mathbb{N}}$ est naturellement un groupe topologique.

Exercice 5.32

1. Soit G un groupe topologique. On suppose que $H \subset G$ est un sous-groupe ouvert de G. Montrer que $H \subset G$ est alors fermé. En particulier, si l’on suppose de plus que G est connexe, il suit que $G = H$.
2. **Un groupe topologique connexe est engendré par ses petits éléments** Soient G un groupe topologique connexe et $U \subset G$ un voisinage de l’élément neutre e. Alors G est égal au sous-groupe de G engendré par U.
3. **Exemple : exponentielle matricielle.** On verra en 8.6 que l’image $\exp(\mathrm{M}_n\mathbb{R}) \subset \mathrm{Gl}^+_n\mathbb{R}$ est un voisinage de Id.
   (a) Montrer que $\exp(-M)\exp(M) = \mathrm{Id}$ pour toute matrice $M \in \mathrm{M}_n\mathbb{R}$.
   (b) Montrer que toute matrice inversible $A \in \mathrm{Gl}^+_n\mathbb{R}$ de déterminant positif peut s’écrire comme un produit d’exponentielles de matrices réelles.
   (c) Montrer que la matrice $M = \begin{pmatrix} -1 & 0 \\ 0 & -2 \end{pmatrix}$ n’est l’exponentielle d’aucune matrice réelle.
4. Soit G un groupe topologique. On désigne par G_e sa “composante neutre”, c’est-à-dire la composante connexe dans G de l’élément neutre e.
   (a) Déterminer les composantes neutres de $\mathbb{R}^*$, de $\mathrm{O}(n)$, de $\mathrm{Gl}_n\mathbb{R}$, et du produit $\mathcal{C} = (\mathbb{Z}/2\mathbb{Z})^{\mathbb{N}}$.
   (b) Montrer que $G_e \subset G$ est un sous-groupe de G.

6. Différentielle

Dans tous les chapitres concernant la différentiabilité, $U \subset \mathbb{R}^n$ désignera un ouvert non vide de l'espace vectoriel de dimension finie $\mathbb{R}^n$ ($n \geq 1$).

A Dérivée d'une fonction de variable réelle

Avant de définir la différentielle d'une application de plusieurs variables, revenons sur le cas de la dimension 1 qui nous servira de ligne directrice.

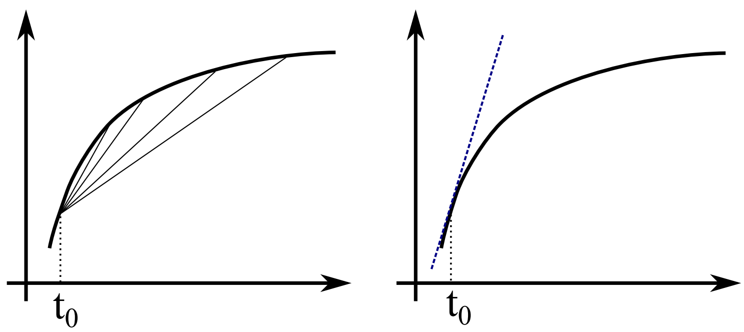
Définition 6.1 *Soient $f : U \subset \mathbb{R} \to \mathbb{R}$ une fonction, et $t_0 \in U$. On dit que la fonction f est dérivable au point t_0 lorsque les taux d'accroissement*

$$\frac{f(t_0+h) - f(t_0)}{h}$$

(pentes des sécantes) ont une limite lorsque $h \to 0$. On dit alors que f est dérivable au point t_0, et on définit la dérivée de f en t_0 comme étant

$$f'(t_0) = \lim_{h \to 0} \frac{f(t_0+h) - f(t_0)}{h} \in \mathbb{R}\,.$$

Géométriquement, le graphe de la fonction f admet alors en $(t_0, f(t_0))$ une "droite affine tangente", qui est le graphe de la fonction affine $t \to f(t_0) + f'(t_0)\,(t - t_0)$.



Les sécantes, la tangente

On remarquera que f est dérivable en t_0, avec une dérivée $f'(t_0) = \alpha$, si et seulement si

$$f(t_0 + h) = f(t_0) + \alpha h + o(h) . \qquad (*)$$

Autrement dit, au voisinage de t_0, la fonction f est "proche" de la fonction affine $t \to f(t_0) + \alpha(t - t_0)$, la différence de ces deux fonctions étant un $o(h)$ (notation de Landau, qui signifie que ce terme d'erreur $o(h)$ tend vers 0 plus vite que h lorsque $h \to 0$, c'est-à-dire que les quotients $o(h)/h$ tendent vers 0 avec h).

Toujours pour des fonctions de variable réelle, on peut s'intéresser aux fonctions à valeurs dans $\mathbb{R}^p$, muni bien sûr de sa topologie d'espace vectoriel normé. Rien ne change.

Définition 6.2 *Soient $f : U \subset \mathbb{R} \to \mathbb{R}^p$ une fonction, et $t_0 \in U$. On dit que la fonction f est dérivable au point t_0 lorsque les taux d'accroissement de f en t_0 ont une limite ; cette limite*

$$f'(t_0) = \lim_{h \to 0} \frac{f(t_0 + h) - f(t_0)}{h} \in \mathbb{R}^p$$

est la dérivée de f en t_0. On peut alors écrire (de façon équivalente)

$$f(t_0 + h) = f(t_0) + h\, f'(t_0) + o(h) . \qquad (*)$$

Remarque 6.3 Remarquons que dans ce cas $f'(t_0) \in \mathbb{R}^p$ est un vecteur. On a donc pris soin d'écrire $h\, f'(t_0)$, avec le scalaire h devant le vecteur $f'(t_0)$ sur lequel il opère.

Cette remarque innocente est prendre au sérieux. Savoir à chaque instant "de quoi on parle", c'est-à-dire faire en permanence l'effort d'avoir en tête la nature de chacun des objets que vous manipulez, vous évitera bien des déboires et facilitera votre apprentissage du calcul différentiel.

B Différentielle

Nous allons généraliser la notion de dérivée aux applications de plusieurs variables, et introduire la différentielle d'une application $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$ $(n, p \geq 1)$. On ne divise pas par des vecteurs... la dimension supérieure passera donc par la formulation (*).

Munissons chacun des espaces $\mathbb{R}^n$ et $\mathbb{R}^p$ d'une norme, qui seront toutes deux désignées par $\| \;\|$.

Définition 6.4 *Soient $U \subset \mathbb{R}^n$ un ouvert, $f : U \to \mathbb{R}^p$ et $x_0 \in U$. On dit que f est différentiable en x_0 lorsqu'il existe une application linéaire $L : \mathbb{R}^n \to \mathbb{R}^p$ telle que*

$$f(x_0 + h) = f(x_0) + L(h) + o(h)\,. \tag{6.1}$$

Autrement dit, on demande que le quotient

$$\frac{\|f(x_0 + h) - f(x_0) - L(h)\|}{\|h\|}$$

tende vers 0 *lorsque $h \to 0$.*

L'application f est donc différentiable en x_0 si et seulement si il existe une application linéaire $L \in L(\mathbb{R}^n, \mathbb{R}^p)$ telle que, pour tout $\varepsilon > 0$, il existe $\eta > 0$ avec pour $\|h\| \leq \eta$:

$$\|f(x_0 + h) - f(x_0) - L(h)\| \leq \varepsilon \, \|h\|\,.$$

L'interprétation géométrique est encore pertinente : au voisinage de x_0, le graphe de f reste proche de celui de l'application affine

$$x \to f(x_0) + L(x - x_0)\,,$$

qui est un sous-espace affine de dimension n de $\mathbb{R}^n \times \mathbb{R}^p$.

Remarque 6.5 *Lorsque $f : U \to \mathbb{R}^p$ est différentiable au point $x_0 \in U$, elle est* a fortiori *continue en ce point.*

Montrons l'unicité de l'application linéaire L qui apparaît dans (6.1).

Lemme 6.6 *Soit une application linéaire $L : \mathbb{R}^n \to \mathbb{R}^p$ pour laquelle*

$$L(h) = o(h)\,.$$

Alors L est identiquement nulle.

Preuve Cela résulte de l'homogénéité de l'application linéaire L. Fixons un vecteur $u \in \mathbb{R}^n$ non nul. L'hypothèse implique que le ratio

$$\frac{\|L(tu)\|}{\|tu\|} = \frac{|t| \, \|L(u)\|}{|t| \, \|u\|} = \frac{\|L(u)\|}{\|u\|}\,,$$

qui est donc indépendant de t, est également un $o(tu)/\|tu\|$ donc tend vers 0 lorsque $t \to 0$. On a donc $L(u) = 0$. Le résultat suit. □

Le lecteur, s'il n'est pas suffisamment familier avec les o, est invité à réécrire cette preuve avec des ε.

On en déduit immédiatement la définition suivante.

Lemme-Définition 6.7 *Lorsque l'application f est différentiable en x_0, il existe une unique application linéaire $L : \mathbb{R}^n \to \mathbb{R}^p$ telle que*

$$f(x_0 + h) = f(x_0) + L(h) + o(h)\,.$$

Cette application linéaire est appelée différentielle de f en x_0. On la note

$$D_{x_0} f := L \in L(\mathbb{R}^n, \mathbb{R}^p)\,.$$

L'intérêt de cette notion est que, avec quelques précautions, une application différentiable se comportera localement “comme” son application affine tangente : on récupérera une information locale à partir d'informations de nature infinitésimale. Voir notamment le théorème des accroissements finis 6.13, le théorème d'inversion locale 8.4 ou encore le théorème des fonctions implicites 8.18.

Remarque Et si l'on change de norme ? Pas d'inquiétude à avoir, nous sommes en dimension finie donc toutes les normes sont équivalentes. Le fait que f soit différentiable en x_0, ainsi que le calcul de sa différentielle $D_{x_0} f$, ne dépendent donc pas du choix des normes.

Remarque Soient $f : U \subset \mathbb{R} \to \mathbb{R}^p$ une fonction de variable réelle. Elle est dérivable en $t_0 \in U$ si et seulement si elle est différentiable en ce point. On a alors

$$D_{t_0} f\,(1) = f'(t_0) \in \mathbb{R}^p$$

(et bien sûr $D_{t_0} f\,(h) = h\,f'(t_0)$ pour tout $h \in \mathbb{R}$).

Définition 6.8 *Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$. Si l'application f est différentiable en chaque point x de U, on dit simplement que f est différentiable sur U.*

On dispose alors, pour chaque point $x \in U$, de la différentielle de f en x, soit $D_x f \in L(\mathbb{R}^n, \mathbb{R}^p)$. On appelle différentielle de f l'application

$$\begin{aligned} Df : U &\longrightarrow L(\mathbb{R}^n, \mathbb{R}^p) \\ x &\longrightarrow D_x f\,. \end{aligned}$$

Exemple 6.9 L'archétype de l'application différentiable !

Une application linéaire $L : \mathbb{R}^n \to \mathbb{R}^p$ est différentiable. Sa différentielle en chaque point $x \in \mathbb{R}^n$ est $D_x L = L$. Son application différentielle est donc l'application constante

$$\begin{aligned} DL : \mathbb{R}^n &\longrightarrow L(\mathbb{R}^n, \mathbb{R}^p) \\ x &\longrightarrow L\,. \end{aligned}$$

Remarque On peut définir de même la différentielle d'une application $f : U \subset E \to F$ définie sur un ouvert d'un espace vectoriel E de dimension finie n, et à valeurs dans un autre espace vectoriel de dimension finie p. Il n'y a rien à changer, un choix de bases pour E et F permettant d'identifier respectivement ces espaces vectoriels à $\mathbb{R}^n$ et $\mathbb{R}^p$.

Nous nous limitons dans ce cours au cadre de la dimension finie, le besoin d'aller vers plus de généralité ne se faisant pas encore sentir à notre niveau.

Noter cependant que l'on peut définir de même la différentielle d'une application

$$f : U \subset E \to F\,,$$

où E et F sont deux espaces vectoriels normés de dimension quelconque, et $U \subset E$ est une partie ouverte. On demandera alors que les applications linéaires tangentes, i.e. les différentielles $D_x f$ ($x \in U$), soient des applications linéaires continues $L \in \mathcal{L}_c(E, F)$. Dans ce contexte, le choix des normes sur les espaces E et F est important.

C Propriétés de stabilité

• Une somme d'applications $f, g : U \subset \mathbb{R}^n \to \mathbb{R}^p$ différentiables en x_0 est encore différentiable en ce point, avec

$$D_{x_0}(f+g) = D_{x_0} f + D_{x_0} g\,.$$

En effet une somme de deux "o" est encore un "o".

• Le produit d'une application $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$ différentiable en x_0 par un scalaire $\lambda \in \mathbb{R}$ est encore différentiable en ce point, avec

$$D_{x_0}(\lambda f) = \lambda\, D_{x_0} f\,.$$

En effet le produit d'un "o" par une constante est encore un "o".

• Une composée d'applications différentiable l'est. Précisément :

Proposition 6.10 *Soient $U \subset \mathbb{R}^n$ et $V \subset \mathbb{R}^p$ deux ouverts. On se donne deux applications $f : U \subset \mathbb{R}^n \to V \subset \mathbb{R}^p$ et $g : V \subset \mathbb{R}^p \to \mathbb{R}^q$. On suppose que f est différentiable en x_0, et que g est différentiable en $f(x_0)$.*

Alors $g \circ f$ est différentiable en x_0 et sa différentielle en ce point est la composée

$$D_{x_0}(g \circ f) = D_{f(x_0)} g \circ D_{x_0} f\,. \tag{6.2}$$

Autrement dit, on a pour tout vecteur $h \in \mathbb{R}^n$:

$$D_{x_0}(g \circ f)\,(h) = D_{f(x_0)} g\,\big((D_{x_0} f)(h)\big)\,.$$

On retrouve en particulier que, pour deux fonctions réelles de variable réelle f et g, on a

$$(g \circ f)' = (g' \circ f)\, f'\,.$$

Avant de démontrer cette assertion on se convainc que, si l'on admet que $g \circ f$ est effectivement différentiable en x_0, la formule ci-dessus – qui donne l'expression de la différentielle de $g \circ f$ en x_0 – est la seule qui soit raisonnable : données une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^p$ (ici $D_{x_0}f$) et une application linéaire de $\mathbb{R}^p$ dans $\mathbb{R}^q$ (ici $D_{f(x_0)}g$) on doit en effet fabriquer $D_{x_0}(g \circ f)$, linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^q$...

Preuve Il n'y a pas d'initiative à prendre, on se laisse porter par les définitions. Puisque f est différentiable en x_0, on a pour $h \in \mathbb{R}^n$ tel que $x_0 + h \in U$

$$(g \circ f)(x_0 + h) = g\big(f(x_0 + h)\big) = g\big(f(x_0) + D_{x_0}f\,(h) + o_f(h)\big)\,.$$

Puisque g est différentiable en $f(x_0)$, il vient alors

$$\begin{aligned}(g \circ f)(x_0 + h) &= g\big(f(x_0)\big) + D_{f(x_0)}g\,(D_{x_0}f\,(h) + o_f(h)) + o_g\big(D_{x_0}f\,(h) + o_f(h)\big)\\ &= (g \circ f)(x_0) + (D_{f(x_0)}g \cdot D_{x_0}f)\,(h)\\ &\quad + [D_{f(x_0)}g\,(o_f(h)) + o_g\big(D_{x_0}f\,(h) + o_f(h)\big)]\,.\end{aligned}$$

On vérifie facilement que le terme entre crochets est un "$o(h)$", ce qui donne le résultat. En effet d'une part

$$\|D_{f(x_0)}g\,(o_f(h))\| \leq \|D_{f(x_0)}g\|\;\|o_f(h))\|\,.$$

Et d'autre part $\|D_{x_0}f\,(h) + o_f(h)\| \leq (\|D_{x_0}f\| + 1)\;\|h\|$ lorsque h est assez petit, donc la quantité $o_g\big(D_{x_0}f\,(h) + o_f(h)\big)$ est un $o(h)$. □

Le lecteur, si il n'est pas suffisamment familier avec les o, est de nouveau invité à réécrire cette preuve avec des ε.

D Quelques exemples, fondamentaux bien sûr

Les premiers exemples d'applications différentiables sont évidemment donnés par les applications linéaires (ou affines) elles-mêmes.

• Redisons qu'une application linéaire $L : \mathbb{R}^n \to \mathbb{R}^p$ est différentiable puisque

$$L(x + h) = L(x) + L(h)$$

pour tous $x, h \in \mathbb{R}^n$. Sa différentielle est l'application constante

$$DL : x \in \mathbb{R}^n \to L \in L(\mathbb{R}^n, \mathbb{R}^p)\,.$$

En d'autres termes, la différentielle de L en chaque point $x \in \mathbb{R}^n$ est l'application L elle-même. De plus, dans ce cas, le "terme d'erreur" en $o(h)$ est nul.

Par exemple l'application "Trace" tr : $M_n\mathbb{R} \to \mathbb{R}$ est linéaire, donc différentiable.

• Une application affine $A : x \in \mathbb{R}^n \to L(x) + y \in \mathbb{R}^p$ (où $L \in L(\mathbb{R}^n, \mathbb{R}^p)$ et $y \in \mathbb{R}^p$) est différentiable. Sa différentielle est également l'application constante

$$DA : x \in \mathbb{R}^n \to L \in L(\mathbb{R}^n, \mathbb{R}^p)\,.$$

• Une application $f = (f_1, \ldots, f_p) : \mathbb{R}^n \to \mathbb{R}^p$ est différentiable si et seulement si chacune des applications coordonnées f_k ($1 \leq k \leq p$) l'est. En effet les projections

$$\pi_k : (x_1, \ldots, x_p) \in \mathbb{R}^p \to x_k \in \mathbb{R}$$

et les injections

$$i_k : x \in \mathbb{R} \to (0, \ldots, x, 0 \ldots, 0) \in \mathbb{R}^p$$

sont linéaires donc différentiables ($1 \leq k \leq p$). Et l'on a

$$f_k = \pi_k \circ f \quad \text{et} \quad f = \sum_{k=1}^{p} i_k \circ f_k\,.$$

• Une application multilinéaire est différentiable. Soient $E_1, \ldots, E_k$ et F des espaces vectoriels de dimension finie sur $\mathbb{R}$, et $f : E_1 \times \cdots \times E_k \to F$ une application k-linéaire. La multilinéarité de f permet d'écrire, pour tous $x = (x_1, \ldots, x_k)$ et $h = (h_1, \ldots, h_k)$,

$$\begin{aligned} f(x+h) &= f(x_1 + h_1, x_2 + h_2, \cdots, x_k + h_k) \\ &= f(x_1, x_2, \ldots, x_k) \\ &\quad + f(h_1, x_2, \ldots, x_k) + f(x_1, h_2, x_3, \ldots, x_k) + \cdots + f(x_1, x_2, \ldots, x_{k-1}, h_k) \\ &\quad + R\,, \end{aligned}$$

autrement dit comme somme du terme constant $f(x_1, x_2, \ldots, x_k)$, d'un terme linéaire en h, et d'un "reste" R qui est une expression comportant un nombre fini de termes, chacun d'entre eux étant obtenu en évaluant f sur k vecteurs dont au moins deux sont choisis parmi $h_1, \ldots, h_k$ et les autres parmi $x_1, \ldots, x_k$.

Rappelons que l'application $(y_1, \ldots, y_k) \in E_1 \times \cdots \times E_k \to \sup_{i=1}^{k} \|y_i\|$ est une norme sur l'espace produit $E_1 \times \cdots \times E_k$. De plus f, multilinéaire en dimension finie, est continue. Il existe donc (exercice 4.14) une constante $C > 0$ telle que, pour tout $(y_1, \ldots, y_k) \in E_1 \times \cdots \times E_k$, on ait

$$|f(y_1, \ldots, y_k)| \leq C \prod_{i=1}^{k} \|y_i\|\,.$$

Pour $\|h\| \leq 1$, le reste R est donc majoré par $C\|h\|^2$ où C est une constante. Ce reste est donc un $o(h)$. On en déduit que l'application f est différentiable et que

$$D_x f(h) = f(h_1, x_2, \ldots, x_k) + f(x_1, h_2, x_3, \ldots, x_k) + \cdots + f(x_1, x_2, \ldots, x_{k-1}, h_k)\,.$$

• On en déduit que si $B : \mathbb{R}^p \times \mathbb{R}^p \to \mathbb{R}^q$ est une application bilinéaire, et $u, v : \mathbb{R}^n \to \mathbb{R}^p$ sont deux applications différentiables, l'application

$$F : x \in \mathbb{R}^n \to B(u(x), v(x)) \in \mathbb{R}^q$$

est différentiable et sa différentielle est donnée par

$$DF = B(Du, v) + B(u, Dv)\,,$$

ce qui signifie plus explicitement que pour tous $x, h \in \mathbb{R}^n$ on a

$$D_x F(h) = B(D_x u(h), v(x)) + B(u(x), D_x v(h))\,.$$

▴ Pour $B : (s,t) \in \mathbb{R} \times \mathbb{R} \to st \in \mathbb{R}$, on retrouve la formule de Leibniz

$$(uv)' = u'v + uv'\,,$$

valable pour un couple (u, v) de fonctions réelles de variable réelle.
De même, le produit de deux fonctions différentiables $u, v : \mathbb{R}^n \to \mathbb{R}$ à valeurs réelles est différentiable. On laisse au lecteur le soin de déterminer la différentielle du produit.
Il suit qu'une application polynômiale $P \in \mathbb{R}[X_1, \ldots, X_n] : \mathbb{R}^n \to \mathbb{R}$ est différentiable. Il suffit de le voir pour un monôme et, pour cela, de raisonner par récurrence sur son degré.

▴ On pourra également utiliser ce résultat lorsque l'application bilinéaire $B : \mathbb{R}^p \times \mathbb{R}^p \to \mathbb{R}$ est le produit scalaire associé à une structure euclidienne.

• Une autre illustration importante concerne l'application déterminant, soit $\det : \mathrm{M}_n\mathbb{R} \to \mathbb{R}$, qui est différentiable en chaque point $M \in \mathrm{M}_n\mathbb{R}$. En effet $\det : \mathrm{M}_n\mathbb{R} \simeq (\mathbb{R}^n)^n \to \mathbb{R}$ est multilinéaire (et alternée). Sa différentielle en la matrice identité est la trace :

$$D_{\mathrm{Id}} \det = \mathrm{Tr} \in L(\mathrm{M}_n\mathbb{R}, \mathbb{R})\,.$$

Voir également l'exercice 7.11.

E Le gradient d'une fonction scalaire

Dans ce paragraphe, nous nous intéressons à une application différentiable à valeurs réelles, soit $f : U \subset \mathbb{R}^n \to \mathbb{R}$. Nous venons d'introduire la différentielle de f. Celle-ci fournit, en chaque point $x \in U$, une forme linéaire

$$D_x f \in L(\mathbb{R}^n, \mathbb{R})\,.$$

Ceci n'est pas forcément facile à visualiser et peut donc paraître un peu abstrait.

Munissons donc $\mathbb{R}^n$ d'une structure euclidienne, c'est-à-dire d'une forme bilinéaire $B : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ définie positive. Ce peut être par exemple le produit scalaire canonique

$$(u,v) = \big((u_1,\ldots,u_n),(v_1,\ldots,v_n)\big) \in \mathbb{R}^n \times \mathbb{R}^n \to < u,v >= \sum_{i=1}^{n} u_i v_i \in \mathbb{R}\,.$$

A chaque vecteur $u \in \mathbb{R}^n$, on associe la forme linéaire

$$\ell_u = B(u,.) : v \in \mathbb{R}^n \to B(u,v) \in \mathbb{R}\,.$$

Le choix du produit scalaire B permet d'identifier les formes linéaires sur $\mathbb{R}^n$ aux vecteurs de $\mathbb{R}^n$ grâce à l'isomorphisme d'espaces vectoriels

$$j : u \in \mathbb{R}^n \longrightarrow \ell_u \in L(\mathbb{R}^n, \mathbb{R})$$

(linéaire et injectif entre deux espaces vectoriels de même dimension n). Lorsque le vecteur u est non nul, le noyau de la forme linéaire ℓ_u est l'hyperplan orthogonal à u (relativement à B). L'isomorphisme $j : \mathbb{R}^n \to L(\mathbb{R}^n, \mathbb{R})$ n'est pas canonique, et dépend du produit scalaire B que l'on a choisi.

Revenons à nos différentielles. Lorsque $f : U \subset \mathbb{R}^n \to \mathbb{R}$ est une fonction à valeurs réelles différentiable en x, l'unique vecteur $u \in \mathbb{R}^n$ pour lequel $\ell_u = D_x f$ est appelé vecteur gradient de f au point x et noté $u := \mathrm{grad}_x f$. Il vérifie, pour tout vecteur $h \in \mathbb{R}^n$,

$$D_x f(h) = B(\mathrm{grad}_x f, h)\,.$$

Interprétation géométrique Partant du point x, il faut s'en éloigner dans la direction de $\mathrm{grad}_x f$ si l'on veut que f croisse au plus vite (au premier ordre, et à vitesse donnée) : c'est Cauchy-Schwarz.
Par contre, si l'on part du point x dans une direction orthogonale à $\mathrm{grad}_x f$, la fonction f reste constante au premier ordre. On y reviendra à la proposition 10.8.

Exemple : cartes de randonnées, et fonction hauteur
Les lignes de niveau sont orthogonales au gradient de la fonction “hauteur”. Les lignes de plus grande pente sont tangentes au gradient de la fonction “hauteur”.

Exercice 6.11 Fonction distance à un point
Dans $\mathbb{R}^n$ euclidien, on se donne un point Ω et on introduit la fonction distance à Ω, soit $f : m \in \mathbb{R}^n \to \|m - \Omega\| \in \mathbb{R}$. Déterminer les points où f est différentiable, et calculer son gradient en ces points. Vérifier que le résultat obtenu est conforme à l'intuition géométrique.

F Le théorème des accroissements finis

C'est un résultat fondamental, qui nous servira moult fois.

La différentiabilité d'une application f au point x apporte une information sur le comportement asymptotique de f lorsqu'on s'approche du point x. Dire que

$$f(x+h) = f(x) + D_x f(h) + o(h)$$

n'apporte rien lorsque l'accroissement h est fixé, puisque la seule information dont on dispose sur le terme d'erreur $o(h)$ est qu'il tend plus rapidement vers 0 que h lui-même. On ne s'intéresse donc dans cette définition qu'aux accroissements “infinitésimaux” de x. Par contraste, on parle dans ce qui suit d'accroissements “finis”.

Le théorème des accroissements finis s'énonce d'abord pour des fonctions d'une variable réelle. Rappelons l'énoncé élémentaire suivant, valable pour les fonctions scalaires.

Rappel 6.12 Théorème de Rolle
Soit $f : [a,b] \subset \mathbb{R} \to \mathbb{R}$ une fonction continue. On suppose que f est dérivable sur l'intervalle ouvert $]a,b[$.
Il existe alors $c \in]a,b[$ pour lequel $f(b) - f(a) = f'(c)\,(b-a)$ (voir 2.13).

Attention, cet énoncé ne se généralise pas tel quel à une fonction $f : [a,b] \subset \mathbb{R} \to \mathbb{R}^n$ lorsque $n \geq 2$. Méditer, sur l'intervalle $[0, 2\pi]$, l'exemple de l'exponentielle complexe

$$f : t \in \mathbb{R} \longrightarrow (\cos t, \sin t) \in \mathbb{R}^2$$

pour laquelle $f(0) = f(2\pi)$ sans que f' ne s'annule.

Par contre, lorsqu'on munit $\mathbb{R}^n$ d'une norme $\|\ \|$, on a le résultat suivant.

Théorème 6.13 des accroissements finis

Soient $f : [a,b] \to \mathbb{R}^n$ et $g : [a,b] \to \mathbb{R}$ deux applications continues, que l'on suppose dérivables sur l'intervalle ouvert $]a,b[$. On suppose qu'on a

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

pour tout réel $t \in]a,b[$. Alors on a l'inégalité

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

Remarque 6.14 Pour une fonction réelle de variable réelle, le théorème de Rolle exprime l'accroissement $f(b) - f(a)$ à l'aide de la dérivée de f en un point convenable. Pour une application à valeurs dans $\mathbb{R}^n$, on sait seulement majorer l'accroissement $\|f(b)-f(a)\|$ à partir d'une majoration de la dérivée (ou de la différentielle) de f ; voir également le corollaire 6.16 et, pour les formules de Taylor, la remarque 9.28.

Preuve On va d'abord se donner un peu de marge, c'est-à-dire fixer $\varepsilon > 0$, et démontrer que pour tout $t \in [a, b]$ on a

$$\|f(t) - f(a)\| \leq g(t) - g(a) + \varepsilon(t - a) + \varepsilon\,.$$

Le résultat voulu en découlera immédiatement en faisant tendre ε vers 0.

Soit donc $\varepsilon > 0$. On introduit $I_\varepsilon \subset [a, b]$ défini par

$$I_\varepsilon = \{s \in [a, b] \text{ tels que } \forall t \in [a, s] \text{ on a } \|f(t) - f(a)\| \leq g(t) - g(a) + \varepsilon(t - a) + \varepsilon\}\,.$$

On observe que :

- $I_\varepsilon \subset [a, b]$ est un intervalle de $[a, b]$ contenant le point a
- $I_\varepsilon \subset [a, b]$ est fermé (applications continues, inégalités larges).

Soit $c = \sup I_\varepsilon \in [a, b]$. On a donc $I_\varepsilon = [a, c]$. Nous allons montrer que $c = b$. La continuité de f et g au point a montre immédiatement que $c > a$. Supposons par l'absurde que $c < b$. Les applications f et g sont différentiables au point $c \in]a, b[$. L'hypothèse et le fait que $c \in I_\varepsilon$ assurent que, pour tout $h > 0$ tel que $c + h \leq b$:

$$\begin{aligned}\|f(c+h) - f(a)\| &\leq \|f(c+h) - f(c)\| + \|f(c) - f(a)\| \\ &\leq h\,\|f'(c)\| + o_f(h) + g(c) - g(a) + \varepsilon(c - a) + \varepsilon \\ &\leq h\,g'(c) + g(c) - g(a) + \varepsilon(c - a) + \varepsilon + o_f(h)\,.\end{aligned}$$

Puisque

$$g(c+h) = g(c) + h\,g'(c) + o_g(h)\,,$$

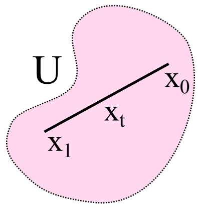
on obtient finalement

$$\|f(c+h) - f(a)\| \leq g(c+h) - g(a) + \varepsilon(c - a) + \varepsilon - o_g(h) + o_f(h)\,,$$

et donc $c+h \in I_\varepsilon$ dès que $h > 0$ est assez petit pour que $o_f(h) - o_g(h) \leq \varepsilon h$.
□

Remarque 6.15 *Noter que le raisonnement que l'on vient de mener est un raisonnement par connexité. On a en effet commencé par montrer que I_ε est une partie non vide et fermée de l'intervalle $[a, b]$. La dernière étape consistant à montrer que $I_\varepsilon \subset [a, b]$ est ouvert.*

Pour évaluer l'accroissement d'une application f entre les points x_0 et x_1 en dimension quelconque, on se ramène à la dimension 1 en travaillant sur le segment $[x_0, x_1] \subset \mathbb{R}^n$, que l'on paramètre par l'intervalle $[0,1]$ via l'application $t \in [0,1] \to x_t = x_0 + t(x_1 - x_0) \in [x_0, x_1]$. On doit supposer que ce segment est entièrement contenu dans le domaine de définition de f.



Le segment $[x_0, x_1] \subset U$

Corollaire 6.16 *Soient $U \subset \mathbb{R}^n$ un ouvert, $f : U \to \mathbb{R}^p$ différentiable et $[x_0, x_1] \subset U$ un segment. Si on a $\|D_{x_t} f\| \leq M$ pour tout réel $t \in [0,1]$, alors*

$$\|f(x_1) - f(x_0)\| \leq M \, \|x_1 - x_0\| \, .$$

Preuve Rappelons que $x_t = x_0 + t(x_1 - x_0)$. On applique le théorème des accroissements finis à

$$\begin{aligned} F &: t \in [0,1] \to f(x_t) \in \mathbb{R}^p \\ g &: t \in [0,1] \to M\|x_1 - x_0\| t \in \mathbb{R} \, . \end{aligned}$$

Les applications F et g sont en effet dérivables avec, pour tout $t \in [0,1]$,

$$\|F'(t)\| = \|D_{x_t} f \; (x_1 - x_0)\| \leq M\|x_1 - x_0\| = g'(t) \, .$$

$\square$

Corollaire 6.17 *Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$ différentiable.*

1. *On suppose que U est convexe, et que pour tout point $x \in U$, on a $\|D_x f\| \leq M$. Alors f est M-lipschitzienne.*
2. *On suppose que U est connexe, et que pour tout point $x \in U$, on a $D_x f = 0$. Alors f est constante.*

Preuve Le premier point suit immédiatement du corollaire 6.16.

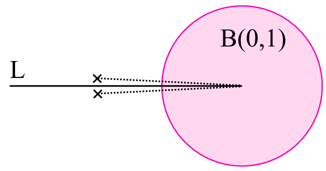
Pour le second, on fixe $x_0 \in U$ et montre que l'ensemble

$$V := \{x \in U \mid f(x) = f(x_0)\}$$

est non vide (il contient x_0), fermé (l'application f est continue) et ouvert (ceci suit du premier point qui assure que toute boule $B(x,r) \subset U$ centrée en un point $x \in V$ est incluse dans V, par convexité de la boule). $\square$

Remarques L'hypothèse de connexité dans 2 est indispensable. Considérer par exemple la fonction $f : x \in \mathbb{R}^* \to \text{signe}(x) \in \{-1, 1\}$. Lorsque $U \subset \mathbb{R}^n$ est un ouvert quelconque, il suit seulement que f est localement constante sur U et donc constante sur chaque composante connexe de U.

L'hypothèse de convexité dans 1 est indispensable. Penser à la fonction "détermination principale de l'argument" $\vartheta : \mathbb{R}^2 \setminus L \to]-\pi, \pi[$, définie sur le complémentaire de l'axe réel négatif L. Ses dérivées partielles sont



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

La différentielle $D_{(x,y)}\vartheta$ est donc uniformément bornée en dehors de la boule unité. On constate pourtant que $\lim_{\varepsilon \to 0} \vartheta(-2, \varepsilon) - \vartheta(-2, -\varepsilon) = 2\pi$.

Continuons avec un petit raffinement. Dans ce dernier énoncé, le réel α a vocation à être petit : il faut penser que x_1 sera choisi proche de x_0 et que l'application f sera de classe C^1 (définition 7.8).

Corollaire 6.18 *Soient $U \subset \mathbb{R}^n$ un ouvert, $f : U \to \mathbb{R}^p$ différentiable et $[x_0, x_1] \subset U$ un segment.*

Si, pour tout réel $t \in [0, 1]$, on a $\|D_{x_t} f - D_{x_0} f\| \leq \alpha$, alors

$$\|f(x_1) - f(x_0) - D_{x_0} f\ (x_1 - x_0)\| \leq \alpha\, \|x_1 - x_0\|\,.$$

Preuve On applique le corollaire 6.17 à l'application

$$\tilde{f} : t \in [0, 1] \to f(x_t) - D_{x_0} f\ (x_t) \in \mathbb{R}^p$$

de dérivée

$$\tilde{f}'(t) = (D_{x_t} f - D_{x_0} f)\ (x_1 - x_0)\,.$$

□

Le théorème des accroissements finis et ses applications sont valables, avec les mêmes démonstrations, pour une application $f : U \subset E \to F$, où E et F sont des espaces vectoriels normés de dimension quelconque.

7. Applications de classe C^1

Ce chapitre est pour l'essentiel consacré aux applications de classe C^1. On commence par définir les dérivées directionnelles et les dérivées partielles.

A Dérivées directionnelles

Définition 7.1 *Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$. Soient $x \in U$ et $u \in \mathbb{R}^n$ un vecteur (non nul). On dit que f admet au point x une dérivée dans la direction du vecteur u lorsque la restriction de f à la droite affine $x+\mathbb{R}\, u$ est dérivable au point x, ou encore lorsque la fonction de variable réelle (définie au voisinage de l'origine)*

$$t \in \mathbb{R} \to f(x+tu) \in \mathbb{R}^p$$

admet une dérivée en $t = 0$. On note alors

$$f'_x(u) = \lim_{t\to 0} \frac{f(x+tu) - f(x)}{t}$$

cette dérivée.

Exemple 7.2 Pour $f : \mathbb{R} \to \mathbb{R}^p$ différentiable au point $x \in \mathbb{R}$, la dérivée $f'(x)$ est égale à la dérivée directionnelle $f'_x(1)$ associée au vecteur $1 \in \mathbb{R}$.

Lemme 7.3 *Si f est différentiable en x, elle admet en ce point des dérivées dans toutes les directions. De plus, on a pour tout $u \in \mathbb{R}^n$ l'égalité*

$$f'_x(u) = D_x f(u)\,,$$

et la dérivée directionnelle $f'_x(u)$ dépend alors linéairement du vecteur u.

Preuve Immédiat. □

Remarque Par contre, il se peut que f possède en x des dérivées dans toutes les directions, sans pour autant être différentiable en ce point. Considérer par exemple, en l'origine, l'application $f : \mathbb{R}^2 \to \mathbb{R}$ définie par

$$f(te^{i\theta}) = \theta\, t, \text{ pour } t \in \mathbb{R} \text{ et } \theta \in [0, \pi[.$$

On peut même construire une application qui admet en x des dérivées dans toutes les directions, et qui n'est pas continue en ce point : l'existence de dérivées directionnelles au point x apporte donc peu d'information quant à la régularité de f en x.

B Dérivées partielles, matrice jacobienne

Les dérivées partielles d'une application $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$ sont ses dérivées dans la direction des vecteurs de base. Notons $(e_1, \cdots, e_n)$ la base canonique de $\mathbb{R}^n$, $x = (x_1, \cdots, x_n)$ les coordonnées d'un point x de $\mathbb{R}^n$ relativement à cette base.

Définition 7.4 *Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$. Soit $x \in U$. Les dérivées partielles de f en x, lorsqu'elles existent, sont les dérivées de f en x dans la direction des vecteurs de la base canonique. On note alors, pour $i = 1...n$,*

$$\frac{\partial f}{\partial x_i}(x) = f'_x(e_i) = \lim_{t \to 0} \frac{f(x + te_i) - f(x)}{t}\,.$$

Bien entendu, si f est différentiable en x, elle admet des dérivées partielles en ce point, avec

$$\frac{\partial f}{\partial x_i}(x) = D_x f(e_i)\,.$$

Dans ce cas, on a pour tout vecteur $h = (h_1, \cdots, h_n) \in \mathbb{R}^n$

$$(D_x f)(h) = \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(x)\,.$$

L'existence de dérivées en x dans toutes les directions prend en compte la restriction de f à chaque droite affine passant par x. On a déjà dit que ça laissait à f la possibilité de ne même pas être continue en x. L'existence de dérivées partielles en ce point ne prend en compte que la restriction de f aux droites passant par x, et qui sont parallèles aux axes de coordonnées. C'est donc une information *a priori* minime. Voir cependant le théorème 7.14.

Définition 7.5 *Soit $f = (f_1, \cdots, f_p) : U \subset \mathbb{R}^n \to \mathbb{R}^p$, différentiable au point $x \in U$. La matrice de l'application linéaire $D_x f \in L(\mathbb{R}^n, \mathbb{R}^p)$ exprimée dans les bases canoniques de $\mathbb{R}^n$ et $\mathbb{R}^p$ est la matrice jacobienne de f en x, soit*

$$J_x f = \left(\frac{\partial f_i}{\partial x_j}\right)_{\substack{1 \le i \le p \\ 1 \le j \le n}} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_p}{\partial x_1} & \cdots & \frac{\partial f_p}{\partial x_n} \end{pmatrix} = \begin{pmatrix} \vdots & & \vdots \\ \frac{\partial f}{\partial x_1} \vdots & \cdots & \vdots \frac{\partial f}{\partial x_n} \\ \vdots & & \vdots \end{pmatrix} \in M_{p,n}\mathbb{R}$$

(toutes ces dérivées partielles étant évaluées au point x).

Pour une composée d'applications différentiables $g \circ f$ avec f différentiable en x et g différentiable en $f(x)$ il suit donc de la proposition 6.10 que le jacobienne de la composée est le produit (matriciel) des jacobiennes, soit :

$$J_x(g \circ f) = J_{f(x)}g \; J_x f \, .$$

Exercice 7.6 On désigne respectivement par $(x_i)_{1\leq i\leq n}$, $(y_j)_{1\leq j\leq p}$ et $(z_k)_{1\leq k\leq q}$ les coordonnées de $\mathbb{R}^n$, de $\mathbb{R}^p$ et de $\mathbb{R}^q$.
Soient $f : \mathbb{R}^n \to \mathbb{R}^p$ et $g : \mathbb{R}^p \to \mathbb{R}^q$ des fonctions différentiables, et la composée $h = g \circ f = (h_k)_{1\leq k\leq q}$. Démontrer, pour $x \in \mathbb{R}^n$, $1 \leq i \leq n$ et $1 \leq k \leq q$, l'égalité

$$\frac{\partial h_k}{\partial x_i}(x) = \sum_{j=1}^{p} \frac{\partial g_k}{\partial y_j}(f(x)) \, \frac{\partial f_j}{\partial x_i}(x) \, .$$

Exercice 7.7 On note (x, y) les coordonnées dans $\mathbb{R}^2$. Soit $f : \mathbb{R}^2 \to \mathbb{R}$ une fonction différentiable. Soit $u : t \in \mathbb{R} \to f(t, f(t, f(t, t))) \in \mathbb{R}$. Comme composée de fonctions différentiables, u est dérivable. On se propose de déterminer sa dérivée. A première vue, ce n'est pas très engageant... Alors on prend les choses posément !

1. Soit $g : t \in \mathbb{R} \to f(t, t) \in \mathbb{R}^2$. Montrer, pour tout $t \in \mathbb{R}$, l'égalité $g'(t) = \frac{\partial f}{\partial x}(t, t) + \frac{\partial f}{\partial y}(t, t)$.
2. Soit $h : t \in \mathbb{R} \to f(t, f(t, t)) = f(t, g(t)) \in \mathbb{R}^2$. Exprimer $h'(t)$ en fonction des dérivées partielles de f.
3. En remarquant que $u(t) = f(t, h(t))$, déterminer enfin $u'(t)$.

C Applications de classe C^1

Nous nous intéressons désormais à une application $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$ différentiable (c'est-à-dire, rappelons le, différentiable en chaque point x de l'ouvert U). Nous disposons donc de l'application “différentielle” de f, soit

$$Df : x \in U \to D_x f \in L(\mathbb{R}^n, \mathbb{R}^p) \, .$$

Définition 7.8 *Une application $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$ différentiable est dite de classe C^1 lorsque sa différentielle*

$$Df : U \subset \mathbb{R}^n \to L(\mathbb{R}^n, \mathbb{R}^p)$$

est une application continue.

L'espace vectoriel $L(\mathbb{R}^n, \mathbb{R}^p)$ est de dimension finie (égale à np). Nous n'avons donc pas besoin de préciser la norme avec laquelle nous travaillons.

Lemme 7.9 *Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$ une application différentiable. Alors f est de classe C^1 si et seulement si ses dérivées partielles*

$$\frac{\partial f}{\partial x_i} : x \in U \to \frac{\partial f}{\partial x_i}(x) \in \mathbb{R}^p$$

sont continues ($1 \leq i \leq n$).

On prendra soin de ne pas confondre cette remarque, banale, avec l'énoncé plus conséquent du théorème 7.14 dans lequel on ne suppose pas *a priori* l'application f différentiable.

Preuve L'application

$$\mathcal{J} : L \in L(\mathbb{R}^n, \mathbb{R}^p) \to (L(e_1), \cdots, L(e_n)) \in (\mathbb{R}^p)^n$$

(évaluation de l'application linéaire L sur les vecteurs de base $(e_1, \cdots, e_n)$) est un isomorphisme d'espaces vectoriels. Elle est donc bicontinue puisque nous sommes en dimension finie. Le résultat suit. □

Exemple 7.10 Une fonction réelle de variable réelle $f : U \subset \mathbb{R} \to \mathbb{R}$ est donc de classe C^1 si et seulement si elle est dérivable en chaque point, et si sa dérivée $f' : U \to \mathbb{R}$ est continue : on retrouve la définition habituelle.

Rappelons bien sûr que, même en dimension 1, il existe des fonctions partout dérivables qui ne sont pas de classe C^1. Par exemple la fonction $t \to t^2 \sin(1/t)$.

D Exemples

• Une somme d'applications $f, g : U \subset \mathbb{R}^n \to \mathbb{R}^p$ de classe C^1 est de classe C^1.

• Le produit d'une application $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$ de classe C^1 par un scalaire $\lambda \in \mathbb{R}$ est encore de classe C^1.

• Une application linéaire ou affine est de classe C^1.

En effet, sa différentielle est une application constante.

• Une application bilinéaire $B : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^q$ est de classe C^1.

On a vu (6.D) que l'application bilinéaire B est différentiable avec, pour tous $x = (x_1, x_2)$ et $h = (h_1, h_2)$ dans $\mathbb{R}^n \times \mathbb{R}^p$,

$$D_xB\ (h) = B(x_1, h_2) + B(h_1, x_2)\,.$$

On constate que l'application différentielle de B

$$DB : x \in \mathbb{R}^n \times \mathbb{R}^q \to D_xB \in L(\mathbb{R}^n \times \mathbb{R}^p, \mathbb{R}^q)$$

est une application linéaire, donc continue (dimension finie).

• Plus généralement, une application k-linéaire $f : E_1 \times \cdots \times E_k \to F$ (où les E_i et F sont des espaces vectoriels de dimension finie) est de classe C^1.

On a vu en (6.D) que f est différentiable avec, pour tous $x = (x_1, \cdots, x_k)$ et $h = (h_1, \cdots, h_k)$ dans $E := E_1 \times \cdots \times E_k$,

$$D_x f(h) = \sum_{i=1}^{k} f(x_1, \cdots, x_{i-1}, h_i, x_{i+1}, \cdots, x_k) = \sum_{i=1}^{k} (\ell_i(x))(h)\,.$$

Pour voir que la différentielle $Df = \sum_{i=1}^{k} \ell_i : E \to L(E, F)$ est continue, nous allons montrer que chacune des applications $\ell_i : E \to L(E, F)$ est continue. On peut supposer sans perte de généralité que $i = k$. On écrit alors l'application $\ell_k = \Phi_k \circ p_k$ comme composée de l'application de projection

$$p_k : (x_1, \cdots, x_k) \in E_1 \times \cdots \times E_k \to (x_1, \cdots, x_{k-1}) \in E_1 \times \cdots \times E_{k-1}$$

(linéaire, donc continue) et de l'application $(k-1)$-linéaire

$$\Phi_k : (y_1, \cdots, y_{k-1}) \in E_1 \times \cdots \times E_{k-1} \to [h \in E \to f(y_1, \cdots, y_{k-1}, h_k)] \in L(E, F)$$

qui est donc également continue.

Exercice 7.11 Montrer que l'application déterminant, soit $\det : \mathrm{M}_n\mathbb{R} \to \mathbb{R}$, est de classe C^1. Déterminer sa différentielle en toute matrice $M \in \mathrm{Gl}_n\mathbb{R}$ inversible, puis sur tout $\mathrm{M}_n\mathbb{R}$.

Une application multilinéaire continue $f : E_1 \times \cdots \times E_k \to F$ entre espaces vectoriels normés de dimensions quelconques est également de classe C^1.

Proposition 7.12 *Une composée d'applications $f : U \subset \mathbb{R}^n \to V \subset \mathbb{R}^p$ et $g : V \subset \mathbb{R}^p \to \mathbb{R}^q$ de classe C^1 est encore de classe C^1.*

Preuve Ceci résulte de l'expression pour la différentielle d'une composée (6.2). En effet, la différentielle de $g \circ f$ s'écrit comme composée des applications

$$x \in U \longrightarrow (D_{f(x)}g, D_x f) \in L(\mathbb{R}^p, \mathbb{R}^q) \times L(\mathbb{R}^n, \mathbb{R}^p)$$

(continue car f, Df et Dg le sont par hypothèse) et de l'application "composer deux applications linéaires"

$$(\Phi, \Psi) \in L(\mathbb{R}^p, \mathbb{R}^q) \times L(\mathbb{R}^n, \mathbb{R}^p) \longrightarrow \Phi \circ \Psi \in L(\mathbb{R}^n, \mathbb{R}^q)$$

(continue, car bilinéaire en dimension finie). □

Exercice 7.13 Soit $k \in \mathbb{N}$ un entier. L'application

$$e_k : M \in \mathrm{M}_n\mathbb{R} \to M^k \in \mathrm{M}_n\mathbb{R}$$

est de classe C^1. On a, pour tout $H \in \mathrm{M}_n\mathbb{R}$,

$$D_M(e_k)(H) = H\ M^{k-1} + M\ H\ M^{k-2} + M^2\ H\ M^{k-3} + \cdots + M^{k-1}\ H$$

(k termes).

E Différentiabilité versus dérivées partielles

Nous avons remarqué plus haut que l'existence de dérivées partielles en un point pour une application $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$ est très peu restrictive. Par contraste, on a le résultat suivant qui va découler du théorème des accroissements finis.

Théorème 7.14 *Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$. Soit $m_0 \in U$. On suppose que f possède des dérivées partielles en chaque point $m \in U$, et que celles-ci sont continues au point m_0. Alors f est différentiable en m_0.*

Remarque L'hypothèse signifie que chacune des dérivées partielles

$$(x_1, \cdots, x_n) \longrightarrow \frac{\partial f}{\partial x_i}(x_1, \cdots, x_n)$$

est continue, au point m_0, par rapport à l'ensemble des variables.

On pourra pour l'essentiel se contenter de retenir le corollaire suivant.

Corollaire 7.15 *Une application $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$ est de classe C^1 si et seulement si ses dérivées partielles existent et sont continues sur U.*

Preuve La partie directe a déjà été évoquée. La réciproque est conséquence immédiate du théorème 7.14, et du lemme 7.9. □

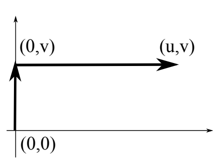
Preuve du théorème 7.14 Pour alléger les notations, on se limite à la dimension 2 et on laisse au lecteur le soin d'écrire la preuve en dimension quelconque : elle ne requiert aucune initiative supplémentaire. On notera (x, y) les coordonnées dans $\mathbb{R}^2$.

On se ramène par translation au cas où m_0 est l'origine. On veut montrer que f est différentiable en 0. Si c'est bien le cas, on sait que sa différentielle en ce point est alors l'application linéaire

$$(u, v) \to u\,\frac{\partial f}{\partial x}(0,0) + v\,\frac{\partial f}{\partial y}(0,0)\,.$$

Nous devons donc montrer que la quantité suivante est un $o(u, v)$:

$$\begin{aligned} f(u,v) - f(0,0) - u\,\frac{\partial f}{\partial x}(0,0) - v\,\frac{\partial f}{\partial y}(0,0) \\ &= f(u,v) - f(0,v) - u\,\frac{\partial f}{\partial x}(0,0) \\ &+ f(0,v) - f(0,0) - v\,\frac{\partial f}{\partial y}(0,0)\,. \end{aligned}$$



Fixons donc $\varepsilon > 0$. Concernant la seconde ligne, la définition de la dérivée partielle $\partial f/\partial y$ en l'origine montre qu'il existe $R_1 > 0$ tel que

$$\|f(0,v) - f(0,0) - v\,\frac{\partial f}{\partial y}(0,0)\| \leq \varepsilon|v|\,,$$

lorsque $|v| \leq R_1$.

Pour estimer l'expression figurant sur la première ligne, on procède comme dans le corollaire 6.18 (conséquence des accroissements finis) et on introduit la famille de fonctions d'une variable réelle, paramétrée par v, définie par

$$g_v(t) := f(t,v) - t\,\frac{\partial f}{\partial x}(0,0)$$

pour $t \in [0,u]$. On a

$$\frac{d}{dt}(g_v(t)) = \frac{\partial f}{\partial x}(t,v) - \frac{\partial f}{\partial x}(0,0)\,.$$

La continuité de $\partial f/\partial x$ en l'origine assure donc qu'il existe $R_2 > 0$ tel que $|d/dt\ g_v(t)| \leq \varepsilon$ dès que $|t| + |v| \leq R_2$. Il suit alors du théorème des accroissements finis que, lorsque $|u| + |v| \leq R_2$, on a

$$|g_v(u) - g_v(0)| = |f(u,v) - f(0,v) - u\,\frac{\partial f}{\partial x}(0,0)| \leq \varepsilon|u|\,,$$

ce qu'on voulait. □

F Suites d'applications de classe C^1

On sait qu'une limite simple d'applications continues n'est pas forcément continue. Par contre la continuité est conservée lorsque la convergence est uniforme (théorème 1.46).

De même, sans hypothèse sur les différentielles, une limite uniforme d'applications de classe C^1 ne sera pas forcément de classe C^1, ni même différentiable. Considérer la suite $f_n : t \in \mathbb{R} \to (t^2 + 1/n^2)^{1/2} \in \mathbb{R}$ qui converge uniformément vers la fonction $f : t \in \mathbb{R} \to |t| \in \mathbb{R}$ qui n'est pas dérivable en l'origine, alors que les f_n sont toutes de classe C^1.

Par contre, tout s'arrange si l'on impose à la suite des différentielles d'être uniformément convergente, au moins localement.

Définition 7.16 *Soient X et Y deux espaces métriques, et $\varphi_k, \varphi : X \to Y$ ($k \in \mathbb{N}$) des applications. On dit que la suite $(\varphi_k)_{k\in\mathbb{N}}$ converge localement uniformément vers φ si, pour tout point $x \in X$, il existe un voisinage $V_x \subset X$ de x sur lequel les restrictions $\varphi_{k|V_x}$ convergent uniformément vers $\varphi_{|V_x}$ lorsque $k \to \infty$.*

Théorème 7.17 *Soient $U \subset \mathbb{R}^n$ un ouvert et $f_k : U \to \mathbb{R}^p$ une suite d'applications de classe C^1. On suppose*

1. *qu'il existe un point $x_0 \in U$ dont la suite $(f_k(x_0))_{k\in\mathbb{N}}$ des images converge dans $\mathbb{R}^p$*
2. *que l'ouvert U est connexe*
3. *que la suite $Df_k : U \to L(\mathbb{R}^n, \mathbb{R}^p)$ des différentielles des f_k converge localement uniformément vers une application $\Phi : U \to L(\mathbb{R}^n, \mathbb{R}^p)$.*

Alors la suite d'applications (f_k) converge localement uniformément vers une application $f : U \to \mathbb{R}^p$. Cette application f est de classe C^1 et sa différentielle est

$$Df = \lim_k Df_k\,.$$

En d'autres termes, la différentielle de la limite est la limite des différentielles : on peut intervertir limite et différentielle.

Remarques On a donné un énoncé un peu élaboré dans lequel on ne suppose pas *a priori* la convergence, même simple, de la suite (f_k) sur tout U. On suppose juste que celle-ci converge en un point. Dans ce cas, la connexité de U (ainsi que l'hypothèse faite sur les différentielles) permet de propager la convergence de (f_k) à tout l'ouvert U.

On pourrait remplacer les conditions 1. et 2. par le fait que la suite d'applications (f_k) converge simplement sur U (dans la pratique, c'est souvent assez évident de s'en rendre compte). Cependant, on serait de toutes façons amenés à écrire l'estimation (7.1).

Il est quand même nécessaire d'imposer la convergence en un point. Penser à la suite de fonctions constantes $f_k : t \in \mathbb{R} \to k \in \mathbb{R}$ $(k \in \mathbb{N})$.

Preuve 1. *On commence par montrer que la suite d'applications (f_k) converge simplement sur U.*

On introduit l'ensemble de convergence de la suite, soit

$$\mathcal{C} = \{x \in U \mid \text{la suite } (f_k(x)) \text{ converge}\} \subset U\,.$$

On veut montrer que $\mathcal{C} = U$. Comme on l'a déjà dit, on va utiliser la connexité de U. Par hypothèse, $\mathcal{C}$ est non vide. Il nous reste donc à montrer que $\mathcal{C}$ est une partie ouverte et fermée de U (attention, topologie induite !). Avec cette définition de $\mathcal{C}$, ce n'est pas bien commode. On remarque donc que, puisque l'espace $\mathbb{R}^p$ est complet, on a aussi

$$\mathcal{C} = \{x \in U \mid \text{la suite } (f_k(x)) \text{ est de Cauchy}\}\,.$$

Soit donc $B(x,R) \subset U$ une boule. Cette boule est convexe. Le théorème des accroissements finis (corollaire 6.17) assure donc que, pour tous $y, z \in B(x,R)$ et tous $k, \ell \in \mathbb{N}$, on a

$$\|(f_\ell(y) - f_k(y)) - (f_\ell(z) - f_k(z))\| \leq \sup_{B(x,R)} \|Df_\ell - Df_k\|\, \|y - z\|\,. \tag{7.1}$$

Supposons maintenant que la suite (Df_k) converge uniformément sur la boule $B(x,R) \subset U$, et qu'il existe $y \in B(x,R)$ tel que la suite $(f_k(y))$ converge (ou soit de Cauchy). Il suit alors de l'inégalité (7.1) que, pour tout $z \in U$, la suite $(f_k(z))$ est de Cauchy, donc converge. Ceci assure bien que $\mathcal{C} \subset U$ est ouvert et fermé dans U (le lecteur doit s'en convaincre !) et donc que la suite (f_k) converge simplement sur U.

2. *La suite d'applications (f_k) converge localement uniformément sur U.*
Il suit immédiatement de la majoration (7.1) que la suite d'applications (f_k) satisfait, localement, un critère de Cauchy uniforme.

3. *L'application $f = \lim f_k$ est différentiable, et $Df = \lim Df_k = \Phi$.*
Soient $x \in U$. On veut montrer que f est différentiable au point x, de différentielle $D_x f = \Phi(x)$. On reprend l'inégalité (7.1), que l'on applique à $y = x + h \in B(x,R)$ et $z = x$. En faisant tendre ℓ vers ∞, il vient

$$\|(f(x+h) - f_k(x+h)) - (f(x) - f_k(x))\| \leq \sup_{a \in B(x,R)} \|\Phi(a) - D_a f_k\| \, \|h\| \,. \quad (7.2)$$

L'inégalité triangulaire donne alors

$$\begin{aligned} &\|f(x+h) - f(x) - (\Phi(x))(h)\| \\ &\qquad \leq \|(f(x+h) - f_k(x+h)) - (f(x) - f_k(x))\| \\ &\qquad + \|D_x f_k(h) - (\Phi(x))(h)\| \\ &\qquad + \|f_k(x+h) - f_k(x) - D_x f_k(h)\| \,. \end{aligned}$$

On veut montrer que le terme de gauche est un o de h. On va estimer chacun des trois termes intervenant dans cette majoration. Attention, il faut s'occuper de ces termes dans le bon ordre !

Soit $\varepsilon > 0$. Choisissons $k \in \mathbb{N}$ tel que $\sup_{a \in B(x,R)} \|\Phi(a) - D_a f_k\| \leq \varepsilon$ (nous insistons sur le fait que k est désormais fixé). Il suit donc de (7.2) que, pour tout $\|h\| < R$, chacun des deux premiers termes est au plus égal à $\varepsilon \|h\|$.

Pour cette valeur de k, l'application f_k est différentiable au point x. Il existe donc η pour lequel, pour tout $\|h\| \leq \eta$, le troisième est majoré par

$$\|f_k(x+h) - f_k(x) - D_x f_k(h)\| \leq \varepsilon \|h\| \,.$$

4. *L'application $f = \lim f_k$ est de classe C^1.*
C'est maintenant immédiat puisque $Df = \lim Df_k$ est continue, comme limite uniforme locale d'applications continues.

$\square$

Ce résultat s'étend, avec la même démonstration, aux suites (f_k) d'applications de classe C^1, avec $f_k : U \subset E \to F$ où E et F sont deux espaces vectoriels normés de dimension quelconque, en supposant F complet.

Si l'on suppose d'emblée la suite (f_k) ponctuellement convergente, l'énoncé reste vrai pour deux deux espaces vectoriels normés E et F quelconques (F n'a plus besoin d'être complet) et l'hypothèse de connexité de U est alors superflue.

G Un exemple : l'exponentielle matricielle

Il va sans dire que le théorème que nous venons de voir se transpose sans difficulté aux séries d'applications C^1. Il suffit d'introduire la suite des sommes partielles. On a alors le

Corollaire 7.18 *Soient $U \subset \mathbb{R}^n$ un ouvert connexe, $f_k : U \to \mathbb{R}^p$ une suite d'applications de classe C^1. On suppose que la série $\sum f_k$ converge en un point de U, et que la série des différentielles $\sum Df_k$ converge localement uniformément sur U. Alors la série d'applications $\sum f_k$ converge localement uniformément sur tout U, sa somme f est de classe C^1 et on a*

$$Df = D(\sum f_k) = \sum (Df_k)\,.$$

Autrement dit on peut différentier la somme terme à terme.
Attention, ne pas confondre ce résultat (somme infinie) avec le fait, immédiat, qu'une somme finie d'applications de classe C^1 est encore de classe C^1 !

Un critère commode pour assurer la convergence uniforme d'une série de fonctions à valeurs dans un Banach est donné par la convergence normale de cette série.

Proposition-Définition 7.19 *On munit $\mathbb{R}^n$ d'une norme. Soient X un espace métrique et $f_k : X \to \mathbb{R}^n$ une suite d'applications bornées . On note $\|f_k\|_\infty = \sup_{x\in X} \|f_k(x)\|$ pour $k \in \mathbb{N}$.*
On dit que la série de fonctions $\sum f_k$ converge normalement sur X lorsque $\sum \|f_k\|_\infty < \infty$. Dans ce cas, la série de fonctions $\sum f_k$ est uniformément convergente sur X.

Preuve Lorsque la série de fonctions $\sum f_k$ converge normalement sur X, la suite des sommes partielles $s_p = \sum_0^p f_k$ satisfait une condition de Cauchy uniforme sur X. On conclut car $\mathbb{R}^n$ est complet. □

Une illustration fondamentale du corollaire 7.18 est fournie par l'exponentielle matricielle, qui interviendra notamment lors de l'étude des équations différentielles linéaires à coefficients constants.

Théorème 7.20 *L'application exponentielle matricielle*

$$\exp : M \in \mathrm{M}_n\mathbb{R} \to \sum_{k=0}^{\infty} \frac{M^k}{k!} \in \mathrm{M}_n\mathbb{R}$$

est de classe C^1 et sa différentielle se calcule en dérivant terme à terme.

En particulier,

$$D_0\exp = \mathrm{Id}\,.$$

Plus généralement, lorsque deux matrices M et H commutent c'est-à-dire lorsque $MH = HM$, on a

$$D_M\exp H = \exp(M)\; H = H\,\exp(M)\,.$$

Remarque 7.21 *On verra dans l'exercice 9.22 que l'exponentielle matricielle est une application de classe C^∞. Cela suivra du théorème 9.21 (suites d'applications C^k).*

Preuve D'après le corollaire 7.18, on pourrait se contenter de constater que la série converge en un point, par exemple en $M = 0$, ce qui n'est pas sorcier. Rappelons cependant que nous avons déjà montré l'existence de l'application exponentielle (théorème 4.25). Montrons maintenant qu'elle est de classe C^1.

Nous avons vu à l'exercice 7.13 que chaque application

$$\mathrm{e}_k : M \in \mathrm{M}_n\mathbb{R} \to M^k \in \mathrm{M}_n\mathbb{R}$$

est de classe C^1, de différentielle

$$D_M(\mathrm{e}_k)(H) = H\;M^{k-1} + M\;H\;M^{k-2} + M^2\;H\;M^{k-3} + \cdots + M^{k-1}\;H\,.$$

On a donc (en munissant $\mathrm{M}_n\mathbb{R}$ d'une norme d'algèbre)

$$\|D_M(\mathrm{e}_k)(H)\| \leq k\;\|M^{k-1}\|\;\|H\|\,,$$

soit

$$\|D_M(\mathrm{e}_k)\| \leq k\;\|M^{k-1}\|\,.$$

Cette estimée assure la convergence normale, donc uniforme, sur chaque partie bornée de $\mathrm{M}_n\mathbb{R}$ de la série d'applications

$$\sum_{k=0}^{\infty} \frac{D\mathrm{e}_k}{k!} : \mathrm{M}_n\mathbb{R} \to L(\mathrm{M}_n\mathbb{R}, \mathrm{M}_n\mathbb{R})\,.$$

Le corollaire 7.18 s'applique donc et montre que exp est de classe C^1 avec, pour tous $M, H \in \mathrm{M}_n\mathbb{R}$,

$$D_M(\exp)(H) = \sum_{k=0}^{\infty} \frac{1}{k!} D_M(\mathrm{e}_k)(H)\,.$$

En particulier, lorsque H et M commutent, il vient

$$D_M(\exp)(H) = \sum_{k=0}^{\infty} \frac{1}{k!} M^k H = (\sum_{k=0}^{\infty} \frac{1}{k!} M^k)\, H = \exp(M)\, H\,,$$

l'avant-dernière dernière égalité venant de la continuité de la multiplication matricielle, qui assure que

$$\lim_{K\to\infty} \big(\sum_{k=0}^{K} \frac{1}{k!} M^k H\big) = \lim_{K\to\infty} \big((\sum_{k=0}^{K} \frac{1}{k!} M^k) H\big) = \big(\lim_{K\to\infty} \sum_{k=0}^{K} \frac{1}{k!} M^k\big)\, H\,.$$

$\square$

8. Inversion locale, fonctions implicites

A Théorème d'inversion locale

Définition 8.1 *Un C^1 difféomorphisme entre deux ouverts $U, V \subset \mathbb{R}^n$ est une application $f : U \to V$ bijective, qui est de classe C^1 ainsi que sa réciproque $f^{-1} : V \to U$.*

Remarque 8.2 *Observer alors que, puisque $f^{-1} \circ f = \mathrm{Id} : U \to U$, on obtient pour tout point $x \in U$ que $D_{f(x)}f^{-1} \cdot D_x f = \mathrm{Id}$. En particulier une condition nécessaire pour que f puisse être un difféomorphisme est que chacune des application linéaire $D_x f \in L(\mathbb{R}^n, \mathbb{R}^n)$ $(x \in U)$ soit inversible.*

Définition 8.3 *On dit que deux ouverts U, V de $\mathbb{R}^n$ sont difféomorphes lorsqu'il existe un difféomorphisme $f : U \to V$.*

Il suit de la remarque que deux ouverts $U \subset \mathbb{R}^n$ et $V \subset \mathbb{R}^p$ ne peuvent être difféomorphes lorsque les dimensions $n \neq p$ sont différentes. En effet une application linéaire $L \in L(\mathbb{R}^n, \mathbb{R}^p)$ ne peut être inversible que lorsque $n = p$. Rappelons que nous avons déjà évoqué le "théorème d'invariance du domaine" 5.17, beaucoup plus délicat à démontrer (sauf lorsqu'une des dimensions est 1), et qui stipule qu'un ouvert $\mathbb{R}^n$ ne peut être homéomorphe à un ouvert de $\mathbb{R}^p$ que lorsque $n = p$.

Dans ce qui suit, on se donne une application $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ de classe C^1 dont la différentielle en chaque point de Ω est une application linéaire inversible, et on se demande dans quelle mesure cette condition (nécessaire) suffit à assurer que f soit un difféomorphisme.

Commençons par le cas modèle. Soit $x \in \mathbb{R}^n \to b + Ax \in \mathbb{R}^n$ une application affine, d'application linéaire tangente $A \in \mathrm{Gl}_n\mathbb{R}$ inversible. Cette application est un difféomorphisme global de $\mathbb{R}^n$ sur lui-même.

Revenons à $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$, application de classe C^1 définie sur un ouvert Ω de $\mathbb{R}^n$. On suppose qu'il existe un point $x_0 \in \Omega$ où la différentielle $D_{x_0}f$ de f est inversible. Au voisinage de x_0, f "ressemble à son application affine tangente"

$$x \to f(x_0) + D_{x_0}f(x - x_0)\,.$$

Cela va nous permettre de montrer que f est localement injective, et que son application réciproque est également de classe C^1. C'est le contenu du théorème d'inversion locale que nous énonçons maintenant.

Théorème 8.4 Inversion locale
Soit $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ une application de classe C^1 définie sur un ouvert Ω de $\mathbb{R}^n$. Soit $x_0 \in \Omega$. On suppose que la différentielle $D_{x_0}f \in \mathrm{Gl}_n\mathbb{R}$ est inversible. Alors il existe un voisinage ouvert $U \subset \Omega$ de x_0 tel que $V = f(U) \subset \mathbb{R}^n$ soit un voisinage ouvert de $f(x_0)$, et tel que $f : U \to V$ réalise un C^1 difféomorphisme entre ces deux ouverts.

Remarque 8.5 *Remarquer que lorsque $D_{x_0}f \in L(\mathbb{R}^n, \mathbb{R}^n)$ est une application linéaire inversible, la différentielle $D_x f \in L(\mathbb{R}^n, \mathbb{R}^n)$ est encore une application linéaire inversible pour $x \in U$ proche de x_0. En effet l'application $x \in \Omega \to \det D_x f \in \mathbb{R}$ est continue, et non nulle au point x_0.*

Donnons maintenant la preuve du théorème d'inversion locale. Nous allons la décomposer en plusieurs étapes. On munit $\mathbb{R}^n$ d'une norme, et $L(\mathbb{R}^n, \mathbb{R}^n)$ de la norme subordonnée.

Preuve • On commence par se simplifier la vie. On peut en effet supposer, sans perte de généralité, que $x_0 = 0$ et $f(x_0) = 0$ (effectuer des translations). Puisque l'application linéaire $D_{x_0}f$ est un difféomorphisme de $\mathbb{R}^n$, on peut même supposer que $D_{x_0}f = \mathrm{Id}$, quitte à composer f avec l'application $(D_{x_0}f)^{-1}$. En d'autres termes, on remplace f par l'application $x \to (D_{x_0}f)^{-1}(f(x + x_0) - f(x_0))$.

• En un premier temps, nous montrons que f est un homéomorphisme local au voisinage de l'origine. Pour cela, nous comparons f à son application affine tangente en 0, i.e. à l'application Id. On pose donc, pour tout $x \in \Omega$, $f(x) = x + u(x)$. Autrement dit, on écrit $f = \mathrm{Id} + u$ comme une perturbation de l'identité.

Montrons déjà que f réalise une bijection entre deux voisinages de 0. Soient $x \in \Omega$ et $y \in \mathbb{R}^n$. Il est équivalent de demander que $y = f(x)$ ou que $x = y - u(x)$, i.e. que x est point fixe de $\varphi_y : x \in \Omega \to y - u(x) \in \mathbb{R}^n$. On a dit "point fixe", on pense donc à Picard : allons-y !

Puisque l'application $Df : x \in \Omega \to D_x f \in L(\mathbb{R}^n, \mathbb{R}^n)$ est continue, il existe une boule fermée $B_f(0, R) \subset \Omega$ pour laquelle on ait, pour tout $x \in B_f(0, R)$,

$$\|D_x\varphi_y\| = \|D_x u\| = \|D_x f - \mathrm{Id}\| = \|D_x f - D_0 f\| \leq 1/2\,. \qquad (8.1)$$

Le théorème des accroissements finis (théorème 6.13) assure alors que la fonction φ_y (ainsi que u) est 1/2-lipschitzienne sur la boule $B_f(0, R) \subset \Omega$.

Soit maintenant $y \in B(0, R/2)$. Puisque u est $(1/2)$-lipschitzienne sur $B_f(0, R)$ et vérifie $u(0) = 0$, on a pour $\|x\| \leq R$:

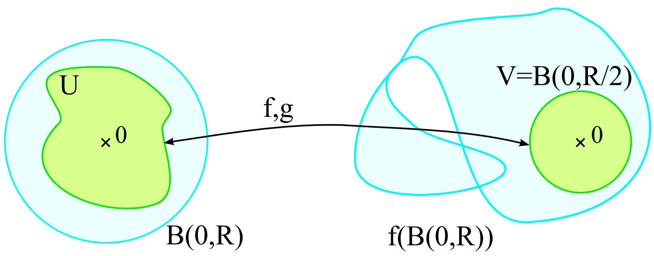
$$\|\varphi_y(x)\| = \|y - u(x)\| \leq \|y\| + (1/2)\|x\| < R\,.$$

Le théorème du point fixe (th. 3.20) s'applique donc à l'application

$$\varphi_y : B_f(0, R) \longrightarrow B(0, R) \subset B_f(0, R)$$

qui est contractante de l'espace complet $B_f(0, R)$ dans lui-même, et assure qu'il existe un unique $x \in B_f(0, R)$ pour lequel $\varphi_y(x) = x$. On a même $x \in B(0, R)$.

Posons alors $V = B(0, R/2)$ et $U = B(0, R) \cap f^{-1}(B(0, R/2))$. Puisque f est continue, $U \subset \Omega$ est ouvert. La discussion précédente assure que $f : U \to V$ est une bijection. Nous noterons $g : V \to U$ son inverse. Il nous faut encore montrer que g est de classe C^1.



• Préoccupons nous en un premier temps de la continuité de g, et montrons déjà que g est 2-lipschitzienne sur V. Tel est bien le cas puisque, pour $y_1, y_2 \in B(0, R/2)$ images respectives par f de $x_1, x_2 \in B(0, R)$, on a

$$\|y_1 - y_2\| \geq \|x_1 - x_2\| - \|u(x_1) - u(x_2)\| \geq (1/2)\, \|x_1 - x_2\| = (1/2)\, \|g(y_1) - g(y_2)\|\,.$$

• On montre ensuite que g est différentiable sur V. Soient $x \in U$ et $y \in V$ avec $y = f(x)$. Notons que, puisque $\|D_x f - \mathrm{Id}\| \leq 1/2$, il suit $D_x f$ est encore inversible avec $\|(D_x f)^{-1}\| \leq 2$ (voir l'exercice 4.32). Montrons que g est différentiable en y. Si c'est le cas, on sait *a priori* que $D_y g = (D_x f)^{-1}$. On estime donc, pour $y + k = f(x + h) \in V$ où $x + h \in U$:

$$\begin{aligned} g(y+k) - g(y) - (D_x f)^{-1}(k) &= h - (D_x f)^{-1}\left(f(x+h) - f(x)\right) \\ &= -(D_x f)^{-1}\left(f(x+h) - f(x) - D_x f(h)\right) \\ &= (D_x f)^{-1}\left(\|h\|\, \varepsilon(h)\right), \end{aligned}$$

où $\varepsilon(h)$ tend vers 0 avec h. Le terme de droite est bien un $o(\|k\|)$ puisque $\|h\| \leq 2\|k\|$ (g étant 2-lipschitzienne) et $\|(D_x f)^{-1}\| \leq 2$.

• Reste à voir, pour conclure, que g est bien de classe C^1. Sa différentielle

$$Dg : y \in V \to (D_{g(y)}f)^{-1} \in L(\mathbb{R}^n, \mathbb{R}^n)$$

est en effet continue comme composée

$$Dg = \mathcal{I} \circ Df \circ g$$

d'applications continues, avec $\mathcal{I} : L \in \mathrm{Gl}_n\mathbb{R} \to L^{-1} \in \mathrm{Gl}_n\mathbb{R}$. □

Exemple 8.6 L'application exponentielle matricielle $\exp : \mathrm{M}_n\mathbb{R} \to \mathrm{M}_n\mathbb{R}$ est un difféomorphisme d'un voisinage de 0 sur un voisinage de l'identité. En effet, cette application est de classe C^1 et sa différentielle en l'origine, $D_0\exp = \mathrm{Id}$, est inversible.

B Difféomorphisme local

Définition 8.7 *Une application $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ de classe C^1 est un C^1 difféomorphisme local si, pour tout point $x \in \Omega$, il existe un voisinage ouvert $U_x \subset \Omega$ de x et un voisinage ouvert $V_x \subset \mathbb{R}^n$ de $f(x)$ tels que la restriction $f : U_x \to V_x$ soit un C^1 difféomorphisme.*

Proposition 8.8 *Une application $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ de classe C^1 est un C^1 difféomorphisme local si et seulement si sa différentielle $D_xf \in L(\mathbb{R}^n, \mathbb{R}^n)$ en chaque point $x \in U$ est une application linéaire inversible.*

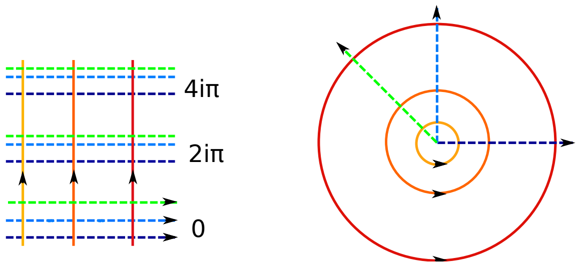
Preuve Suit immédiatement du théorème d'inversion locale. □

Corollaire 8.9 *Soit $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ un difféomorphisme local. L'application f est ouverte : l'image $f(U)$ de tout ouvert $U \subset \Omega$ est un ouvert de $\mathbb{R}^n$. En particulier, l'image $f(\Omega)$ est ouverte dans $\mathbb{R}^n$.*

Preuve Soit $U \subset \Omega$ un ouvert. On applique 8.7 et 8.8 à $f : U \to \mathbb{R}^n$. On obtient donc, pour tout $x \in U$, un ouvert $U_x \subset U$ contenant x dont l'image $V_x = f(U_x) \subset f(U)$ est ouverte dans $\mathbb{R}^n$. On a $f(U) = \cup_{x \in U} f(U_x)$ ouvert comme union d'ouverts. □

Remarque 8.10 Un difféomorphisme local $f : \Omega \to \mathbb{R}^n$ n'est pas toujours injectif. Considérer par exemple l'application exponentielle complexe

$$(x, y) \in \mathbb{R}^2 \simeq \mathbb{C} \longrightarrow (e^x \cos y, e^x \sin y) \in \mathbb{R}^2 \simeq \mathbb{C}\,.$$



L'application exponentielle complexe

Par contre, pour un difféomorphisme local injectif, on peut faire la remarque suivante.

Proposition 8.11 *Soit f un difféomorphisme local $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ injectif. Alors, f est un difféomorphisme $f : \Omega \to f(\Omega)$ sur son image.*

Preuve L'application $f : \Omega \to f(\Omega)$ est une bijection sur son image $f(\Omega)$. Par le théorème d'inversion locale, $f(\Omega)$ est ouverte dans $\mathbb{R}^n$ et l'application réciproque f^{-1} est localement de classe C^1, donc C^1. □

On retrouve le cas de la dimension 1.

Rappel 8.12 *Soit $f :]a, b[\to \mathbb{R}$ une application de classe C^1. On suppose que la dérivée de f ne s'annule pas sur $]a, b[$.*
Alors f est une bijection sur son image, qui est un intervalle ouvert $J \subset \mathbb{R}$, et l'application réciproque $f^{-1} : J \to]a, b[$ est également de classe C^1. En d'autres termes, $f :]a, b[\to J$ est un C^1 difféomorphisme.

L'injectivité de f est ici conséquence de ce que f' ne s'annule pas et donc que f est strictement monotone.

C * Théorème de Hadamard *

Un petit complément, le théorème de Hadamard. On commence par introduire la notion d'application propre. Cette partie peut être considérée comme un exercice.

Nous avons vu que l'image directe d'un compact par une application continue est compacte. Par contre, l'image inverse d'un compact par cette même application peut ne pas être compacte (considérer une application constante, ou bornée, par exemple).

Définition 8.13 Application propre
Soit $f : X \to Y$ une application entre deux espaces métriques. On dit que f est propre si f est continue, et si l'image réciproque $f^{-1}(K)$ de tout compact $K \subset Y$ de Y est un compact de X.

Propriété 8.14 *Une application continue $f : \mathbb{R}^n \to \mathbb{R}^p$ est propre si et seulement si $\|f(x)\| \to \infty$ lorsque $\|x\| \to \infty$.*

Preuve Supposons que $\|f(x)\| \to \infty$ lorsque $\|x\| \to \infty$. L'image réciproque par f d'une partie bornée de $\mathbb{R}^p$ est donc bornée dans $\mathbb{R}^n$. Puisque f est continue, l'image réciproque par f d'une partie fermée de $\mathbb{R}^p$ est fermée dans $\mathbb{R}^n$. La conclusion suit de la description des compacts de $\mathbb{R}^n$ et $\mathbb{R}^p$.

Si maintenant f est propre, l'image réciproque de toute boule fermée $B_f(x,r) \subset \mathbb{R}^p$ est un compact de $\mathbb{R}^n$ donc est bornée. □

Proposition 8.15 *Une application propre $f : X \to Y$ entre deux espaces métriques est fermée, c'est-à-dire dire que l'image directe par f de tout fermé de X est un fermé de Y.*

Preuve Soit $F \subset X$ un fermé. Soit $y_n = f(x_n)$ une suite de $f(X) \subset Y$, qui converge vers $\alpha \in Y$. On veut montrer que $\alpha \in f(X)$.

La réunion $K = \{y_n, n \in \mathbb{N}\} \cup \{\alpha\}$ est une partie compacte de Y (une suite convergente et sa limite, exemple 2.46). Son image réciproque $f^{-1}(K)$ est donc une partie compacte de X. La suite (x_n) prenant ses valeurs dans $f^{-1}(K)$ compact, elle admet une sous-suite convergente (x_{n_p}). Notons $x_\infty \in X$ la limite de cette sous-suite. Par continuité de f, on a

$$f(x_\infty) = \lim_{p\to\infty} f(x_{n_p}) = \lim_{p\to\infty} y_{n_p} = \alpha\,,$$

et le résultat. □

Le résultat suivant précise le théorème d'inversion locale, lorsque le difféomorphisme local est propre et défini sur tout $\mathbb{R}^n$.

Théorème 8.16 Théorème de Hadamard

Soit $f : \mathbb{R}^n \to \mathbb{R}^n$ un difféomorphisme local. On suppose que f est propre. Alors $f(\mathbb{R}^n) = \mathbb{R}^n$, et $f : \mathbb{R}^n \to \mathbb{R}^n$ est un difféomorphisme.

Le difféomorphisme local f est ici défini sur $\mathbb{R}^n$ tout entier. On demande de plus que $f : \mathbb{R}^n \to \mathbb{R}^n$ soit propre, c'est-à-dire que $\|f(x)\| \to \infty$ lorsque $\|x\| \to \infty$. Sous cette hypothèse, on peut montrer d'une part que $f(\mathbb{R}^n) = \mathbb{R}^n$, et d'autre part que f est injective. Le résultat suit alors du lemme **??**. Ce second point (injectivité) est un peu délicat ; une façon élégante de l'aborder est d'utiliser la notion de revêtement et d'espace simplement connexe. Nous nous contenterons donc de montrer le résultat suivant, en admettant l'injectivité.

Proposition 8.17 *Soit $f : \mathbb{R}^n \to \mathbb{R}^n$ un difféomorphisme local. On suppose que f est injective et propre. Alors $f(\mathbb{R}^n) = \mathbb{R}^n$, et $f : \mathbb{R}^n \to \mathbb{R}^n$ est un difféomorphisme.*

Preuve Il suffit donc de montrer que $f : \mathbb{R}^n \to \mathbb{R}^n$ est surjective. Puisque f est un difféomorphisme local, l'image $f(\mathbb{R}^n) \subset \mathbb{R}^n$ est ouverte, et bien sûr non vide. Le fait qu'elle soit fermée résulte de la proposition 8.15. Le résultat suit par connexité de $\mathbb{R}^n$. □

D Théorème des fonctions implicites

> Le théorème d'inversion locale va nous permettre dans certains cas d'étudier localement les "ensembles de niveau" d'une application $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^q$ de classe C^1.

Commençons par évoquer les applications scalaires ($q = 1$). Nous verrons plus loin (exercice 9.23) que tout fermé F de $\mathbb{R}^n$ peut s'écrire comme l'ensemble $F = \{m \in \mathbb{R}^n \mid f(m) = 0\}$ des zéros d'une fonction $f : \mathbb{R}^n \to \mathbb{R}$ de classe C^1 (et même C^∞). Si l'on veut pouvoir dire quelque chose de pertinent sur les ensembles de niveau de f, il faudra donc faire des hypothèses sur cette application.

En supposant que la différentielle de $f : \mathbb{R}^n \to \mathbb{R}$ ne s'annule pas au point $m_0 \in \mathbb{R}^n$, nous montrerons que le lieu F des zéros de f n'est, au voisinage de m_0, "pas n'importe quoi" : c'est un graphe au-dessus d'un hyperplan affine (voir le chapitre 10 pour aller plus loin dans cette approche géométrique).

Soit, plus généralement, une application $f : \Omega \subset \mathbb{R}^p \times \mathbb{R}^q \to \mathbb{R}^q$ de classe C^1. Soit $m_0 \in \Omega$. On désignera respectivement par

$$D^{(1)}_{m_0} f = D_{m_0} f_{|\mathbb{R}^p \times 0} \in L(\mathbb{R}^p, \mathbb{R}^q) \quad \text{et} \quad D^{(2)}_{m_0} f = D_{m_0} f_{|0 \times \mathbb{R}^q} \in L(\mathbb{R}^q, \mathbb{R}^q)$$

les restrictions de la différentielle $D_{m_0} f$ au premier (resp. second) facteur du produit $\mathbb{R}^p \times \mathbb{R}^q$.

Théorème 8.18 Fonctions implicites *Soit $f : \Omega \subset \mathbb{R}^p \times \mathbb{R}^q \to \mathbb{R}^q$ une application de classe C^1. Soient $m_0 \in \Omega$ et $c = f(m_0)$.*

On suppose que, au point $m_0 \in \Omega$, la "différentielle partielle" de f dans la direction du second paquet de coordonnées est une application linéaire inversible, soit

$$D^{(2)}_{m_0} f = D_{m_0} f_{|0 \times \mathbb{R}^q} \in GL(\mathbb{R}^q, \mathbb{R}^q)\,.$$

Il existe alors un voisinage ouvert $U \times V \subset \Omega \subset \mathbb{R}^p \times \mathbb{R}^q$ de m_0 et une application $h : U \to V$ de classe C^1 tels que

$$(x, y) \in U \times V \textit{ vérifie } f(x, y) = c \quad \text{si et seulement si} \quad x \in U \textit{ et } y = h(x).$$

Notons $m_0 = (x_0, y_0)$. On sait de plus calculer la différentielle de h en x_0 :

$$D_{x_0} h = -(D^{(2)}_{m_0} f)^{-1} \cdot D^{(1)}_{m_0} f\,.$$

Ceci nous dit que l'ensemble de niveau $\{(x,y) \mid f(x,y) = f(m_0)\}$ est, au voisinage du point m_0, le graphe $\{(x,y) \in \mathbb{R}^p \times \mathbb{R}^q \mid y = h(x)\}$ d'une application h de classe C^1. La fonction h n'est pas connue explicitement, mais est définie de façon “implicite” par la condition $f(x, h(x)) = c$.

Pour comprendre cet énoncé, on se réfère au cas modèle, celui d'une application affine

$$\varphi : (x,y) \in \mathbb{R}^p \times \mathbb{R}^q \longrightarrow Ax + By + \alpha \in \mathbb{R}^q$$

avec $A \in L(\mathbb{R}^p, \mathbb{R}^q)$, $B \in L(\mathbb{R}^q, \mathbb{R}^q)$ et $\alpha \in \mathbb{R}^q$. La différentielle partielle de φ dans la direction du second facteur est, en tout point $(x,y) \in \mathbb{R}^p \times \mathbb{R}^q$,

$$D^{(2)}_{(x,y)}\varphi = B\,.$$

Supposons $B \in L(\mathbb{R}^q, \mathbb{R}^q)$ inversible. Soit $c \in \mathbb{R}^q$ quelconque. L'ensemble de niveau $\mathcal{E}_c = \{(x,y) \mid \varphi(x,y) = c\}$ coïncide avec le graphe de l'application $h : x \in \mathbb{R}^p \to B^{-1}\,(c-\alpha-Ax) \in \mathbb{R}^q$. Le théorème des fonctions implicites est donc vrai “globalement” dans ce cas modèle. Cela s'applique en particulier à l'application affine tangente à f au point a. Conformément à notre principe général il restera vrai, localement, pour f elle-même.

Remarque 8.19 Lorsqu'on cherchera à appliquer le théorème des fonctions implicites à une application $f : \mathbb{R}^n \to \mathbb{R}^q$ au voisinage du point m_0, l'espace de départ $\mathbb{R}^n$ ne sera en général pas donné “clés en mains” sous la forme d'un produit $\mathbb{R}^n = \mathbb{R}^p \times \mathbb{R}^q$ avec $D^{(2)}_{m_0}f$ inversible.

L'hypothèse naturelle sera que la différentielle de $f : \mathbb{R}^n \to \mathbb{R}^q$ au point m_0 est une application linéaire $D_{m_0}f \in L(\mathbb{R}^n, \mathbb{R}^q)$ surjective. Sous cette hypothèse, un raisonnement d'algèbre linéaire élémentaire (voir l'exercice 8.22) montre alors que, *après permutation des coordonnées*, l'application $f : \mathbb{R}^n = \mathbb{R}^p \times \mathbb{R}^q \to \mathbb{R}^q$ vérifie bien l'hypothèse $D^{(2)}_{m_0}f$ inversible. Voir l'exemple ci-dessous et l'exercice 8.21.

Exemple 8.20 **Cas d'école** Prenons l'exemple de l'application

$$f : (x,y) \in \mathbb{R}^2 \to x^2 + y^2 - 1 \in \mathbb{R}$$

dont l'ensemble des zéros est le cercle $\mathbb{S}^1$. C'est une application de classe C^1 dont la différentielle en chaque point m_0 distinct de l'origine (en particulier, en chaque point du cercle) est non nulle, donc surjective. Le théorème des fonctions implicites s'applique en tout point $m_0 = (x_0, y_0) \in \mathbb{S}^1$ pour montrer que

— si $y_0 \neq 0$, le cercle est au voisinage de m_0 un graphe y fonction de x

— si $x_0 \neq 0$, le cercle est au voisinage de m_0 un graphe x fonction de y.

Bien entendu, dans ce cas extrêmement simple, on n'a pas besoin du théorème des fonctions implicites pour s'apercevoir que $\mathbb{S}^1$, selon l'endroit où on regarde, est localement le graphe d'au moins une des applications

$$x \to y = \pm\sqrt{1-x^2} \quad \text{ou} \quad y \to x = \pm\sqrt{1-y^2}.$$

Nous reverrons cet exemple dans le paragraphe 10.C.

Exercice 8.21 Soit $n \geq 2$. Etudier de même la sphère $\mathbb{S}^{n-1}$, ensemble des zéros de l'application

$$f : (x_1, \cdots, x_n) \in \mathbb{R}^n \to \sum_{k=1}^{n} x_k^2 - 1 \in \mathbb{R}\,.$$

Exercice 8.22

1. Soit $M \in \mathrm{M}_{q,n}\mathbb{R}$ une matrice. Montrer que le rang de M est égal au rang du plus grand mineur inversible extrait de M. On pourra utiliser le théorème de la base incomplète.
2. Soit $L : \mathbb{R}^n \to \mathbb{R}^q$ une application linéaire surjective. Montrer qu'il existe $1 \leq i_1 < i_2 < \cdots < i_q \leq n$ tels que les vecteurs $(L(e_{i_1}), L(e_{i_2}), \cdots, L(e_{i_q}))$ soient linéairement indépendants.

Passons à la preuve. Nous allons déduire très facilement le théorème des fonctions implicites du théorème d'inversion locale.

Preuve du théorème • On cherche à construire un difféomorphisme F défini sur un voisinage de m_0 dans $\mathbb{R}^p \times \mathbb{R}^q$ et qui envoie les ensembles de niveau de f sur des sous-espaces affines de $\mathbb{R}^p \times \mathbb{R}^q$. Introduisons l'application

$$F : (x, y) \in \Omega \subset \mathbb{R}^p \times \mathbb{R}^q \longrightarrow (x, f(x, y)) \in \mathbb{R}^p \times \mathbb{R}^q\,.$$

L'application F est de classe C^1 comme f. Il est très facile de voir que F est un C^1-difféomorphisme local au voisinage du point m_0. Pour cela il suffit en effet, grâce au théorème d'inversion locale, de vérifier que sa différentielle au point m_0 est inversible. On vérifie que c'est bien le cas en examinant sa matrice jacobienne, qui est triangulaire par blocs avec deux blocs diagonaux inversibles :

$$J_{m_0}F = \begin{pmatrix} \mathrm{Id}_p & 0 \\ J_{m_0}^{(1)}f & J_{m_0}^{(2)}f \end{pmatrix} \in \mathrm{M}_{p+q}\mathbb{R}\,,$$

$J_{m_0}^{(1)}f$ et $J_{m_0}^{(2)}f$ désignant respectivement les matrices jacobiennes associées à $D_{m_0}^{(1)}f$ et à $D_{m_0}^{(2)}f$.

On obtient donc un voisinage ouvert de m_0, que l'on prendra sous forme de produit $U_1 \times V \subset \Omega$, et tel que la restriction

$$F : U_1 \times V \longrightarrow F(U_1 \times V)$$

soit un C^1 difféomorphisme sur son image $F(U_1 \times V)$. L'application réciproque est de classe C^1 et s'écrit donc

$$(x, H(x,z)) \in U_1 \times V \longleftarrow (x,z) \in F(U_1 \times V) : F^{-1}$$

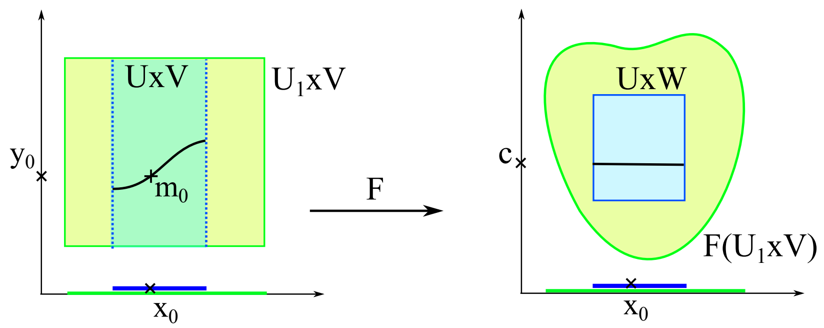
avec H de classe C^1.

• On construit maintenant h. L'image $F(U_1 \times V)$ est un voisinage ouvert de $F(x_0, y_0) = (x_0, c)$ donc contient un produit $U \times W$, où $U \subset U_1$ et W sont respectivement des voisinages ouverts de x_0 et c. On pose

$$h : x \in U \longrightarrow H(x,c) \in V\,.$$

Soit $(x,y) \in U \times V$ tel que $f(x,y) = c$. On a alors $F(x,y) = (x,c) \in U \times W$, donc $(x,y) = F^{-1}(x,c) = (x, h(x))$ et $y = h(x)$.

• Passons à la réciproque et supposons maintenant que $x \in U$ et $y = h(x) \in V$. Il suit alors que $(x,y) = (x, H(x,c)) = F^{-1}(x,c)$, donc que $c = f(x,y)$.



• Reste à déterminer la différentielle de h en x_0. On a, pour tout $x \in U$, $f(x, h(x)) \equiv c$. En différentiant, on obtient pour tous $x \in U$ et $u \in \mathbb{R}^p$:

$$D^{(1)}_{(x,h(x))} f\ (u) + D^{(2)}_{(x,h(x))} f \cdot D_x h\ (u) = 0\,,$$

comme annoncé. □

9. Différentielles d'ordre supérieur

Une application différentiable f s'approche, localement, par des applications affines. Pour obtenir des approximations plus fines, on cherchera plus généralement à approcher f par des applications polynomiales (formules de Taylor). Ceci repose sur la notion de différentielle d'ordre supérieur.

On va d'abord introduire très tranquillement la différentielle d'ordre 2. On se permettra d'aller plus rapidement pour les différentielles d'ordre supérieur, en laissant au lecteur le soin de compléter les quelques détails manquants.

A Différentielle seconde

A.1 Différentier deux fois

Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$ une fonction différentiable définie sur un ouvert U de $\mathbb{R}^n$. On dispose donc d'une application "différentielle de f", soit

$$Df : U \subset \mathbb{R}^n \to L(\mathbb{R}^n, \mathbb{R}^p) \simeq \mathbb{R}^{np} .$$

Supposons que l'application $Df : U \subset \mathbb{R}^n \to L(\mathbb{R}^n, \mathbb{R}^p)$ soit elle-même différentiable en $x \in U$. Sa différentielle au point x, soit

$$D_x(Df) \in L(\mathbb{R}^n, L(\mathbb{R}^n, \mathbb{R}^p)) ,$$

est une application linéaire de $\mathbb{R}^n$ vers $L(\mathbb{R}^n, \mathbb{R}^p)$.

Définition 9.1 *Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$. On suppose que f est différentiable au voisinage de $x \in U$. Lorsque Df est elle-même différentiable en x, on dit que f est deux fois différentiable au point x.*

En pratique, on pourra se servir du lemme suivant pour montrer qu'une application est deux fois différentiable.

Lemme 9.2 *On note $(e_1, \cdots, e_n)$ la base canonique de $\mathbb{R}^n$ et $x = (x_1, \cdots, x_n)$ les coordonnées d'un point dans cette base.*

Soient $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$ une application différentiable et $x_0 \in U$. Les assertions suivantes sont équivalentes :

1. *l'application f est deux fois différentiable au point $x_0 \in U$*
2. *chaque application $x \in U \to D_x f(e_i) = \frac{\partial f}{\partial x_i} \in \mathbb{R}^p$ est différentiable en x_0 $(1 \leq i \leq n)$*
3. *pour tout vecteur $h \in \mathbb{R}^n$, l'application $x \in U \to D_x f(h) \in \mathbb{R}^p$ est différentiable en x_0.*

Preuve L'équivalence entre les conditions 1. et 2. suit (comme dans le lemme 7.9) de ce que l'application

$$\mathcal{J} : L \in L(\mathbb{R}^n, \mathbb{R}^p) \to (L(e_1), \cdots, L(e_n)) \in (\mathbb{R}^p)^n$$

est un isomorphisme d'espaces vectoriels et donc un C^1 difféomorphisme.

La condition 3. implique trivialement la condition 2. La réciproque suit de la linéarité de chaque différentielle $D_x f$. □

Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$. Lorsqu'elle est définie, la différentielle $D_x(Df)$ appartient par construction à l'espace $L(\mathbb{R}^n, L(\mathbb{R}^n, \mathbb{R}^p)$. Ce n'est pas très joli, et promet d'empirer lorsqu'on introduira les différentielles d'ordre supérieur. Heureusement, on a le lemme suivant.

Lemme 9.3 *L'espace vectoriel $L(\mathbb{R}^n, L(\mathbb{R}^n, \mathbb{R}^p))$ des applications linéaires sur $\mathbb{R}^n$ à valeurs dans $L(\mathbb{R}^n, \mathbb{R}^p)$ s'identifie naturellement à l'espace vectoriel $L_2(\mathbb{R}^n, \mathbb{R}^p)$ des applications bilinéaires de $\mathbb{R}^n \times \mathbb{R}^n$ dans $\mathbb{R}^p$.*

Preuve Soit $\ell \in L(\mathbb{R}^n, L(\mathbb{R}^n, \mathbb{R}^p))$ une application linéaire. On lui associe l'application bilinéaire $\Phi(\ell) \in L_2(\mathbb{R}^n, \mathbb{R}^p)$ définie par

$$\Phi(\ell) : (u, v) \in \mathbb{R}^n \times \mathbb{R}^n \to \big(\ell(u)\big)(v) \in \mathbb{R}^p \,.$$

L'application linéaire $\Phi : L(\mathbb{R}^n, L(\mathbb{R}^n, \mathbb{R}^p)) \to L_2(\mathbb{R}^n, \mathbb{R}^p)$, qui est injective entre deux espaces vectoriels de même dimension n^2p, est un isomorphisme d'espaces vectoriels.

Son inverse $\Phi^{-1} : L_2(\mathbb{R}^n, \mathbb{R}^p) \to L(\mathbb{R}^n, L(\mathbb{R}^n, \mathbb{R}^p))$ est défini, pour chaque $B \in L_2(\mathbb{R}^n, \mathbb{R}^p)$, par $\Phi^{-1}(B)(u) = B(u, \cdot) \in L(\mathbb{R}^n, \mathbb{R}^p)$ pour tout $u \in \mathbb{R}^n$. □

Définition 9.4 *La différentielle $D_x(Df) \in L(\mathbb{R}^n, L(\mathbb{R}^n, \mathbb{R}^p))$ de Df au point x, lorsqu'elle existe, s'identifie donc naturellement une à application bilinéaire de $\mathbb{R}^n \times \mathbb{R}^n$ dans $\mathbb{R}^p$ que nous noterons $D_x^2 f \in L_2(\mathbb{R}^n, \mathbb{R}^p)$: c'est la différentielle seconde de f en x.*

Attention à ne pas confondre la différentielle seconde $D_x^2 f$ au point x avec la notation, utilisée au chapitre précédent, où f était définie sur un espace produit et $D^{(2)}f$ désignait la différentielle (d'ordre 1) dans la direction du second facteur du produit.

Nous allons voir que la différentielle seconde est une application bilinéaire symétrique. Voir également le corollaire 9.14 pour les différentielles d'ordre supérieur.

Théorème 9.5 Symétrie de Schwartz *Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$, que l'on suppose différentiable sur U et deux fois différentiable au point $x \in U$. Alors la différentielle seconde $D_x^2 f \in L_2(\mathbb{R}^n, \mathbb{R}^p)$ est bilinéaire symétrique, i.e. on a pour tous vecteurs $u, v \in \mathbb{R}^n$*

$$D_x^2 f(u, v) = D_x^2 f(v, u)\,.$$

Soient $u, v \in \mathbb{R}^n$. La preuve consistera à comparer l'expression

$$\Delta = f(x + u + v) + f(x) - f(x + u) - f(x + v)$$

avec $D_x^2 f(u, v)$ et $D_x^2 f(v, u)$ pour constater que la différence

$$B : (u, v) \longrightarrow D_x^2 f(u, v) - D_x^2 f(v, u)$$

est un $o(\|u\| + \|v\|)^2$. On concluera alors avec le lemme suivant, à rapprocher du lemme 6.6.

Lemme 9.6 *On munit $\mathbb{R}^n$ et $\mathbb{R}^p$ de normes. Soit $B : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^p$ une application bilinéaire. On suppose que*

$$B(u, v) = o(\|u\| + \|v\|)^2\,,$$

c'est-à-dire que pour tout $\varepsilon > 0$ on a

$$\|B(u, v)\| \leq \varepsilon\,(\|u\| + \|v\|)^2$$

dès que $\|u\| + \|v\|$ est assez petit. Alors la forme bilinéaire B est nulle.

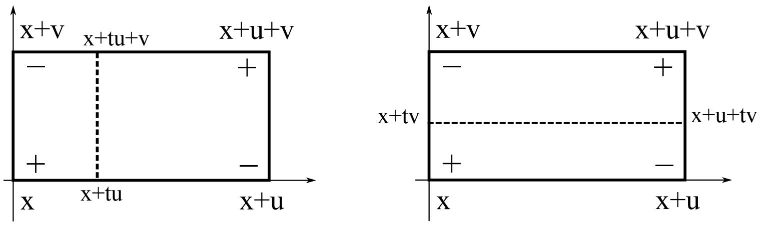
Preuve Immédiat par homogénéité de B. On fixe en effet $u, v \in \mathbb{R}^n$ et on obtient pour tout $t > 0$ assez petit

$$t^2\|B(u, v)\| = \|B(tu, tv)\| \leq \varepsilon\, t^2(\|u\| + \|v\|)^2\,,$$

autrement dit $\|B(u, v)\|$ est arbitrairement petit, donc nul. □

Preuve du théorème 9.5 Soient $u, v \in \mathbb{R}^n$. Nous allons utiliser le théorème des accroissements finis pour évaluer de deux façons différentes la quantité

$$\Delta = f(x + u + v) + f(x) - f(x + u) - f(x + v)\,.$$



• Pour la première estimation, on observe que $\Delta = g(1) - g(0)$, où

$$g(t) = f(x + tu + v) - f(x + tu)\,.$$

L'application g est dérivable, avec $g'(t) = D_{x+tu+v}f(u) - D_{x+tu}f(u)$.

Pour estimer chacune des dérivées $g'(t)$, on va utiliser le fait que la différentielle Df est différentiable au point x. Soit $\varepsilon > 0$. Il existe donc $r_\varepsilon > 0$ tel que, pour tout vecteur $h \in \mathbb{R}^n$ de norme au plus r_ε, on ait $\|D_{x+h}f - D_x f - D^2_x f(h,\cdot)\| \leq \varepsilon\|h\|$. Supposons maintenant que les vecteurs u et v satisfassent $\|u\| + \|v\| \leq r_\varepsilon$. On a alors pour tout $t \in [0, 1]$

$$\begin{aligned}
\|D_{x+tu+v}f - D_x f - D^2_x f(tu + v, \cdot)\| &\leq \varepsilon\|tu + v\| \leq \varepsilon(\|u\| + \|v\|)\\
\|D_{x+tu}f - D_x f - D^2_x f(tu, \cdot)\| &\leq \ \varepsilon\|tu\| \ \leq \varepsilon\|u\|\,.
\end{aligned}$$

Il suit alors, en utilisant l'inégalité triangulaire, que

$$\begin{aligned}
\|g'(t) - D^2_x f(v, u)\| &= \|g'(t) - D^2_x f(tu + v, u) + D^2_x f(tu, u)\|\\
&\leq 2\varepsilon(\|u\| + \|v\|)\ \|u\|\,.
\end{aligned}$$

Le théorème 6.13 appliqué à $\varphi : t \in [0, 1] \to g(t) - tD^2_x f(v, u) \in \mathbb{R}^p$ donne

$$\|\Delta - D^2_x f(v, u)\| \leq 2\varepsilon(\|u\| + \|v\|)\ \|u\|\,.$$

• Pour la seconde estimation, on procède de même en remarquant cette fois que $\Delta = h(1) - h(0)$, où

$$h(t) = f(x + u + tv) - f(x + tv)\,.$$

On obtient alors, toujours lorsque $\|u\| + \|v\| \leq r_\varepsilon$, l'inégalité

$$\|\Delta - D^2_x f(u, v)\| \leq 2\varepsilon(\|u\| + \|v\|)\ \|v\|\,.$$

• Il suit de ces deux estimations que, pour $\|u\| + \|v\| \leq r_\varepsilon$, on a

$$\|D^2_x f(v, u) - D^2_x f(u, v)\| \leq 2\varepsilon(\|u\| + \|v\|)^2\,.$$

La conclusion suit alors, comme annoncé, du lemme 9.6. $\square$

Exercice 9.7 Exemples fondamentaux

1. Soit $f : \mathbb{R} \to \mathbb{R}^n$. Montrer que f est deux fois dérivable au point $t_0 \in \mathbb{R}$ si et seulement si elle est deux fois différentiable en ce point. Montrer alors l'égalité $f''(t_0) = D^2_{t_0} f(1,1)$.
2. Déterminer la différentielle seconde d'une application linéaire $\ell : \mathbb{R}^n \to \mathbb{R}^p$.
3. Déterminer la différentielle seconde de $B : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$ bilinéaire.

Exercice 9.8

1. Soit $f : \mathbb{R}^n \to \mathbb{R}^p$ une application deux fois différentiable. Soit $v \in \mathbb{R}^n$. Montrer que l'application $x \in \mathbb{R}^n \to D_x f(v) \in \mathbb{R}^p$ est différentiable et déterminer sa différentielle.
2. Soient $k \geq 1$ et $\mathrm{e}_k : A \in \mathrm{M}_n\mathbb{R} \to A^k \in \mathrm{M}_n\mathbb{R}$. Montrer que cette application est deux fois différentiable, et déterminer sa différentielle. On pourra se limiter à $k = 2$ et $k = 3$.

A.2 Dérivées partielles d'ordre 2

Tout comme à l'ordre 1, il sera commode d'introduire les dérivées partielles d'ordre 2. Nous reprenons les notations du chapitre 7 et noterons $(e_1, \cdots, e_n)$ la base canonique de $\mathbb{R}^n$, et $(x_1, \cdots, x_n)$ les coordonnées correspondantes. La fonction f étant supposée différentiable sur U, nous disposons des dérivées partielles

$$\frac{\partial f}{\partial x_i} : x \in U \to D_x f(e_i) \in \mathbb{R}^p\,.$$

Lemme 9.9 *Supposons $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$ deux fois différentiable au point $x \in U$. Pour tous $1 \leq i, j \leq n$, on a*

$$\frac{\partial}{\partial x_i}\Big(\frac{\partial f}{\partial x_j}\Big)(x) = D^2_x(f)(e_i, e_j)\,.$$

Preuve Immédiat par différentiation de fonctions composées. On a en effet pour tout $y \in U$

$$\frac{\partial f}{\partial x_j}(y) = D_y f(e_j) = B((Df)(y), e_j)\,,$$

$B : (\ell, u) \in L(\mathbb{R}^n, \mathbb{R}^p) \times \mathbb{R}^n \to \ell(u) \in \mathbb{R}^p$ désignant ici l'application, bilinéaire, d'évaluation de l'application linéaire ℓ sur le vecteur u. □

On note, lorsque c'est défini,

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i}\Big(\frac{\partial f}{\partial x_j}\Big)\,.$$

La symétrie de Schwartz équivaut à dire que, lorsque f est deux fois différentiable au point x, on a pour tous $1 \leq i, j \leq n$,

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x) . \tag{9.1}$$

Remarque 9.10 Supposons la fonction f différentiable sur U, et que chacune des dérivées partielles $\frac{\partial f}{\partial x_j}$ admette elle-même des dérivées partielles au point x. Ceci ne permet pas d'assurer que f soit elle-même deux fois différentiable en x. Dans ce cas, la symétrie (9.1) peut être mise en défaut. Cela dit, il ne faut ne pas trop s'attarder sur ces pathologies, même s'il faut être conscient qu'elles existent.

A.3 Fonctions de classe C^2

Définition 9.11 *Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$. On dit que f est de classe C^2 lorsqu'elle est deux fois différentiable en chaque point de U, et lorsque sa différentielle seconde*

$$D^2 f : x \in U \to D_x^2 f \in L_2(\mathbb{R}^n, \mathbb{R}^p)$$

est une application continue.

On déduit du lemme 7.9, du théorème 7.14 et du corollaire 7.15 des énoncés analogues pour la différentielle seconde. En particulier :

Lemme 9.12 *Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$ deux fois différentiable. Alors f est de classe C^2 si et seulement si toutes ses dérivées partielles*

$$\frac{\partial^2 f}{\partial x_i \partial x_j} : U \to \mathbb{R}^p$$

($1 \leq i, j \leq n$) sont continues.

Preuve Conséquence immédiate de ce que l'application

$$B \in L_2(\mathbb{R}^n, \mathbb{R}^p) \to (B(e_i, e_j))_{1 \leq i,j \leq n} \in (\mathbb{R}^p)^{n^2}$$

(évaluation de l'application bilinéaire B sur les vecteurs de base $(e_1, \cdots, e_n)$) est un isomorphisme d'espaces vectoriels et donc un homéomorphisme. □

Théorème 9.13 *Si les dérivées partielles d'ordre 2 existent et sont continues sur U, la fonction f est de classe C^2 sur U.*

Preuve Supposons que les dérivées partielles d'ordre 2 existent et sont continues sur U. Soit $1 \leq i \leq n$. Le corollaire 7.15, appliqué à la dérivée partielle $x \to \frac{\partial f}{\partial x_i}(x) = D_x f(e_i)$, assure que celle-ci est une fonction de classe C^1 sur U. L'application

$$L \in L(\mathbb{R}^n, \mathbb{R}^p) \to (L(e_1), \cdots, L(e_n)) \in (\mathbb{R}^p)^n$$

est un isomorphisme d'espace vectoriels, et donc un C^1 difféomorphisme. Il suit que la différentielle $Df : U \to L(\mathbb{R}^n, \mathbb{R}^p)$ est elle-même de classe C^1, ce qu'on voulait. □

B Différentielle d'ordre k

Soit $k \geq 1$. On désigne par $L_k(\mathbb{R}^n, \mathbb{R}^p)$ l'espace vectoriel des applications k-linéaires sur le produit $\mathbb{R}^n \times \cdots \times \mathbb{R}^n = (\mathbb{R}^n)^k$, et à valeurs dans $\mathbb{R}^p$. Le lemme 9.3 se généralise à tout entier $k \geq 2$ et fournit un isomorphisme naturel

$$L_k(\mathbb{R}^n, \mathbb{R}^p) \simeq L(\mathbb{R}^n, L_{k-1}(\mathbb{R}^n, \mathbb{R}^p)) .$$

Soit maintenant $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$ une application définie sur un ouvert U de $\mathbb{R}^n$. On définit récursivement, lorsqu'elles existent, les différentielles d'ordre k de f. Supposons que f soit $k-1$ fois différentiable sur U, et k fois différentiable en $x \in U$. Cela signifie que l'application

$$D^{k-1} f : x \in U \longrightarrow D_x^{k-1} f \in L_{k-1}(\mathbb{R}^n, \mathbb{R}^p)$$

est différentiable au point x. Sa différentielle est alors la différentielle d'ordre k de f en x, soit

$$D_x^k f \in L(\mathbb{R}^n, L_{k-1}(\mathbb{R}^n, \mathbb{R}^p)) \simeq L_k(\mathbb{R}^n, \mathbb{R}^p) .$$

La symétrie de Schwartz s'étend à la différentielle d'ordre k.

Corollaire 9.14 Symétrie de Schwartz *Soient $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$ et $k \geq 2$. On suppose que f est $k-1$ fois différentiable sur U et admet une différentielle k-ième en $x \in U$. Alors $D_x^k f \in L_k(\mathbb{R}^n, \mathbb{R}^p)$ est une application k-linéaire symétrique.*

De façon équivalente, on a pour tout k-uplet $(i_1, \cdots, i_k)$ d'indices avec $1 \leq i_1, \cdots, i_k \leq n$ et toute permutation $\sigma \in \mathfrak{S}_k$, l'égalité

$$\frac{\partial^k f}{\partial x_{i_1} \cdots \partial x_{i_k}}(x) = \frac{\partial^k f}{\partial x_{\sigma(i_1)} \cdots \partial x_{\sigma(i_k)}}(x) .$$

Preuve Cet énoncé suit immédiatement de l'énoncé lorsque $k = 2$ (lemme 9.5) et de ce que toute permutation de $\{1, \cdots, k\}$ s'obtient comme produit des transpositions $\tau_{i,i+1}$ pour $1 \leq i \leq k-1$. □.

Définition 9.15 *Soit $k \geq 1$. On dit que $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$ est de classe C^k lorsqu'elle est k fois différentiable sur U, et lorsque sa différentielle d'ordre k, soit*

$$D^k f : x \in U \to D^k_x f \in L_k(\mathbb{R}^n, \mathbb{R}^p)\,,$$

est continue.
On dit que f est de classe C^∞ lorsqu'elle est de classe C^k pour tout entier $k \geq 1$.

Dans la pratique la proposition suivante, qui est une généralisation facile du théorème 7.14, pourra fournir un critère commode pour montrer qu'une application est de classe C^k.

Proposition 9.16 *L'application $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$ est de classe C^k si et seulement elle admet sur U des dérivées partielles d'ordre k continues.*

Revenons maintenant sur nos exemples favoris.

Lemme 9.17 *Une application linéaire, ou multilinéaire, est de classe C^∞.*

Preuve La preuve est laissée au lecteur. On lui demande également de déterminer les différentielles de tous ordres de ces applications. □

Passons aux propriétés de stabilité des applications k fois différentiables.

Proposition 9.18 *Soient $f : U \subset \mathbb{R}^n \to V \subset \mathbb{R}^p$ et $g : V \subset \mathbb{R}^p \to \mathbb{R}^q$ deux applications $k-1$ fois différentiables. On suppose que f est k fois différentiable en x et g est k fois différentiable en $f(x)$. Alors $g \circ f$ est k fois différentiable en x.*

Si f et g sont toutes deux de classe C^k, alors $g \circ f$ est également de classe C^k.

Preuve La démonstration se fait par récurrence sur k. Pour $k = 1$, ce sont les propositions 6.10 et 7.12. Supposons donc $k \geq 2$, et que la propriété a été démontrée à l'ordre $k-1$. On a vu en (6.2) que, pour tout $x \in U$, on a

$$D_x(g \circ f) = D_{f(x)}g \cdot D_x f\,.$$

Autrement dit on obtient $D(g \circ f)$ comme composée des applications

$$x \in U \to \big((Dg) \circ f, Df\big) \in L(\mathbb{R}^p, \mathbb{R}^q) \times L(\mathbb{R}^n, \mathbb{R}^p)$$

et de l'application bilinéaire (donc C^∞)

$$(A, B) \in L(\mathbb{R}^p, \mathbb{R}^q) \times L(\mathbb{R}^n, \mathbb{R}^p) \longrightarrow A \cdot B \in L(\mathbb{R}^n, \mathbb{R}^q)$$

(composée des applications linéaires A et B). Le résultat suit, puisque $g \circ f$ est k fois différentiable en x (resp. C^k) si et seulement si $D(g \circ f)$ est $(k-1)$ fois différentiable en x (resp. C^{k-1}). $\square$

Enfin, que se passe-t-il dans le théorème d'inversion locale ou des fonctions implicites, lorsque la donnée est de classe C^k (ou bien C^∞) ?

Proposition 9.19 Inversion locale C^k

Soit $f : U \subset \mathbb{R}^n \to V \subset \mathbb{R}^n$ un C^1 difféomorphisme. On suppose que f est de classe C^k pour un entier $k \geq 2$. Alors la réciproque $g := f^{-1}$ est également de classe C^k. On dit alors que f est un C^k difféomorphisme.

Preuve On a vu dans la preuve du théorème d'inversion locale (th. 8.4) que

$$Dg = \mathcal{I} \circ Df \circ g \tag{9.2}$$

où l'application $\mathcal{I} : L \in \mathrm{Gl}_n\mathbb{R} \to L^{-1} \in \mathrm{Gl}_n\mathbb{R}$, dont les composantes sont des fonctions rationnelles sur $\mathrm{M}_n\mathbb{R}$, est de classe C^∞. Supposons avoir montré que g est de classe C^j pour un entier $1 \leq j \leq k-1$. Il suit alors de (9.2) que la fonction Dg, comme composée de fonctions de classe C^j, est également de classe C^j. Ainsi, g est de classe C^{j+1}. $\square$

Proposition 9.20 Fonctions implicites C^k

Soient $f : \Omega \subset \mathbb{R}^p \times \mathbb{R}^q \to \mathbb{R}^q$ de classe C^k avec $k \geq 2$ et $m_0 = (x_0, y_0) \in \Omega$. On suppose que la dérivée partielle $D^{(2)}_{m_0} f \in L(\mathbb{R}^q, \mathbb{R}^q)$ est linéaire inversible. La fonction h définie implicitement au voisinage de x_0 par la condition

$$f(x, h(x)) = f(m_0)$$

est de classe C^k.

Preuve Reprenons les notations de la démonstration du théorème des fonctions implicites 8.18. La fonction implicite h est définie par $h(x) = H(x, c)$ où $(x, z) \to (x, H(x, z))$ est l'application réciproque de F. Lorsque f est supposée de classe C^k, le difféomorphisme local F est un C^k difféomorphisme local (9.19) donc H, puis h, sont bien de classe C^k. $\square$

C Suites ou séries de fonctions de classe C^k

Le théorème 7.17 se généralise immédiatement aux suites ou aux séries de fonctions de classe C^k $(k \geq 1)$.

Théorème 9.21 *Soient $U \subset \mathbb{R}^n$ un ouvert, $f_i : U \to \mathbb{R}^p$ une suite de fonctions de classe C^k. On suppose que la suite (f_i) converge localement uniformément sur U, ainsi que chacune des suites des différentielles $(D^p f_i)$ pour $1 \leq p \leq k$. Alors la limite $f := \lim f_i$ de la suite est de classe C^k et chacune des différentielles d'ordre $1 \leq p \leq k$ s'obtient en intervertissant limite et différentielle :*

$$D^p f = \lim D^p f_i \,.$$

Exemple 9.22 L'exponentielle matricielle $\exp : \mathrm{M}_n\mathbb{R} \to \mathrm{M}_n\mathbb{R}$ est une application de classe C^∞.

Exercice 9.23 Théorème de Whitney

1. Construire une fontion $g : \mathbb{R} \to [0, \infty[$ de classe C^∞ telle que $g > 0$ sur l'intervalle $]-1, 1[$ et soit nulle en dehors de cet intervalle.
2. Soit $U \subset \mathbb{R}^n$ un ouvert. Montrer que U peut s'écrire comme une réunion dénombrable de boules ouvertes.
3. Soit $B(a, r) \subset \mathbb{R}^n$ une boule ouverte. Construire une fonction $h : \mathbb{R}^n \to \mathbb{R}^+$ de classe C^∞, telle que h soit strictement positive en tout point de $B(a, r)$ et nulle en dehors de $B(a, r)$.
4. En déduire que tout fermé $F \subset \mathbb{R}^n$ (aussi compliqué soit-il !) peut s'écrire comme l'ensemble des zéros d'une fonction $F : \mathbb{R}^n \to \mathbb{R}$ de classe C^∞.
   On pourra écrire le complémentaire $U = {}^cF$ comme réunion dénombrable de boules ouvertes, et construire F comme une série convergente (dans C^∞) de fonctions de classe C^∞.

D Formules de Taylor

Une fois introduites les différentielles d'ordre supérieur, on va chercher à approcher une fonction k fois différentiable en un point par une fonction polynomiale de degré k.

En guise d'introduction, on rappelle la formule de Taylor pour les fonctions polynomiales.

D.1 La formule de Taylor pour les fonctions polynomiales

On fixe un entier $k \geq 0$ et on note $\mathbb{R}_k[X]$ l'espace vectoriel de dimension $k+1$ des fonctions polynomiales de degré au plus k sur $\mathbb{R}$.

Soit $x_0 \in \mathbb{R}$. La famille des polynômes $e_j : x \to (x - x_0)^j$ $(0 \leq j \leq k)$ étant linéairement indépendante, elle constitue une base de $\mathbb{R}_k[X]$.

Soit maintenant un polynôme $P \in \mathbb{R}_k[X]$. Il existe donc une unique famille de coefficients $(a_0, \cdots, a_k) \in \mathbb{R}^{k+1}$ pour laquelle $P = \sum_{j=0}^{k} a_j e_j$. En dérivant cette identité à l'ordre $0 \leq j \leq k$, et en évaluant au point x_0, on identifie les coefficients et l'on obtient $a_j = P^{(j)}(x_0)/j\,!$ pour $0 \leq j \leq k$. Autrement dit, on a pour tout $x \in \mathbb{R}$ l'égalité

$$P(x) = \sum_{j=0}^{k} \frac{P^{(j)}(x_0)}{j\,!}\,(x - x_0)^j\,. \tag{9.3}$$

La même démonstration fournit une formule analogue pour les fonctions polynomiales $P \in \mathbb{R}[X_1, \cdots, \mathbb{R}_n]$ à plusieurs variables.

Dans ce qui suit, on cherche à généraliser cette identité à des fonctions quelconques (i.e. non polynomiales). La fonction ne sera pas toujours égale à son "polynôme de Taylor" comme en (9.3). La différence, ou reste, pourra prendre plusieurs formes : Taylor-Young, Lagrange ou reste intégral, selon la précision recherchée et ce que permet la régularité de la fonction.

D.2 Polynôme de Taylor

Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$, que l'on suppose k fois différentiable au point $x \in U$. Sa différentielle d'ordre k au point x est une application k-linéaire symétrique $D_x^k f \in L_k(\mathbb{R}^n, \mathbb{R}^p)$ sur le produit $\mathbb{R}^n \times \cdots \times \mathbb{R}^n = (\mathbb{R}^n)^k$.

Notation 9.24 *Pour tout vecteur $h \in \mathbb{R}^n$, on notera*

$$D_x^k f \cdot h^{(k)} := D_x^k f(h, \cdots, h)\,.$$

Il est important de savoir écrire l'expression ci-dessus en termes de dérivées partielles (en particulier en petite dimension et pour k petit). Allons-y !

Si $h = (h_1, \cdots, h_n)$, le fait que $D_x^k f$ soit k-linéaire assure que

$$D_x^k f \cdot h^{(k)} = \sum_{i_1, \cdots, i_k} \frac{\partial^k f(x)}{\partial x_{i_1} \cdots \partial x_{i_k}}\, h_{i_1} \cdots h_{i_k}\,,$$

la somme étant prise sur tous les k-uplets $(i_1, \cdots, i_k)$ d'entiers de 1 à n.

Nous tirons maintenant partie de ce que $D_x^k f$ est symétrique (corollaire 9.14) en regroupant dans cette somme tous les termes faisant intervenir le même nombre de dérivées partielles par rapport à x_1, à x_2,... et à x_n respectivement. Il vient alors (compter les termes semblables)

$$\frac{1}{k!}\, D_x^k f \cdot h^{(k)} = \sum_{|\alpha| = k} \frac{h^\alpha}{\alpha!}\, \frac{\partial^k f}{\partial x^\alpha}(x)\,,$$

la somme étant désormais prise sur les multi-indices $\alpha = (\alpha_1, \cdots, \alpha_n)$ de longueur $|\alpha| = \sum_1^n \alpha_i$ égale à k, avec les conventions

$$\frac{\partial^k}{\partial x^\alpha} = \Big(\frac{\partial}{\partial x_1}\Big)^{\alpha_1} \cdots \Big(\frac{\partial}{\partial x_n}\Big)^{\alpha_n}, \quad \alpha! = \textstyle\prod_1^n \alpha_i! \quad \text{et} \quad h^\alpha = \prod_1^n h_i^{\alpha_i}.$$

Définition 9.25 Polynôme et reste de Taylor
Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$ que l'on suppose k fois différentiable au point $x_0 \in U$. Le polynôme de Taylor d'ordre k de f en x_0, est

$$x \in \mathbb{R}^n \longrightarrow \sum_{j=0}^{k} \frac{1}{j!} D_x^j f \cdot (x - x_0)^{(j)} = \sum_{|\alpha| \leq k} \frac{\partial^\alpha f(x)}{\alpha!} (x - x_0)^\alpha .$$

Le reste de Taylor d'ordre k de f en x_0 est la différence entre f et son polynôme de Taylor, soit

$$R_k f(x_0, x) := f(x) - \sum_{j=0}^{k} \frac{1}{j!} D_{x_0}^j f \cdot (x - x_0)^{(j)} .$$

L'objectif des différentes formules de Taylor qui suivent est d'estimer, pour $x \in U$ ou bien pour x proche de x_0, ce reste de Taylor. On pourra, pour l'essentiel des applications, supposer que la fonction f est de classe C^{k+1} sur U (voire C^∞ !) Dans ce cas, les estimations de Taylor-Young et Taylor-Lagrange sont toutes deux conséquences de Taylor-intégrale. Nous les énonçerons cependant avec des hypothèses de régularité moindre.

D.3 Taylor-Young

La formule Taylor-Young à l'ordre 1, pour une application $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$ différentiable au point $x_0 \in U$, n'est autre que la définition de la différentielle. Elle s'écrit

$$f(x_0 + h) = f(x_0) + D_{x_0} f \cdot h + o(\|h\|) .$$

A l'ordre supérieur, elle s'énonce comme suit.

Proposition 9.26 Taylor-Young
Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$ k fois différentiable au point $x_0 \in U$. Alors

$$f(x_0 + h) = f(x_0) + D_{x_0} f \cdot h + \cdots + \frac{1}{k!} D_{x_0}^k f \cdot h^{(k)} + o(\|h\|^k) . \qquad (TY_k)$$

La formule de Taylor-Young donne donc un développement limité (ici à l'ordre k) pour f en x_0, c'est-à-dire une information sur le comportement asymptotique de $f(x)$ lorsque $x \to x_0$. La notation $o(\|h\|^k)$ signifie bien entendu que le "reste"

$$R_k f(x_0, x_0 + h) := f(x_0 + h) - \sum_{j=0}^{k} \frac{1}{j!} D_{x_0}^j f \cdot h^{(j)}$$

vérifie $\|R_k f(x_0, x_0 + h)\| / \|h\|^k \to 0$ quand $h \to 0$.

Esquisse de preuve Lorsque la régularité de f le permet, la formule de Taylor-Young est conséquence immédiate de la formule de Taylor avec reste intégral ou bien de l'inégalité de Taylor-Lagrange prouvées ci-dessous.

Nous nous contentons d'esquisser la preuve lorsque f est seulement supposée k fois différentiable au point. On démontre par récurrence sur k que l'assertion (TY_k) est vraie pour toute fonction $F : U \subset \mathbb{R}^n \to \mathbb{R}^N$ (N quelconque) k fois différentiable au point x_0. Sachant que (TY_{k-1}) est vraie, on l'applique à la différentielle Df de f. On en déduit (TY_k) pour f en utilisant les accroissements finis. □

D.4 Taylor-Lagrange et Taylor-intégrale

Dans tout ce paragraphe, on se donne une application $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$, et deux points $x_0, x_1 \in U$. On écrit

$$f(x_1) = \sum_{j=0}^{k} \frac{1}{j!} D^j_{x_0} f \cdot (x_1 - x_0)^{(j)} + R_k f(x_0, x_1)\,,$$

et on veut estimer le reste $R_k f(x_0, x_1)$.

Proposition 9.27 *Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$ et x_0, x_1 deux points de U.On suppose que tout le segment $[x_0, x_1]$ est inclus dans U.*

Inégalité de Taylor-Lagrange *On suppose que la fonction f est $(k+1)$ fois différentiable sur U. On pose $M = \sup_{t\in[0,1]} \|D^{(k+1)}_{x_t} f\|$ (voir 4.14). Alors*

$$\|R_k f(x_0, x_1)\| \leq \frac{M}{(k+1)!} \|x_1 - x_0\|^{k+1}\,.$$

Egalité de Taylor-Lagrange (fonctions scalaires) *On suppose que $p = 1$ et que la fonction $f : U \to \mathbb{R}$ est $(k+1)$ fois différentiable sur U. Il existe alors $\theta \in]0, 1[$ tel que*

$$R_k f(x_0, x_1) = \frac{1}{(k+1)!} D^{k+1}_{x_\theta} f \cdot (x_1 - x_0)^{(k+1)}\,.$$

Remarque 9.28 Lorsque $k = 0$, on retrouve respectivement le théorème des accroissements finis (corollaire 6.16), et le théorème de Rolle (rappel 6.12). L'égalité de Taylor-Lagrange n'est vraie que pour des fonctions à valeurs réelles. Cela a déjà été signalé lorsque $k = 0$.

Les deux formules de Tayor-Lagrange permettent d'estimer chaque reste de Taylor $R_k f(x_0, x_1)$, et non plus seulement le comportement asymptotique de $R_k f(x_0, x)$ lorsque $x \to x_0$ comme dans la formule de Taylor-Young.

Cependant, ce ne sont que des estimations (même dans l'égalité de Taylor-Lagrange, on ne sait pas "qui est θ"). Pour une expression exacte du reste, on utilisera la formule de Taylor avec reste intégral.

Proposition 9.29 Taylor-intégrale *Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$ et x_0, x_1 deux points de U tels que tout le segment $[x_0, x_1]$ soit inclus dans U. On suppose que la fonction f est de classe C^{k+1}. Alors on a une expression exacte du reste, soit*

$$R_k f(x_0, x_1) = \frac{1}{k!} \int_0^1 (1-t)^k D_{x_t}^{k+1} f \cdot (x_1 - x_0)^{(k+1)}\, dt\,.$$

Preuve des propositions 9.27 et 9.29
Pour $t \in [0,1]$, on note $x_t = x_0 + t(x_1 - x_0) \in U$ et on introduit la fonction d'une variable réelle définie pour pour $t \in [0,1]$ par

$$v(t) := f(x_t) = f(x_0 + t(x_1 - x_0))\,.$$

La fonction v a la même régularité que f. On a pour tout $1 \leq j \leq k+1$

$$v^{(j)}(t) = D_{x_t}^{j} f \cdot (x_1 - x_0)^{(j)}\,,$$

et donc l'égalité des restes $R_k f(x_0, x_1) = R_k v(0,1)$. Il nous suffira donc de montrer les deux énoncés ci-dessus pour la fonction d'une variable réelle v.

- On introduit

$$w(t) = v(t) + (1-t)v'(t) + \cdots + \frac{1}{k!}(1-t)^k\, v^{(k)}(t)\,.$$

On a $R_k v(0,1) = w(1) - w(0)$. Un calcul élémentaire montre que

$$w'(t) = \frac{1}{k!}(1-t)^k\, v^{(k+1)}(t)\,.$$

Soit $g : t \in [0,1] \to \frac{-1}{(k+1)!}(1-t)^{k+1}\, M \,\|x_1 - x_0\|^{k+1} \in \mathbb{R}$. L'inégalité de Taylor-Lagrange suit alors du théorème des accroissements finis puisque w est alors différentiable avec

$$\|w'(t)\| \leq \frac{1}{k!}(1-t)^k\, M\, \|x_1 - x_0\|^{k+1} = g'(t)\,.$$

- La formule de Taylor-intégrale suit de ce que, la fonction w étant C^1, on a

$$w(1) - w(0) = \int_0^1 w'(t)\, dt\,.$$

- Supposons maintenant que les fonctions f et v soient à valeurs scalaires. On choisit $\alpha \in \mathbb{R}$ de sorte que la fonction

$$\tilde{w}(t) = v(1) - w(t) - \alpha \frac{(1-t)^{k+1}}{(k+1)!}\,,$$

qui s'annule en $t = 1$ pour tout choix de α, s'annule également en $t = 0$. On a alors, en exprimant que $\tilde{w}(0) = 0$, l'égalité

$$v(1) = w(1) + \frac{\alpha}{(k+1)!} = \sum_{j=0}^{k} \frac{1}{j!} v^{(j)}(0) + \frac{\alpha}{(k+1)!} \,.$$

Reste à déterminer α. Le théorème de Rolle assure qu'il existe $\theta \in]0,1[$ tel que $\tilde{w}'(\theta) = 0$. On a

$$\tilde{w}'(t) = -\frac{(1-t)^k}{k!} v^{(k+1)}(t) + \alpha \frac{(1-t)^k}{k!}$$

pour tout $t \in [0,1]$. Puisque $\theta \neq 1$, on obtient finalement $\alpha = v^{(k+1)}(\theta)$. $\square$

E Etude des extrema d'une fonction

Dans cette section, nous étudions les extrema d'une fonction à valeurs réelles, définie sur un ouvert de $\mathbb{R}^n$. Le lecteur pourra se référer systématique–ment au cas $n = 1$ (fonction réelle d'une variable réelle), censé être plus familier.

E.1 Points critiques, condition à l'ordre 1

Définition 9.30 *Une fonction $f : X \to \mathbb{R}$ définie sur un espace métrique X admet au point m_0 un maximum local lorsqu'il existe un voisinage $V \subset X$ de m_0 tel que $f(m) \leq f(m_0)$ pour tout $m \in V$.*
Ce maximum local est strict lorsque $f(m) < f(m_0)$ quand $m \in V$ et $m \neq m_0$.
On définit de même la notion de minimum local (strict).

Lemme 9.31 *Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}$ une fonction différentiable. Si f admet en $m_0 \in U$ un maximum (ou un minimum) local, la différentielle $D_{m_0}f$ de f en ce point s'annule.*

Preuve Par définition de la différentielle, on a

$$f(m_0 + h) = f(m_0) + D_{m_0}f \cdot h + \|h\| \, \varepsilon(h) \,,$$

où la fonction $h \to \varepsilon(h)$ tend vers 0 quand $h \to 0$. Procédons par l'absurde et supposons $D_{m_0}f$ non nulle. Il existe donc un vecteur $u \in \mathbb{R}^n$ pour lequel on a $D_{m_0}f \cdot u > 0$. Appliquons l'identité précédente au vecteur $h = tu$ pour $t > 0$. Il vient

$$f(m_0 + tu) - f(m_0) = t \, [D_{m_0}f \cdot u + \|u\| \varepsilon(t)]$$

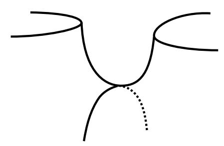
où $\varepsilon(t)$ tend vers 0 avec t. Le terme entre crochets est donc strictement positif pour $t > 0$ petit, et m_0 n'est donc pas un maximum local de f. La même considération, avec $t < 0$, montre de même que m_0 n'est donc pas un minimum local de f. $\square$

Définition 9.32 *Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}$ une fonction différentiable. Un point $m_0 \in U$ est un point critique de f lorsque $D_{m_0} f = 0$.*

Les extrema (locaux) d'une application différentiable f sont donc à rechercher parmi ses points critiques. Comme on le sait en dimension 1, un point critique ne fournit pas toujours un extremum local. Considérer par exemple la fonction $x \to x^3$ en l'origine.

Définition 9.33 *Le point critique $m_0 \in U$ est un point selle (ou col) pour f lorsque dans tout voisinage de m_0 la fonction prend des valeurs strictement supérieures, et strictement inférieures, à $f(m_0)$.*

On a représenté ci-contre le graphe d'une fonction $f : \mathbb{R}^2 \to \mathbb{R}$ autour d'un point selle. Ce graphe ressemble à une selle (de cheval) ou à un col (en montagne).



E.2 Points critiques à l'ordre 2

Dans certains cas, aller à l'ordre 2 permettra d'élucider la nature d'un point critique. C'est la question que nous abordons dans ce paragraphe.

Supposons la fonction $f : U \subset \mathbb{R}^n \to \mathbb{R}$ différentiable sur U, et deux fois différentiable au point m_0. Sa différentielle seconde $D^2_{m_0} f$ est une forme bilinéaire symétrique sur $\mathbb{R}^n$. Il lui est associé une forme quadratique, dont elle est la forme polaire, et qui est définie pour tout $u \in \mathbb{R}^n$ par

$$Q(u) = D^2_{m_0} f(u, u)\,.$$

Rappelons quelques faits concernant les formes quadratiques sur $\mathbb{R}^n$. C'est le moment de vérifier que vous savez les redémontrer !

Proposition 9.34 Rang, signature *Soit Q une forme quadratique sur $\mathbb{R}^n$. Il existe deux entiers p, q tels que $p + q \leq n$ et une base $(\epsilon_1, \cdots, \epsilon_n)$ de $\mathbb{R}^n$ telle que, pour tout vecteur $v \in \mathbb{R}^n$ avec $v = \sum v_i \epsilon_i$, on ait*

$$Q(v) = \sum_{i=1}^{p} v_i^2 - \sum_{j=1}^{q} v_{p+j}^2\,.$$

Le couple (p, q) est unique. C'est la signature de Q. Le rang de Q est $r = p + q$.

Définition 9.35 *On dit que la forme quadratique est positive si $q = 0$. Elle est définie positive si $p = n$.*

On dit que la forme quadratique est négative si $p = 0$. Elle est définie négative si $q = n$.

Munissons désormais $\mathbb{R}^n$ d'une norme.

Lemme 9.36 *La forme quadratique est positive si et seulement si $Q(v) \geq 0$ pour tout vecteur $v \in \mathbb{R}^n$.*
Elle est définie positive lorsque $Q(v) > 0$ pour tout vecteur non nul $v \in \mathbb{R}^n$. Dans ce cas, il existe une constante $c > 0$ telle qu'on ait $Q(v) \geq c\|v\|^2$ pour tout vecteur $v \in \mathbb{R}^n$.

On a bien sûr un énoncé similaire pour les formes négatives.
Preuve Contentons nous de prouver la dernière assertion, les autres étant immédiates. Notons $\mathbb{S} \subset \mathbb{R}^n$ la sphère unité relative à notre norme. La fonction continue

$$v \in \mathbb{S} \longrightarrow Q(v) \in \mathbb{R}$$

est par hypothèse strictement positive. Puisque $\mathbb{S}$ est compacte, cette fonction atteint son minimum $c > 0$. On conclut par homogénéité. □

Voyons maintenant ce que ces considérations, jointes à Taylor-Young, nous disent sur les points critiques.

Théorème 9.37 Nature des points critiques

Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}$ une application différentiable. Soit $m_0 \in U$ un point critique de f. On suppose que f est deux fois différentiable en m_0.

1. *Si f admet un minimum local en m_0, alors la forme quadratique $D^2_{m_0}f$ est positive.*
2. *Si la forme quadratique $D^2_{m_0}f$ est définie positive, alors f admet en m_0 un minimum local strict.*
3. *Si la forme quadratique $D^2_{m_0}f$ n'est ni positive ni négative (i.e. $p \neq 0$ et $q \neq 0$), le point m_0 est point selle pour f.*

Au lecteur d'énoncer et de démontrer les équivalents des points 1. et 2. en un maximum.
Preuve Tout ceci sera conséquence de Taylor-Young 9.26.

1. Soit $u \in \mathbb{R}^n$ un vecteur. On a $f(m_0 + tu) \geq f(m_0)$ pour tout réel t assez petit. Puisque

$$f(m_0 + tu) - f(m_0) = \frac{t^2}{2} D^2_{m_0}f(u,u) + o(t^2) = \frac{t^2}{2}\left(D^2_{m_0}f(u,u) + \varepsilon(t)\right)$$

où $\varepsilon(t)$ tend vers 0 avec t, il suit que $D^2_{m_0}f(u,u) \geq 0$.

2. Le lemme 9.36 assure qu'il existe une constante $c > 0$ avec, pour tout vecteur $u \in \mathbb{R}^n$, $D^2_{m_0} f(u,u) \geq c\|u\|^2$. La formule de Taylor-Young donne que

$$f(m_0 + u) - f(m_0) = \frac{1}{2} D^2_{m_0} f(u,u) + \|u\|^2 \,\varepsilon(u)\,,$$

où $\varepsilon(u) \to 0$ lorsque $u \to 0$. Pour u assez petit on aura $|\varepsilon(u)| \leq c/4$ et donc

$$f(m_0 + u) - f(m_0) \geq \frac{c}{4}\|u\|^2\,.$$

3. Lorsque la forme quadratique $D^2_{m_0} f$ n'est pas définie, on choisit deux vecteurs $u, v \in \mathbb{R}^n$ avec $D^2_{m_0} f(u,u) > 0$ et $D^2_{m_0} f(v,v) < 0$, et l'on montre comme ci-dessus que lorsque $t \neq 0$ est assez petit on a $f(m_0 + tu) > f(m_0)$ tandis que $f(m_0 + tv) < f(m_0)$. □

Par contre, lorsque la forme quadratique $D^2_{m_0} f$ est positive, mais pas définie positive (donc est dégénérée), l'examen de la différentielle seconde ne permet pas de conclure quant à la nature du point critique m_0. Les ennuis commencent dès la dimension 1 avec, en $m_0 = 0$, les fonctions $x \to x^4$, $x \to -x^4$ ou encore $x \to x^3$.

E.3 Formes quadratiques en dimension 2

> En dimension quelconque, il faudra se fatiguer un peu pour déterminer la nature d'une forme quadratique, par exemple en exhibant une base orthogonale comme en 9.34. En dimension 2, c'est bien plus facile.

On note (x, y) les coordonnées dans $\mathbb{R}^2$.

Proposition 9.38 *Soit* $Q(x,y) = rx^2 + 2sxy + ty^2$ *une forme quadratique.*

La forme Q *est de signature* $(1,1)$ *si et seulement si* $rt - s^2 < 0$.

La forme quadratique est définie si et seulement si $rt - s^2 > 0$. *Elle est alors définie positive (resp. définie négative) si de plus* $r > 0$ *(resp.* $r < 0$*).*

Esquisse de preuve Tout vecteur de $\mathbb{R}^2$ est proportionnel à l'un des vecteurs $(x, 1)$ $(x \in \mathbb{R})$ ou bien à $(1, 0)$.

Le discriminant du trinôme $Q(x,1) = rx^2 + 2sx + t$ est $\Delta = s^2 - rt$. Lorsque $\Delta > 0$ (ce qui inclut le cas $r = 0$ et $s \neq 0$), ce trinôme change de signe et la forme quadratique Q prend donc des valeurs positives, et des valeurs négatives.

Lorsque $\Delta < 0$, cela impose que r et t soient de même signe. Dans le cas où ils sont tous deux positifs (resp. négatifs), la forme quadratique est partout positive (resp. négative).

Lorsque $\Delta = 0$, la forme est positive (si $r > 0$) ou négative (si $r < 0$) mais on a des vecteurs isotropes. □

On revient maintenant à une application $f : U \subset \mathbb{R}^2 \to \mathbb{R}$, que l'on suppose deux fois différentiable au point critique m_0.

Corollaire 9.39 Notations de Monge

Soit $f : U \subset \mathbb{R}^2 \to \mathbb{R}$ une application différentiable. Soit $m_0 \in U$ un point critique de f. On suppose que f est deux fois différentiable en m_0 et on note $r = \partial^2 f/\partial x^2$, $s = \partial^2 f/\partial x \partial y$ et $t = \partial^2 f/\partial y^2$ les dérivées partielles d'ordre 2 de f en m_0.

Si $rt - s^2 < 0$, la fonction f possède un point selle en m_0.

Si $rt - s^2 > 0$ et $r > 0$ (resp. $r < 0$), la fonction f possède minimum (resp. maximum) local strict en m_0.

Preuve La proposition précédente s'applique à la forme quadratique associée à $D^2_{m_0} f$, soit

$$Q(x,y) = D^2_{m_0} f\,\big((x,y),(x,y)\big) = rx^2 + 2sxy + ty^2\,.$$

Conclure avec le théorème 9.37. □

E.4 Matrice hessienne

Revenons à la dimension quelconque, et soit $f : U \subset \mathbb{R}^n \to \mathbb{R}$ deux fois différentiable au point $m_0 \in U$. On note $(x_1, \cdots, x_n)$ les coordonnées de $\mathbb{R}^n$.

Définition 9.40 *La matrice hessienne (ou simplement hessienne) de f au point m_0 est la matrice relativement à la base canonique de la forme bilinéaire symétrique $D^2_{m_0} f$, soit*

$$\operatorname{Hess}_{m_0} f = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n} \\ & \ddots & \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \dfrac{\partial^2 f}{\partial x_n^2} \end{pmatrix}.$$

Le théorème de Schwartz nous assure que cette matrice est symétrique. On a en particulier en dimension 2, avec les notations de Monge,

$$\operatorname{Hess}_{m_0} f = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x \partial y} \\ \dfrac{\partial^2 f}{\partial x \partial y} & \dfrac{\partial^2 f}{\partial y^2} \end{pmatrix} = \begin{pmatrix} r & s \\ s & t \end{pmatrix}$$

et le déterminant de la hessienne (discriminant de la forme quadratique associée) est $rt - s^2$.

F Fonctions convexes

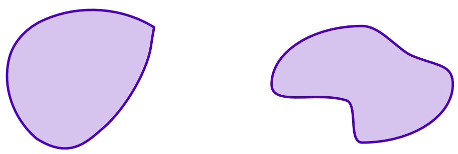
Nous ne ferons qu'effleurer cette notion, qui joue un rôle fondamental notamment en optimisation.

Rappel 9.41 *Une partie $C \subset \mathbb{R}^n$ est convexe lorsque, pour tous points $x_0, x_1 \in C$, le segment*

$$[x_0, x_1] = \{x_t = (1-t)\,x_0 + t\,x_1 \mid t \in [0,1]\}$$

est entièrement inclus dans C.

Les parties convexes de $\mathbb{R}$ sont les intervalles.



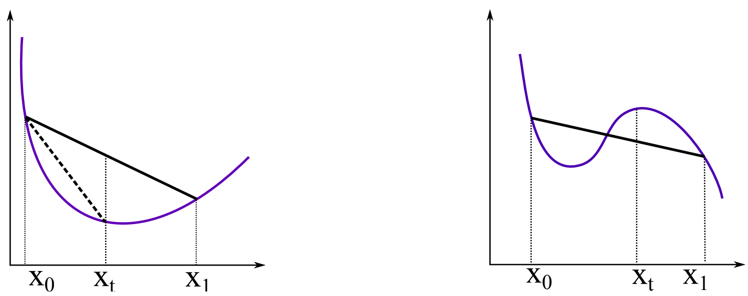
Partie convexe, pas convexe

Définition 9.42 *Soit $C \subset \mathbb{R}^n$ une partie convexe. Une fonction $f : C \to \mathbb{R}$ est dite convexe lorsque, pour tous points $x_0, x_1 \in C$ et tout réel $t \in [0,1]$, on a*

$$f(x_t) \leq (1-t)\,f(x_0) + t\,f(x_1)\,.$$

La fonction est convexe lorsque son graphe est en-dessous des cordes.

Remarque 9.43 *Une fonction est convexe lorsque sa restriction à chaque segment inclus dans son domaine de définition est convexe.*



Fonction convexe, pas convexe

Avant de passer à la dimension quelconque, on se rafraîchit la mémoire en dimension 1.

Rappel 9.44 *Soient $I \subset \mathbb{R}$ un intervalle et $f : I \to \mathbb{R}$ une fonction.*
- *Supposons f dérivable sur I. Alors f est convexe si son graphe est au-dessus de chaque tangente affine de f : pour tous $s, t \in I$ on a l'inégalité $f(t) \geq f(s) + (t-s)\, f'(s)$.*
- *Supposons f deux fois dérivable sur I. Alors f est convexe si et seulement si $f'' \geq 0$.*

On va démontrer cet énoncé, et le généraliser cet énoncé à la dimension quelconque. On commence par caractériser, par leur différentielle, les applications convexes différentiables.

Proposition 9.45 *Soit $\Omega \subset \mathbb{R}^n$ un ouvert convexe. Soit $f : \Omega \to \mathbb{R}$ une fonction différentiable. La fonction f est convexe si et seulement si on a*

$$f(x_1) \geq f(x_0) + D_{x_0} f\,(x_1 - x_0)$$

pour tous $x_0, x_1 \in \Omega$ (la fonction est au-dessus de ses tangentes).

Preuve Supposons f convexe. Soient $x_0, x_1 \in \Omega$. On se ramène à une fonction de variable réelle en introduisant la fonction $g : t \in [0,1] \to \mathbb{R}$ définie par $g(t) = f(x_t) = f(x_0 + t(x_1 - x_0))$. La convexité de f, et donc de g, assure que pour tout $t \in]0,1]$, on a l'inégalité

$$\frac{g(t) - g(0)}{t} \leq \frac{g(1) - g(0)}{1}$$

entre les pentes des cordes (figure ci-dessus). Le résultat suit puisque, lorsque $t \to 0$, le ratio de gauche converge vers $g'(t) = D_{x_0} f(x_1 - x_0)$.

Démontrons la réciproque. On va faire jouer des rôles symétriques aux points x_0 et x_1. Soit $t \in]0,1[$. On a par hypothèse

$$\begin{aligned} f(x_0) &\geq f(x_t) + D_{x_t} f\,(x_0 - x_t) = f(x_t) + t\, D_{x_t} f\,(x_0 - x_1) \\ f(x_1) &\geq f(x_t) + D_{x_t} f\,(x_1 - x_t) = f(x_t) - (1-t)\, D_{x_t} f\,(x_0 - x_1)\,. \end{aligned} \tag{9.4}$$

Le résultat suit en multipliant la première ligne par $1 - t$, la seconde par t, et en les additionant. □

On en déduit immédiatement le corollaire suivant.

Corollaire 9.46 *Soit $\Omega \subset \mathbb{R}^n$ un ouvert convexe. Soit $f : \Omega \to \mathbb{R}$ une fonction convexe différentiable. On suppose que la différentielle de f au point $a \in \Omega$ s'annule. Alors f admet au point a un minimum global.*

Terminons par un critère de convexité portant sur la différentielle seconde.

Proposition 9.47 *Soit $\Omega \subset \mathbb{R}^n$ un ouvert convexe. Soit $f : \Omega \to \mathbb{R}$ une fonction deux fois différentiable. La fonction f est convexe si et seulement si sa différentielle seconde $D_x^2 f$ en chaque point $x \in \Omega$ est une forme bilinéaire positive, i.e. si elle vérifie pour tout vecteur $h \in \mathbb{R}^n$ l'inégalité*

$$D_x^2 f(h,h) \geq 0\,.$$

Preuve On se ramène de nouveau à la dimension 1 en observant que, si $x, h \in \mathbb{R}^n$ et $t \in \mathbb{R}$ sont tels que $x + t_0 h \in \Omega$, la différentielle seconde en t_0 de l'application $g : t \to f(x + th)$ est $g''(t_0) = D^2_{x+t_0 h} f(h, h)$.

Soit donc $g : I \subset \mathbb{R} \to \mathbb{R}$ une application deux fois différentiable sur l'intervalle ouvert I.

Supposons $g''(t) \geq 0$ pour tout $t \in I$. L'égalité de Taylor-Lagrange à l'ordre 2 montre que g est en dessus de ses tangentes. Réciproquement, si g est convexe elle est en dessus de ses tangentes, et la formule de Taylor-Young montre que sa dérivée seconde est partout positive ou nulle. $\square$

10. Sous-variétés

On revient, avec un objectif plus géométrique, sur le théorème des fonctions implicites.

Notre prétexte pour introduire la notion d'hypersurface de $\mathbb{R}^n$ est le problème des extrêma liés. Jusqu'ici, on a fait du calcul différentiel sur des ouverts de $\mathbb{R}^n$. Pour étendre le champ d'utilisation des techniques qu'on a développées sur $\mathbb{R}^n$, on introduit la notion d'hypersurface de $\mathbb{R}^n$ (en codimension 1) ou plus généralement la notion de sous-variété (en codimension quelconque).

A Motivation

Soit $g : X \to \mathbb{R}$ une fonction. On recherche les extrema de g. Lorsque g est définie et continue sur un espace métrique compact X, on a vu que g est bornée et qu'elle atteint ses bornes (corollaire 2.12). Mais nous n'avons cependant pas de méthode pour déterminer ces bornes.

Si maintenant $g : \mathbb{R}^n \to \mathbb{R}$ est définie sur $\mathbb{R}^n$ (ou un ouvert de $\mathbb{R}^n$) et différentiable, on a vu dans le chapitre 9 que g ne peut admettre de maximum ou de minimum (local) qu'en un point critique.

Intéressons-nous maintenant à un problème d'extrema sous contrainte. On se donne une partie $M \subset \mathbb{R}^n$, une fonction $g : \mathbb{R}^n \to \mathbb{R}$ différentiable, et on cherche les extrema de la restriction $g_{|M} : M \to \mathbb{R}$. Lorsque M est quelconque, on ne sera pas beaucoup plus avancés. Par contre, lorsque M est une hypersurface de $\mathbb{R}^n$ (ou plus généralement une sous-variété), on pourra généraliser le résultat précédent à travers la notion de "point critique relatif" (théorème 10.12).

B Hypersurfaces de $\mathbb{R}^n$

Rappelons que tout fermé de $\mathbb{R}^n$ est l'ensemble des zéros d'une fonction scalaire de classe C^1 (exercice 9.23). Lorsqu'on imposera à cette fonction scalaire d'être submersive, on aura par contre un "joli" ensemble de zéros.

Définition 10.1 *Une application $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$ de classe C^1 est submersive (ou est une submersion) si sa différentielle $D_x f \in L(\mathbb{R}^n, \mathbb{R}^p)$ en tout point $x \in U$ est une application linéaire surjective.*

Remarque 10.2 *Si $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$ est submersive, il suit que $n \geq p$.*

Une application linéaire $\ell : \mathbb{R}^n \to \mathbb{R}^p$ surjective est submersive. En particulier, une forme linéaire $\ell : \mathbb{R}^n \to \mathbb{R}$ non nulle est submersive.

Si la différentielle $D_{x_0} f \in L(\mathbb{R}^n, \mathbb{R}^p)$ au point x_0 d'une application f de classe C^1 est surjective il suit que, pour tout point x suffisamment proche de x_0, la différentielle $D_{x_0} f \in L(\mathbb{R}^n, \mathbb{R}^p)$ est encore surjective.

Une application scalaire $f : U \subset \mathbb{R}^n \to \mathbb{R}$ de classe C^1 est submersive si et seulement si sa différentielle en chaque point $x \in U$ est non nulle.

On va définir géométriquement une hypersurface de $\mathbb{R}^n$ comme étant un objet qui ressemble, localement et avec des lunettes différentiables, à un hyperplan affine de $\mathbb{R}^n$ (1).

Dans la pratique, les hypersurfaces seront souvent définies analytiquement, d'où l'intérêt de la définition (2). Par exemple, la sphère unité de $\mathbb{R}^n$ est définie par l'équation

$$\mathbb{S}^n = \{x \in \mathbb{R}^n \mid \|x\|^2 = 1\}\,.$$

Enfin, une hypersurface se lira localement comme un graphe au-dessus d'un hyperplan (3).

Chacune de ces définitions équivalentes des hypersurfaces a son intérêt, et on choisira de travailler avec l'une ou l'autre selon le problème envisagé.

Proposition-Définition 10.3 Hypersurfaces

Soit $n \geq 2$. Une hypersurface $M \subset \mathbb{R}^n$ est une partie de $\mathbb{R}^n$ qui vérifie l'une des propriétés équivalentes suivantes :

1. Redressement : *pour tout point $m \in M$, il existe un voisinage ouvert $m \in U \subset \mathbb{R}^n$ et un C^1-difféomorphisme $\varphi : U \to V$ qui envoie le point m sur l'origine et M sur un morceau d'hyperplan, soit*

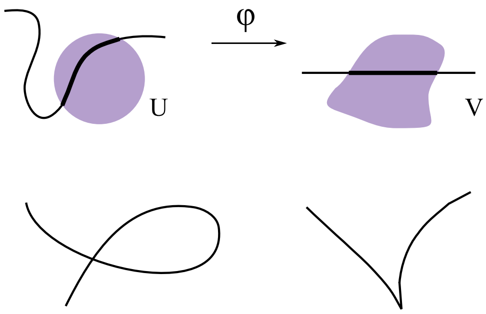
$$\varphi(M \cap U) = (\mathbb{R}^{n-1} \times \{0\}) \cap V\,.$$

2. Equation : *pour tout point $m \in M$, il existe un voisinage ouvert $m \in U \subset \mathbb{R}^n$ et une application scalaire $f : U \to \mathbb{R}$ de classe C^1 submersive, telle que*

$$M \cap U = \{p \in U \mid f(p) = 0\}\,.$$

3. Graphe : *pour tout point $m \in M$, quitte à permuter les coordonnées, il existe un voisinage ouvert $m \in U := W \times I \subset \mathbb{R}^n$ (où $W \subset \mathbb{R}^{n-1}$ et $I \subset \mathbb{R}$ sont ouverts) et une application $h : W \to I$ de classe C^1 dont $M \cap U$ soit le graphe, c'est-à-dire tels que*

$$M \cap U = \{(x, h(x)) \in W \times I\} \subset \mathbb{R}^{n-1} \times \mathbb{R}\,.$$



En haut, une hypersurface de $\mathbb{R}^2$ (courbe) que l'on a redressée (1).
En bas, deux courbes de $\mathbb{R}^2$ qui ne sont pas des hypersurfaces : auto-intersection (exercice 10.5) ou point anguleux (exercice 10.10).

On dit qu'une hypersurface $M \subset \mathbb{R}^n$ est de dimension $n-1$. Via l'une des définition (1) ou (3), elle est en effet localement paramétrée par un ouvert de $\mathbb{R}^{n-1}$. Partant d'un point de $\mathbb{R}^n$ (n degrés de liberté), la condition (2) impose une contrainte ; il est raisonnable de passer à $(n-1)$ degrés de liberté.

Preuve (1)⇒(2) Notons $\varphi_1, \cdots, \varphi_n$ les composantes de l'application φ. Puisque $\varphi : U \to V$ est un difféomorphisme, il suit que les différentielles $D_p\varphi_i$ $(1 \leq i \leq n)$ en chaque point $p \in U$ sont linéairement indépendantes, donc a fortiori non nulles. L'application $f := \varphi_n : U \to \mathbb{R}$ est donc une submersion. Et, par construction, on a $M \cap U = \{p \in U \mid f(p) = 0\}$.

(2)⇒(3) Soit $f : U \to \mathbb{R}$ une application submersive pour laquelle $M \cap U = \{p \in U \mid f(p) = 0\}$. La différentielle $D_m f : \mathbb{R}^n \to \mathbb{R}$ est une forme linéaire surjective, c'est-à-dire non nulle. Quitte à permuter les coordonnées, on peut donc supposer que $D_m f(e_n) \neq 0$ où e_n désigne le dernier vecteur de la base. Le théorème des fonctions implicites s'applique au point m à l'application $f : U \subset \mathbb{R}^{n-1} \times \mathbb{R} \to \mathbb{R}$ et donne le résultat.

(3)⇒(1) Supposons que $M \cap (W \times I)$ soit le graphe de l'application $h : W \subset \mathbb{R}^{n-1} \to I \subset \mathbb{R}$. La matrice jacobienne de l'application

$$\varphi : (x_1, \cdots, x_{n-1}; x_n) \to (x_1, \cdots, x_{n-1}; x_n - h(x_1, \cdots, x_{n-1}))$$

est, en chaque point de $W \times I$, une matrice triangulaire inférieure inversible

$$J\varphi = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ * & & 1 \end{pmatrix}.$$

L'application φ réalise donc un difféomorphisme local d'un voisinage U de m sur son image $V = \varphi(U)$. Et, par construction, φ redresse $M \cap U$ sur un morceau d'hyperplan. □

C Exemples

Quelques exemples, qui sont autant d'exercices !

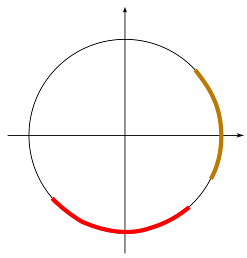
• Un hyperplan affine $H \subset \mathbb{R}^n$, ou un ouvert de cet hyperplan, constituent les hypersurfaces les plus simples de $\mathbb{R}^n$.

• La sphère unité de $\mathbb{R}^n$ euclidien, soit

$$\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n \mid \|x\|^2 = 1\}$$

est une hypersurface de $\mathbb{R}^n$, puisque l'application $x \in \mathbb{R}^n \to \|x\|^2 \in \mathbb{R}$ est submersive en dehors de l'origine. En particulier, le cercle $\mathbb{S}^1$ est une hypersurface de $\mathbb{R}^2$, c'est-à-dire une courbe régulière de $\mathbb{R}^2$.

Au voisinage de tout point $(x, y) \in \mathbb{S}^1$ avec $y \neq 0$, le cercle est le graphe d'une application $x \to y(x)$.
Au voisinage de tout point $(x, y) \in \mathbb{S}^1$ avec $x \neq 0$, le cercle est le graphe d'une application $y \to x(y)$.



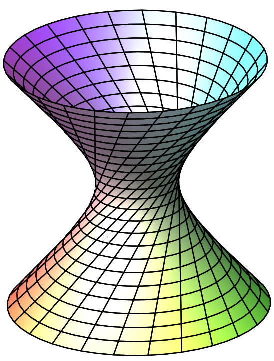
• L'hyperboloïde à une nappe

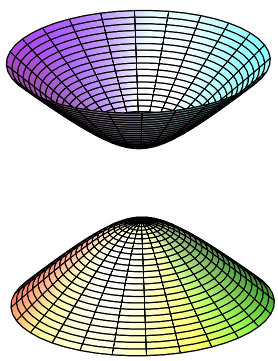
$$H_1 = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 - z^2 = 1\}$$

et l'hyperboloïde à deux nappes

$$H_2 = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 - z^2 = -1\}$$

sont des hypersurfaces de $\mathbb{R}^3$, c'est-à-dire surfaces régulières de $\mathbb{R}^3$.

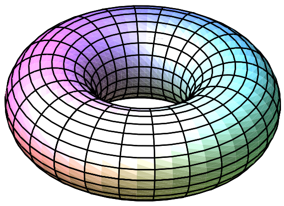




• Le tore de révolution (chambre à air) défini pour $0 < b < a$ par

$$\mathbb{T}^2 = \{(x,y,z) \in \mathbb{R}^3 \mid \big((x^2+y^2)^{1/2} - a\big)^2 + z^2 = b^2\}$$

est une hypersurface de $\mathbb{R}^3$, c'est-à-dire une surface régulière de $\mathbb{R}^3$.



• Le groupe spécial linéaire $\mathrm{Sl}_n\mathbb{R}$, constitué des matrices de déterminant 1, est une hypersurface de $M_n\mathbb{R} \simeq \mathbb{R}^{n^2}$.

Les deux exercices qui suivent ne sont pas bien compliqués. Mais, en première lecture, on suggère de se concentrer sur les exemples d'hypersurfaces plutôt que sur les contre-exemples.

Exercice 10.4 En quels points l'application $f : (x,y) \in \mathbb{R}^2 \to xy \in \mathbb{R}$ est-elle submersive ? Pour quelles valeurs de $c \in \mathbb{R}$ l'ensemble de niveau

$$\mathcal{H}_c = \{(x,y) \in \mathbb{R}^2 \,|\, f(x,y) = c\}$$

est-il une hypersurface de $\mathbb{R}^2$?

Exercice 10.5 Soit l'application

$$g : t \in]-\pi/2, \pi/2[\to \sin t \, \cos t \, (\cos t, \sin t\,) \in \mathbb{R}^2 \,.$$

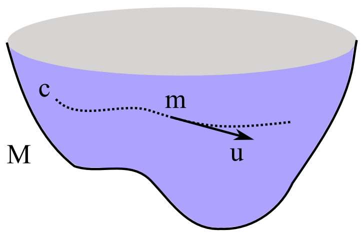
Montrer que g est de classe C^1, que pour tout $t \in]-\pi/2, \pi/2[$ on a $g'(t) \neq 0$ et que g est injective. Dessiner la courbe plane $\mathcal{C}$ image de g : est-ce une hypersurface de $\mathbb{R}^2$? On pourra utiliser un argument de connexité.

D Espace tangent

On va introduire, pour chaque point $m \in M$ d'une hypersurface de $\mathbb{R}^n$, l'espace vectoriel tangent T_mM à l'hypersurface en ce point. Nous en donnerons en 10.8 plusieurs définitions équivalentes, qui reflètent les trois définitions équivalentes des hypersurfaces (10.3).

Définition 10.6 *Soient $M \subset \mathbb{R}^n$ une hypersurface et m un point de M.*
Un vecteur $u \in \mathbb{R}^n$ est tangent à l'hypersurface M au point m si et seulement si il existe une courbe paramétrée $c :]-1, 1[\to \mathbb{R}^n$ de classe C^1 qui est tracée sur M (c'est-à-dire telle que $c(t) \in M$ pour tout t) et pour laquelle $c(0) = m$ et $c'(0) = u$.

0. Vecteurs vitesse : *L'ensemble T_mM de tous les vecteurs tangents à M au point m est l'espace tangent à M en m.*



Une courbe tracée sur l'hypersurface M, et un vecteur tangent u au point $m \in M$.

Exemple 10.7 Soit $M \subset H \subset \mathbb{R}^n$ une hypersurface de $\mathbb{R}^n$ qui est un ouvert d'un hyperplan affine H de $\mathbb{R}^n$. L'espace tangent à M en chaque point $m \in M$ est l'espace vectoriel directeur $\mathcal{H}$ de H.

Proposition 10.8 *Les notations sont celles de la définition 10.3. Soit $M \subset \mathbb{R}^n$ une hypersurface. L'espace tangent au point $m \in M$ s'obtient au choix par*

1. Redressement : *Si φ redresse localement M en $\mathbb{R}^{n-1} \times \{0\}$ avec $\varphi(m) = 0$, on a*

$$T_mM = (D_0\varphi^{-1})(\mathbb{R}^{n-1} \times \{0\}).$$

2. Equation : *Si M est localement définie par l'équation $f = 0$, avec $f : \mathbb{R}^n \to \mathbb{R}$ submersive, on a*

$$T_mM = \mathrm{Ker} D_m f .$$

3. Graphe : *Si M est localement le graphe de $h : W \subset \mathbb{R}^{n-1} \to \mathbb{R}$ avec $m = (0, h(0))$ on a*

$$T_mM = \{(v, D_0h(v)) \in \mathbb{R}^{n-1} \times \mathbb{R} \mid v \in \mathbb{R}^{n-1}\}.$$

En particulier, T_mM est un sous-espace vectoriel de $\mathbb{R}^n$ de dimension $n-1$.

Preuve (0)⇔(1) Le difféomorphisme $\varphi : U \to V$ (avec $\varphi(m) = 0$) induit une bijection entre

d'une part, l'ensemble des courbes $c :]-1,1[\to \mathbb{R}^n$ de classe C^1 tracées sur M et pour lesquelles $c(0) = m$

et, d'autre part, l'ensemble des courbes $\gamma :]-1,1[\to \mathbb{R}^n$ de classe C^1 tracées sur l'hyperplan $\mathbb{R}^{n-1} \times \{0\}$, et pour lesquelles $\gamma(0) = 0$.
Cette bijection est donnée par $\gamma = \varphi \circ c$, et on a $\gamma'(0) = (D_m\varphi)(c'(0))$. Le résultat suit immédiatement de l'exemple 10.7. En particulier, T_mM est un sous-espace vectoriel de dimension $n-1$ de $\mathbb{R}^n$.

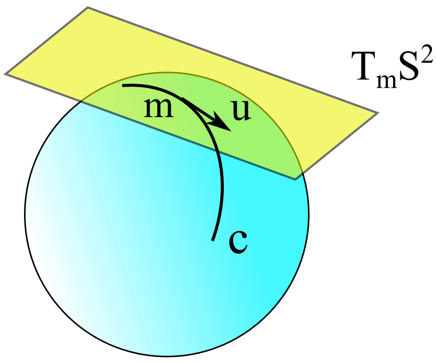
(0)⇔(2) Soit $u \in T_mM$. Cela signifie qu'il existe il existe une courbe $c :]-1,1[\to M \subset \mathbb{R}^n$ de classe C^1 telle que $c(0) = m$ et $c'(0) = u$. On a donc $f(c(t)) = 0$ pour tout $t \in]-1,1[$. En dérivant en $t = 0$, on obtient $D_mf(u) = 0$. On a donc $T_mM \subset \mathrm{Ker} D_mf$, d'où l'égalité de ces deux sous-espaces vectoriels de même dimension $n-1$.

(0)⇔(3) On a cette fois $\{(v, D_0h(v)) \in \mathbb{R}^{n-1} \times \mathbb{R} \mid v \in \mathbb{R}^{n-1}\} \subset T_mM$, puisque chaque courbe $t \to (tv, h(tv))$ est tracée sur M (pour t petit). On conclut comme ci-dessus à l'égalité de ces sous-espaces vectoriels de même dimension. □

Exercice 10.9 Déterminer les espaces tangents, aux points de votre choix, aux hypersurfaces du paragraphe C. En particulier, déterminer l'espace tangent au groupe spécial linéaire $\mathrm{Sl}_n\mathbb{R}$ en l'identité.

Exercice 10.10 Montrer que l'image de l'application $t \in \mathbb{R} \to (t, |t|) \in \mathbb{R}^2$ n'est pas une courbe régulière de $\mathbb{R}^2$.
Sinon, quel serait son espace tangent au point $(0, 0)$?

Remarque 10.11 *Tel qu'il a été défini ci-dessus, T_mM est un sous-espace vectoriel (hyperplan) de $\mathbb{R}^n$. En pratique, on dessinera plutôt l'espace affine tangent, qui passe par le point m et d'espace vectoriel directeur T_mM.*



Le plan affine tangent en m à la sphère $\mathbb{S}^2 \subset \mathbb{R}^3$.

E Extrema liés

Nous nous intéressons maintenant à un problème de recherche d'extrema sous contrainte. Données deux fonctions scalaires $f, g : \mathbb{R}^n \to \mathbb{R}$, on cherche les extrema de la restriction de g à l'ensemble $\{f = 0\}$ des zéros de f.

Par exemple on cherche à déterminer les points du tore de révolution, d'équation $\mathcal{T} = \{(x, y, z) \in \mathbb{R}^3 \mid \left((x^2 + y^2)^{1/2} - 2\right)^2 + z^2 = 1\}$, qui sont à distance minimale du point $p_0 = (3, 4, 1)$.

Théorème 10.12 *Soient $f, g : \mathbb{R}^n \to \mathbb{R}$ deux fonctions de classe C^1. On suppose que, en chaque point $m \in M := \{f = 0\}$, la fonction f est submersive, et donc que M est une hypersurface de $\mathbb{R}^n$.*
On suppose que la restriction $g_{|M} : M \to \mathbb{R}$ admet un extremum local en $m_0 \in M$. Alors le point m_0 est un point critique relatif de g, c'est-à-dire que

— *pour tout vecteur $u \in T_{m_0}M$ on a $D_{m_0}g(u) = 0$*
— *ou, de façon équivalente, il existe un scalaire $\lambda \in \mathbb{R}$ tel que*

$$D_{m_0}g = \lambda\, D_{m_0}f\,.$$

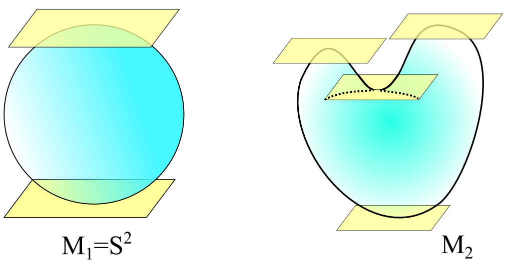
Le scalaire λ est alors unique ; c'est le multiplicateur de Lagrange.

Remarque 10.13 Tout comme dans le lemme 9.31, le fait que m_0 soit un point critique relatif de g sur M est une condition nécessaire, mais non suffisante, pour que g ait en m_0 un extremum local relatif (c'est-à-dire pour que la restriction $g_{|M} : M \to \mathbb{R}$ admette un extremum local en $m_0 \in M$.).

Prenons l'exemple de la fonction hauteur $f : (x, y, z) \in \mathbb{R}^3 \to z \in \mathbb{R}$ dans $\mathbb{R}^3$ euclidien. Son gradient est le vecteur constant $\mathrm{grad}_{(x,y,z)}f = (0, 0, 1)$: il ne s'annule jamais, donc f n'a aucun point critique.

En restriction à la sphère $M_1 = \mathbb{S}^2$, cette fonction hauteur admet exactement deux points critiques relatifs, qui sont les points où l'espace tangent est orthogonal au vecteur $(0, 0, 1)$. Ce sont les pôles $N = (0, 0, 1)$, où $f_{|M_1}$ atteint son maximum, et $S = (0, 0, -1)$, où $f_{|M_1}$ atteint son minimum.

Pour la surface M_2 dessinée à droite, la restriction $f_{|M_2}$ de la fonction hauteur possède quatre points critiques relatifs : un maximum et un minimum global, un maximum local et un point selle.



Les points critiques relatifs pour la fonction hauteur.

Preuve du théorème Supposons que la restriction $g_{|M} : M \to \mathbb{R}$ admette un extremum local au point $m_0 \in M$. Il suit que, pour toute courbe $c :]-1,1[\to \mathbb{R}^n$ de classe C^1 tracée sur M telle que $c(0) = m_0$, la fonction $t \in]-1,1[\to g(c(t)) \in \mathbb{R}$ admet en $t = 0$ un extremum local et a donc une dérivée nulle en $t = 0$. On a donc $D_{m_0}g(u) = 0$ pour $u = c'(0)$. Ainsi $T_{m_0}M \subset \mathrm{Ker}D_{m_0}g$.

Puisque $T_{m_0}M = \mathrm{Ker}D_{m_0}f$, cette condition est équivalente à demander que $\mathrm{Ker}D_{m_0}f \subset \mathrm{Ker}D_{m_0}g$. Soit la forme linéaire $\mathrm{Ker}D_{m_0}g$ est nulle (on a alors $\lambda = 0$). Soit elle a même noyau que $\mathrm{Ker}D_{m_0}f$, et dans ce cas les deux formes linéaires $D_{m_0}g$ et $D_{m_0}f$ sont proportionnelles comme annoncé. □

Exercice 10.14 Soit $M \subset \mathbb{R}^n$ une hypersurface compacte de $\mathbb{R}^n$ euclidien. Soit $p_0 \in \mathbb{R}^n$ un point pris hors de M. On suppose que la distance $d(p_0, M)$ est atteinte en un point $m \in M$. Traduire géométriquement la condition des extrema liés.

On peut notamment se servir des extrema liés pour montrer le théorème de réduction suivant (diagonalisation simultanée de deux matrices symétriques réelles, dont l'une est définie positive).

Exercice 10.15 Diagonalisation simultanée

Soient S_0 et S deux matrices symétriques réelles de taille (n,n). On suppose que la matrice S_0 est définie positive. Alors, il existe une base de $\mathbb{R}^n$ qui est orthonormale pour S_0 et orthogonale pour S. Autrement dit, il existe une matrice inversible $P \in \mathrm{Gl}_n\mathbb{R}$ telle que

$$^tPS_0P = \mathrm{Id} \text{ et } {}^tPSP \text{ est une matrice diagonale.}$$

Ces deux matrices symétriques sont respectivement associées aux formes quadratiques q_0 et q. On note $\mathbb{S} \subset \mathbb{R}^n$ la sphère unité de q_0, qui est compacte (q_0 est une structure euclidienne). L'application continue $q : x \in \mathbb{S} \to \mathbb{R}$ atteint son maximum en un vecteur $v_1 \in \mathbb{S}$, qui est donc unitaire pour q_0.

Utiliser les extrema liés pour montrer que v_1 est vecteur propre de S, et que l'orthogonal de v_1 pour la forme quadratique q_0 est stable par S. Conclure par récurrence sur la dimension.

F * Sous-variétés *

Tout ce qu'on vient de faire (énoncés et démonstrations) pour les hypersurfaces, c'est-à-dire en codimension 1, se transpose à la dimension quelconque. On parle alors de sous-variété de $\mathbb{R}^n$.

Une sous-variété $M \subset \mathbb{R}^n$ de dimension p (et de codimension $n-p$) est un objet qui, au choix :

(1) se redresse localement sur un ouvert d'un sous-espace vectoriel $\mathbb{R}^p \times \{0\} \subset \mathbb{R}^p \times \mathbb{R}^{n-p}$ de dimension p de $\mathbb{R}^n$

(2) est localement l'ensemble des zéros d'une application de classe C^1 submersive $f : U \subset \mathbb{R}^n \to \mathbb{R}^{n-p}$

(3) est localement, et après permutation des coordonnées, le graphe d'une application $h : W \subset \mathbb{R}^p \to \mathbb{R}^{n-p}$.

Les preuves de ces équivalences sont laissées au lecteur. Il suffit d'adapter les preuves données pour les hypersurfaces.

Via l'une des définitions (1) ou (3), notre sous-variété est localement paramétrée par un ouvert de $\mathbb{R}^p$. Partant d'un point de $\mathbb{R}^n$ (n degrés de liberté), la condition (2) impose $n-p$ contraintes "indépendantes" (ce que traduit f submersive) ; il est donc raisonnable de passer à $p = n - (n-p)$ degrés de liberté.

On définit, comme pour les hypersurfaces, l'espace tangent à la sous-variété en chaque point. C'est un sous-espace vectoriel de $\mathbb{R}^n$ de dimension p, et les diverses caractérisations de cet espace tangent (en termes de redressement, d'équation et de graphe) sont encore d'actualité, avec la même démonstration. En particulier, lorsque $M = \{f = 0\}$ avec

$$f = (f_1, \cdots, f_{n-p}) : \mathbb{R}^n \to \mathbb{R}^{n-p}$$

de classe C^1 et submersive, on a en tout point $m \in M$

$$T_m M = \mathrm{Ker} D_m f = \bigcap_{i=1}^{n-p} \mathrm{Ker} D_m f_i\,,$$

sous-espace vectoriel de dimension $n-p$ de $\mathbb{R}^n$.

Théorème 10.16 Extrema liés
Soient $f = (f_1, \cdots, f_{n-p}) : \mathbb{R}^n \to \mathbb{R}^{n-p}$ une application de classe C^1 submersive, et $M := \{f = 0\}$ la sous-variété de dimension p correspondante. Soit $g : \mathbb{R}^n \to \mathbb{R}$ une fonction scalaire de classe C^1.

On suppose que la restriction $g_{|M} : M \to \mathbb{R}$ admet un extrema local au point $m_0 \in M$. Alors le point m_0 est un point critique relatif de g, c'est-à-dire que

— *pour tout vecteur $u \in T_{m_0}M$ on a $D_{m_0}g(u) = 0$*

— *ou, de façon équivalente, il existe des scalaires $\lambda_1, \cdots, \lambda_{n-p}$ tels que*

$$D_{m_0}g = \sum_{i=1}^{n-p} \lambda_i \, D_{m_0}f_i \, .$$

Les scalaires $(\lambda_1, \cdots, \lambda_{n-p})$ sont uniques ; ce sont les multiplicateurs de Lagrange.

Preuve La début de la preuve est le même que pour les hypersurfaces. On arrive à la condition $\mathrm{Ker} D_{m_0}f \subset \mathrm{Ker} D_{m_0}g$, ou encore

$$\bigcap_{i=1}^{n-p} \mathrm{Ker} D_{m_0}f_i \subset \mathrm{Ker} D_{m_0}g \, . \tag{10.1}$$

On introduit, pour toute partie $A \subset L(\mathbb{R}^n, \mathbb{R})$, son orthogonal qui est l'espace vectoriel

$$A^\top = \{x \in \mathbb{R}^n \mid \ell(x) = 0 \text{ pour toute } \ell \in A\} \subset \mathbb{R}^n \, .$$

On a

$$\bigcap_{i=1}^{n-p} \mathrm{Ker} D_{m_0}f_i = \bigcap_{i=1}^{n-p} (D_{m_0}f_i)^\top = \Big(\bigoplus_{i=1}^{n-p} \mathbb{R}\, D_{m_0}f_i\Big)^\top \, .$$

L'inclusion (10.1) se réécrit donc

$$\big(\oplus_{i=1}^{n-p} \mathbb{R}\, D_{m_0}f_i\big)^\top \subset (D_{m_0}g)^\top \, ,$$

ou, de façon équivalente,

$$D_{m_0}g \subset \oplus_{i=1}^{n-p} \mathbb{R}\, D_{m_0}f_i \, ,$$

ce qu'on voulait. L'unicité des multiplicateurs de Lagrange vient de ce que les formes linéaires $D_{m_0}f_i$ sont indépendantes, puisque f est submersive. $\square$

Exercice 10.17 Montrer que le groupe orthogonal $O(n)$ est une sous-variété de $M_n\mathbb{R} \simeq \mathbb{R}^{n^2}$. Déterminer sa dimension, et son espace tangent en l'identité.