Certificat S3 : TOPOLOGIE ET CALCUL DIFFERENTIEL (MATH 205)

B. Helffer (à partir de différents documents de G. David et autres collègues)
Département de Mathématiques
Université Paris-Sud

Version Septembre-Décembre 2012

Chapitre 1

Topologie

1.1 RAPPELS SUR LES RÉELS

On rappelle juste ce qui va nous intéresser spécialement pour parler bientôt de la topologie dans $\mathbb{R}^n$.

a. Bornes supérieures et inférieures

Définition 1.1.1
Soient $E \subset \mathbb{R}$ et $M \in \mathbb{R}$. On dit que M est un majorant de E si $x \leq M$ pour tout $x \in E$. On dit que M est un minorant de E si $x \geq M$ pour tout $x \in E$. On dit que M est le plus grand élément de E si $M \in E$ et M est un majorant de E. On dit que M est le plus petit élément de E si $M \in E$ et M est un minorant de E.

Noter que E (même non vide et majoré) n'a pas forcément de plus grand élément (prendre $E = [0, 1[$), mais que l'unicité est triviale.

Définition 1.1.2
On dit que $E \subset \mathbb{R}$ est majorée s'il existe un majorant $M \in \mathbb{R}$. Sinon, c'est que pour tout $M \in \mathbb{R}$ (ou tout entier $M \geq 0$), il existe $x \in E$ tel que $x > M$. On dit que $E \subset \mathbb{R}$ est minorée s'il existe un minorant $M \in \mathbb{R}$. Sinon, c'est que pour tout $M \in \mathbb{R}$ (ou tout entier $M \geq 0$), il existe $x \in E$ tel que $x < M$. On dit que E est bornée si elle est à la fois majorée et minorée.

Propriété fondamentale de $\mathbb{R}$ (Théorème ou définition, c'est selon) :
Si $E \subset \mathbb{R}$ est une partie non vide et majorée de $\mathbb{R}$, alors E admet une borne supérieure. C'est à dire, un nombre $M \in \mathbb{R}$ tel que :
• *$x \leq M$ pour tout $x \in E$;*

• *$m \geq M$ pour tout majorant de E.*

En bref, M est le plus petit des majorants de E. On le note par exemple $\sup_{x\in E} x$ ou encore mieux $\sup\{x\,;\, x \in E\}$ (notation qui a l'avantage d'être explicite et de se généraliser)

C'est pareil pour la borne inférieure $\inf_{x\in E} x$.

Remarques :

– la borne supérieure peut appartenir, ou pas, à l'ensemble. Si c'est le cas, l'ensemble a un plus grand élément. Contrexemple : $[0, 1[$.

– La propriété fondamentale est fausse sur $\mathbb{Q}$, par exemple, à cause de $E = \{x \in \mathbb{Q}\,; x^2 \leq 2\}$.

Convention acceptée : $\sup\{x\,;\, x \in E\} = +\infty$ si E n'est pas majoré. Mais on ne parlera pas de borne supérieure dans ce cas.

b. Distance, suites de Cauchy On en profite pour donner les définitions en termes de distance.

Définition 1.1.3

Soit E un ensemble. Une **distance** *sur E est une fonction* dist*, définie sur $E \times E$ et à valeurs dans $\mathbb{R}_+ = [0, +\infty[$, avec les trois propriétés suivantes :*

• $\mathrm{dist}(x, y) = \mathrm{dist}(y, x)$ *pour $x, y \in E$;*

• $\mathrm{dist}(x, y) = 0$ *si et seulement si $x = y$;*

• $\mathrm{dist}(x, z) \leq \mathrm{dist}(x, y) + \mathrm{dist}(y, z)$ *pour tout choix de $x, y, z \in E$ (inégalité triangulaire). Un* **espace métrique** *est un couple (E, dist), où E est un ensemble et* dist *est une distance sur E.*

On s'intéresse surtout à la distance usuelle sur $\mathbb{R}$, définie par $\mathrm{dist}(x, y) = |x - y|$. Noter que c'est facile d'en trouver d'autres, par exemple $\mathrm{dist}(x, y) = |e^x - e^y|$, ou $\mathrm{dist}(x, y) = \sqrt{|x - y|}$. Voici un exemple un peu plus pathologique : la distance discrète définie sur un ensemble E (par exemple $E = \mathbb{R}$) par $\mathrm{dist}(x, y) = 0$ si $x = y$ et $\mathrm{dist}(x, y) = 1$ si $x \neq y$.

Définition 1.1.4

Soient $\{x_n\}_{n\geq 0}$ (ou plus simplement $\{x_n\}$) une suite de nombres réels. Soit $l \in \mathbb{R}$. On dit que $\{x_n\}$ converge vers l (et on note $\lim_{n\to+\infty} x_n = l$) quand, pour tout réel $\varepsilon > 0$, il existe $N \geq 0$ tel que $\mathrm{dist}(x_n, l) \leq \varepsilon$ pour tout $n \geq N$.

Noter que cette définition marche aussi très bien pour des suites à valeurs dans un ensemble E, muni d'une distance dist (espace métrique).

Notons aussi une propriété de $\mathbb{R}$ essentiellement équivalente à celle de la borne supérieure :

Toute suite $\{x_n\}$ dans $\mathbb{R}$, qui est croissante et majorée, est convergente (c.à.d., a une limite $l \in \mathbb{R}$).
[Exercice : démontrer chacune à partir de l'autre.]

On a la même propriété pour les suites décroissantes minorées.

Définition 1.1.5
Une suite $\{x_n\}_{n\geq 0}$ est dite **de Cauchy** *quand, pour tout $\varepsilon > 0$, il existe $N \geq 0$ tel que $\operatorname{dist}(x_m, x_n) \leq \varepsilon$ pour tout choix de $m \geq N$ et $n \geq N$. [En gros, $\operatorname{dist}(x_m, x_n)$ tend vers 0 quand m et n tendent vers $+\infty$.]*

Noter que cette définition marche très bien dans un espace métrique. Noter aussi que toute suite convergente est une suite de Cauchy (le faire). Mais il n'est pas vrai en général, c'est que toute suite de Cauchy dans un espace métrique converge, comme dans ce qui suit.

Théorème 1.1.6
Toute suite de Cauchy dans $\mathbb{R}$ est convergente.

Notons que c'est faux dans $\mathbb{Q}$, ou dans $\mathbb{R}\setminus\{0\}$. Cette propriété caractérise les espaces métriques dits “complets”.

Démonstration.

D'abord, toute suite de Cauchy est bornée : On prend $\varepsilon = 1$ et on trouve $N \geq 0$ tel qu'en particulier $\operatorname{dist}(x_n, x_N) \leq 1$ pour $n \geq M$. Alors $|x_n| \leq M$ pour tout n, où l'on peut prendre $M = \operatorname{Max}(1 + |x_N|, |x_0|, |x_1|, \dots |x_{N-1}|)$.

Soit M tel que $|x_n| \leq M$ pour tout n. Puis, pour tout $k \geq 0$, on pose $s_k = \sup_{n\geq k} u_n$ (borne supérieure d'un ensemble). Noter que s_k est bien défini, et que $s_k \in [-M, M]$.

Ensuite, $\{s_k\}$ est une suite décroissante. Elle a une limite l (puisqu'elle est aussi minorée par $-M$). [en fait, l est la limite supérieure de la suite, et on vient de voir qu'elle est définie et finie si la suite est majorée.]

Il reste à voir que $\lim_{n\to+\infty} x_n = l$. On se donne $\varepsilon > 0$. On trouve N_1 tel que $|v_k - l| \leq \varepsilon/3$ pour $k \geq N_1$, et N_2 tel que $|x_m - x_n| \leq \varepsilon/3$ pour m et $n \geq N_1$.

Voyons si $|x_n - l| \leq \varepsilon$ pour $n \geq N = \operatorname{Max}(N_1, N_2)$.

D'abord, $x_n \leq s_n \leq l + \varepsilon/3$, juste par définition de s_n et de N_1. Reste à minorer x_n

Mais $s_n \geq l - \varepsilon/3$. C'est une borne supérieure, donc le plus petit des majorants, donc $s_n - \varepsilon/3$ n'est plus un majorant, donc il existe $m \geq n$ tel que $x_m > s_n - \varepsilon/3$. Donc $x_m > l - 2\varepsilon/3$. Mais $|x_m - x_n| \leq \varepsilon/3$, donc $x_n > l - \varepsilon$. Tout ça vaut pour tout $n \geq N$, donc $\lim_{n\to+\infty} x_n = l$ et le théorème est démontré. □

c. Suites extraites, Bolzano-Weierstrass

Définition 1.1.7
Une **suite extraite** *de la suite* $\{x_n\}$ *est une suite de la forme* $\{x_{\varphi(k)}\}$*, où* $\varphi : \mathbb{N} \to \mathbb{N}$ *est une application* **strictement** *croissante.*

Par exemple, la suite $\{x_{2k+3}\}$ est extraite de la suite $\{x_n\}$ (avec $\varphi(k) = 2k+1$).
Commentaire : Puisque φ est strictement croissante, on vérifie par récurrence que $\varphi(k) \geq k$ pour tout k. En particulier, $\lim_{k\to+\infty} \varphi(k) = +\infty$.
Remarque : Une suite extraite d'une suite extraite est une suite extraite. En effet, on part de $\{x_n\}$, puis on passe à $\{y_k\}$, avec $y_k = x_{\varphi(k)}$, puis à $\{z_l\}$, associée à l'application strictement croissante ψ, ce qui donne $z_l = y_{\psi(l)} = x_{\varphi(\psi(l))}$. Donc, $\{z_l\}$ est la suite extraite de $\{x_n\}$ associée à l'application strictement croissante $\varphi \circ \psi$.

Proposition 1.1.8
Si la suite $\{x_n\}$ *converge vers* l*, alors toute suite extraite* $\{x_{\varphi(k)}\}$ *converge aussi vers* l*.*

En fait, $\{x_{\varphi(k)}\}$ converge vers l dès que $\varphi : \mathbb{N} \to \mathbb{N}$ tend vers $+\infty$ à l'infini. La démonstration est standard par composition des limites.

La réciproque est fausse, comme le montre la suite de terme général $x_n = (-1)^n$.

Théorème 1.1.9 *[(Bolzano-Weierstrass).]*
Si $\{x_n\}$ *est une suite bornée à valeurs dans* $\mathbb{R}$*, alors il existe une suite extraite* $\{x_{\varphi(k)}\}$ *qui converge.*

Ce théorème est très important ; la démonstration vient plus tard.

Définition 1.1.10
Soit $\{x_n\}$ *une suite à valeurs dans* $\mathbb{R}$*. Une* **valeur d'adhérence** *de* $\{x_n\}$ *est un nombre* l *tel que, pour tout* $\varepsilon > 0$*, il existe une infinité de valeurs de* n *telles que* $\mathrm{dist}(x_n, l) < \varepsilon$*.*

Ou encore, pour tout $\varepsilon > 0$ et tout $N > 0$, il existe $n \geq N$ tel que $\mathrm{dist}(x_n, l) < \varepsilon$. [Expliqué à l'oral]

Cette définition marche plus généralement dans un espace métrique.
Exemples :
• Pour $x_n = (-1)^n$, il y a deux valeurs d'adhérence (1 et -1).

• Si la suite $\{x_n\}$ converge vers $\ell \in \mathbb{R}$, ℓ est valeur d'adhérence. [Comparez les définitions.]

Une **convention acceptée** est de dire que $+\infty$ est valeur d'adhérence si la suite n'est pas majorée. [Penser : pour tout voisinage V de $+\infty$, on a $x_n \in V$ pour une infinité de valeurs de n.]

Noter que la propriété d'avoir une valeur d'adhérence estune propriété de la suite, et pas seulement de l'ensemble des valeurs prises par la suite.

Une étape importante pour la démonstration de Bolzano-Weierstrass est le théorème suivant :

Théorème 1.1.11
Toute suite bornée dans $\mathbb{R}$ a au moins une valeur d'adhérence.

Démonstration :
Soit M tel que $x_n \in [-M, M]$ pour tout n (on a utilisé l'hypothèse du théorème). On va définir une suite d'intervalles fermés emboîtés I_k. Chacun sera de longueur $2^{-k}(2M)$, et contiendra x_n pour une infinité de valeurs de n.

On commence par $I_0 = [-M, M]$. Supposons construit I_k, avec les propriétés ci-dessus. On note c_k le centre de I_k, et on coupe I_k en deux morceaux (à gauche et à droite de c_n), à savoir $I_k^- = \{x \in I_k\,;\, x \leq c_k\}$ et $I_k^+ = \{x \in I_k\,;\, x \geq c_k\}$. Si I_k^- contient u_n pour une infinité de valeurs de n, on prend $I_{k+1} = I_k^-$. Sinon, on prend $I_{k+1} = I_k^+$. Dans ce cas aussi, I_{k+1} contient u_n pour une infinité de valeurs de n, parce que c'est le cas pour $I_k = I_k^- \cup I_k^+$, et que seul un nombre fini atterrit dans I_k^-. Ceci termine la construction de I_k par récurrence.

Vérifions maintenant que la suite $\{c_k\}$ des centres converge. Il suffit de voir que c'est une suite de Cauchy. Soit donc $\varepsilon > 0$. Choisissons $N \geq 0$ tel que $2^{-N+1}M < \varepsilon$. Alors pour tout choix de $m \geq N$ et $n \geq N$, $c_m \in I_m \subset I_N$ (parce que les intervalles sont emboîtés et $m \geq N$), et pareillement $c_n \in I_n \subset I_N$. Donc $|x_m - x_n| \leq \text{longueur}(I_N) = 2^{-N+1}M < \varepsilon$, comme souhaité.

Donc $\{c_k\}$ est une suite de Cauchy, et converge. Soit c la limite. On montrerait assez facilement que l'intersection des I_k est exactement $\{c\}$, mais on n'en a pas besoin.

Vérifions juste que c est une valeur d'adhérence de la suite $\{x_k\}$. Soit donc $\varepsilon > 0$, et vérifions que $|x_n - c| < \varepsilon$ pour une infinité de valeurs de n.

Choisissons N tel que $2^{-N+1}M < \varepsilon/2$ d'une part, et $|c_n - c| \leq \varepsilon/2$ pour $n \geq N$ d'autre part. Rappelons que $x_n \in I_N$ pour une infinité de valeurs de n. Et pour ces valeurs, $|x_n - c_n| \leq \varepsilon/2$ puisque les deux points sont dans I_N, et $|c_n - c| \leq \varepsilon/2$ par choix de N, donc $|x_n - c| \leq \varepsilon$, comme souhaité. $\square$

Une observation bien plus simple, et qui marche dans n'importe quel espace métrique, se traduit par la proposition suivante :

Proposition 1.1.12
Si l est valeur d'adhérence de la suite $\{x_n\}$, alors il existe une suite extraite $\{x_{\varphi(k)}\}$ qui converge vers l.

Remarque : Si on ne prend pas une sous-suite, c'est faux (penser à la suite $(-1)^n$).

Démonstration de la proposition.
On va définir $\varphi(k)$ par récurrence sur k. On part de $\varphi(0) = 0$. Soit $k \geq 0$, supposons $\varphi(k)$ défini, et trouvons $\varphi(k+1)$.
Appliquons la définition avec $\varepsilon = 2^{-k-1}$; on trouve que $|x_n - l| < 2^{-k-1}$ pour une infinité de valeurs de n. En particulier, on peut trouver $n > \varphi(k)$ tel que $|x_n - l| < 2^{-k-1}$. On prend cet n pour $\varphi(k+1)$.

Clairement, ceci nous donne une fonction φ strictement croissante, et comme on s'est assuré que $|x_{\varphi(k+1)} - l| < 2^{-k-1}$ pour tout k, on a bien $\lim_{k\to+\infty} x_{\varphi(k)} = l$. $\square$

Maintenant le théorème de Bolzano-Weierstrass est une conséquence immédiate de la proposition et du dernier théorème.

Notons aussi que pour qu'il existe une suite extraite de $\{x_n\}$ qui converge vers l, il faut que l soit valeur d'adhérence de la suite. C'est assez facile à vérifier (il faut que l soit valeur d'adhérence de la suite extraite, donc aussi (en vérifiant les définitions) de la suite de départ) et donne une réciproque au théorème ci-dessus, mais c'est aussi assez peu utile en pratique.

Le théorème de Bolzano-Weierstrass joue un rôle fondamental en analyse. Voici une application parmi beaucoup.

Théorème 1.1.13
Soient I un intervalle fermé borné de $\mathbb{R}$ (on dit aussi un intervalle compact, voir la suite du cours) et $f : I \to \mathbb{R}$ une application continue. Alors f est bornée et atteint ses bornes. Autrement dit, $m = \inf\{f(x)\,;\, x \in I\}$ et $M = \sup\{f(x)\,;\, x \in I\}$ sont des réels, et il existe $\alpha \in I$ et $\beta \in I$ tels que $f(a) = m$ et $f(b) = M$.

Démonstration pour la borne supérieure.
Si $E = \{f(x)\,;\, x \in I\}$ n'est pas majoré, on peut, pour tout $n \geq 0$, trouver $y_n \in E$ tel que $y_n > n$, donc aussi $x_n \in I$ tel que $y_n = f(x_n) > n$. La suite $\{x_n\}$ est bornée (puisque I est bornée). Par Bolzano-Weierstrass, on peut en extraire une sous-suite (= une suite extraite) $\{x_{\varphi(k)}\}$ qui converge vers une

limite l. Et $l \in I$, parce que I est fermé. Plus précisément, si $I = [a, b]$, on a que $a \leq x_{\varphi(k)} \leq b$ pour tout k, donc $a \leq l \leq b$ en passant à la limite. Et alors

$$(*) \qquad f(l) = \lim_{k \to +\infty} f(x_{\varphi(k)})$$

par continuité de f en l. C'est impossible parce que $f(l)$ est fini, mais la limite à droite est $+\infty$ (car $f(x_{\varphi(k)}) = y_{\varphi(k)} \geq \varphi(k) \geq k$).

Donc E est majoré, et a une borne supérieure $M \in \mathbb{R}$. Par définition de M, $f(x) \leq M$ pour tout $x \in I$. Mais aussi, pour tout $n \geq 0$, il existe $y_n \in E$ tel que $y_n > M - 2^{-n}$, donc aussi un $x_n \in I$ tel que $f(x_n) = y$. Donc $M - 2^{-n} < f(x_n) \leq M$.

A nouveau et comme avant, la suite $\{x_n\}$ est bornée donc Bolzano-Weierstrass donne une sous-suite $\{x_{\varphi(k)}\}$ qui converge vers une limite l, et $l \in I$ car I est fermé. De plus, la continuité de f en l donne encore $(*)$.

Mais ici, $M - 2^{-n} \leq f(x_n) \leq M$ pour tout n, donc

$$M - 2^{-\varphi(k)} \leq f(x_{\varphi(k)}) \leq M$$

pour tout k. Comme $\lim_{k \to +\infty} 2^{-\varphi(k)} = 0$, on trouve

$$\lim_{k \to +\infty} f(x_{\varphi(k)}) = M$$

en passant à la limite, et $(*)$ donne $f(l) = M$. Bref, on peut prendre $\beta = l$.

Le cas de la borne inférieure m est traité pareillement. $\square$

1.2 TOPOLOGIES

C'est le moyen général de s'intéresser à la notion de convergence. On commence par l'exemple le plus courant, qu'il conviendra de garder à l'esprit dans la suite. Nous introduisons le langage général, mais au final nous ne travaillerons que dans l'espace $\mathbb{R}^n$.

a) Ouverts et fermés dans un espace métrique
On se donne un espace métrique (E, dist) (voir la définition 1.1.3). L'exemple qu'on utilisera le plus est celui de $\mathbb{R}^n$, muni par exemple de la distance euclidienne $\mathrm{dist}(x,y) = \{\sum_{i=1}^n |x_i - y_i|^2\}^{1/2}$, où bien sûr on a noté $x = (x_1, \ldots, x_n)$ et pareil pour y. [On verra d'autres exemples, qui d'ailleurs peuvent donner les mêmes ouverts et fermés.]

Un autre exemple important est celui d'un sous-ensemble E de $\mathbb{R}^n$, toujours muni de (la restriction à E de) la distance euclidienne.

Notations : Pour $x \in E$ et $r > 0$, $B(x,r)$ est la boule ouverte centrée en x et de rayon r, à savoir $B(x,r) = \{y \in E\,;\, \mathrm{dist}(x,y) < r\}$; pareillement $\overline{B}(x,r)$ est la boule fermée centrée en x et de rayon r, à savoir

$$\overline{B}(x,r) = \{y \in E\,;\, \mathrm{dist}(x,y) \le r\}$$

(et on peut tolérer $r = 0$ si vous voulez).

Définition 1.2.1
Soit donc (E, dist) un espace métrique. Un **ouvert** *de E est un sous-ensemble O de E tel que, pour tout $x \in O$, il existe $r > 0$ tel que $B(x,r) \subset O$. Un* **fermé** *de E est un sous-ensemble F de E tel que $E \backslash O$ (ici le complémentaire de F dans E) est ouvert, c'est à dire, tel que pour tout $x \notin F$, il existe $r > 0$ tel que $B(x,r)$ ne rencontre pas F.*

En général, on note $A \setminus B = \{x \in A\,;\, x \notin B\}$. Parfois le complémentaire de A dans E est noté A^c (si, à cause du contexte, il est absolument clair que l'on sait qui est l'ensemble E).

Exemples :
- $B(x_0, r_0)$ est toujours ouvert, comme le montre l'inégalité triangulaire. [Si $x \in B(x_0, r_0)$, alors la boule $B(x,r)$ est contenue dans $B(x_0, r_0)$, où on a choisi $r = r_0 - \mathrm{dist}(x, x_0) > 0$.
- De même, $\overline{B}(x,r)$ est fermé (par l'inégalité triangulaire à nouveau, qui montre que si $x \notin \overline{B}(x,r)$, alors $B(x,r)$ est contenu dans le complémentaire de $\overline{B}(x,r)$, où l'on a pris $r = \mathrm{dist}(x, x_0) - r_0 > 0$.
  Malheureusement, la notion peut prêter à confusion, parce que dans certains espaces métriques, il peut se faire que l'adhérence de $B(x,r)$

(voir plus bas) soit strictement contenue dans la boule fermée, alors que la notation est pareille. [prendre $B(0,2)$ dans $E_0 = [0,1] \cup [2,3]$ (contenu dans $\mathbb{R}$).]
- Pour la même raison que pour une boule fermée, tout singleton $\{x\}$ dans un espace métrique est fermé.
- Une droite, dans le plan euclidien, est fermée.
- Un sous-ensemble $A \subset E$ peut très bien n'être ni ouvert ni fermé (par exemple $[0,1[$ dans $\mathbb{R}$), ou les deux à la fois (par exemple $B(0,3/2)$ dans l'exemple précédent avec $E = E_0$).
- Quand dist est la distance discrète sur E, tout sous-ensemble est à la fois ouvert et fermé.

Proposition 1.2.2
Soit (E, dist) un espace métrique.
- *Une union quelconque d'ouverts de E est un ouvert de E.*
- *Une intersection finie d'ouverts de E est un ouvert de E.*
- *Une intersection quelconque de fermés de E est un fermé de E.*
- *Une union finie de fermés de E est un fermé de E.*

Démonstration.
Pour une union U d'ouverts O_i : soit $x \in U$. Soit i tel que $x \in O_i$. Il existe $r > 0$ tel que $B(x,r) \subset O_i$. Alors $B(x,r) \subset U$, □.

Pour une intersection finie $O = \cap_{1 \leq i \leq m} O_i$, soit $x \in O$. Comme O_i est ouvert, il existe $r_i > 0$ tel que $B(x_i, r_i) \subset O_i$. On prend pour r le plus petit des r_i ; alors $B(x,r) \subset O_i$ pour tout i, donc $B(x,r) \subset O$, □.

Notez que $\cap_{n \in \mathbb{N}^*}]-\frac{1}{n}, \frac{1}{n}[= \{0\}$. C'est un exemple simple d'une intersection infinie d'ouverts qui n'est pas ouverte.

Le cas des fermés se fait par passage au complémentaire, ou directement par le même genre d'argument. □

Pour bien des choses, le résultat de la proposition est ce qu'il nous faut, par exemple, pour parler de convergence. C'est pourquoi on lui attache une notion (la notion de topologie), comme ci-dessous.

b) Topologie sur un ensemble

Définition 1.2.3
Soit E un ensemble. Une topologie sur E est la donnée d'une famille $\mathcal{T}$ de sous-ensembles de E (la liste des ensembles qu'on déclare ouverts), avec les propriétés suivantes :

• *$E \in \mathcal{T}$ et $\emptyset \in \mathcal{T}$;*
• *toute union d'ensembles de $\mathcal{T}$ est dans $\mathcal{T}$;*
• *toute intersection finie d'ensembles de $\mathcal{T}$ est dans $\mathcal{T}$.*
Un espace topologique est juste un couple $(E, \mathcal{T})$, où E est un ensemble et $\mathcal{T}$ est une topologie sur E.

Exemple majeur :
C'est celui des ensembles ouverts dans un espace métrique (et en particulier dans $\mathbb{R}^n$ ou un sous-ensemble de $\mathbb{R}^n$, muni de la distance euclidienne). D'ailleurs, le plus souvent, quand on travaillera sur $\mathbb{R}^n$, on utilisera cette topologie-là, sans même se donner la peine de préciser.
Voici quelques exemples plus académiques (mais utiles pour les contre-exemples) : la topologie discrète où $\mathcal{T} = \mathcal{P}(E)$ (tout ensemble est ouvert), et la topologie grossière où $\mathcal{T} = \{\emptyset, E\}$ (le minimum syndical).
Dans ce qui suit, on donne les définitions logiques, et on essaie de vérifier un peu au fur et à mesure que c'est bien cohérent avec ce que vous savez de la convergence sur $\mathbb{R}$. On se donne donc maintenant un espace topologique $(E, \mathcal{T})$.

Définition 1.2.4
Soit $x \in E$. Un voisinage de x est un ensemble $V \subset E$, qui contient un ouvert qui contient x.

Dans un espace métrique, cela veut juste dire qu'il existe $r > 0$ tel que $B(x, r) \subset V$. Cela peut paraître un peu ridicule, mais c'est quand même bien utile. Par exemple, $]-1, 1]$ est un voisinage de 0 (dans $\mathbb{R}$, muni de la distance usuelle), et on se moque souvent qu'il ne soit pas fermé ou pas ouvert.

Comme dans le cas métrique on a la définition :

Définition 1.2.5
Un ensemble fermé est un ensemble $F \subset E$ tel que $E \setminus F$ soit ouvert.

Notons que $\emptyset$ et E sont aussi fermés. Toute union finie de fermés est un fermé, et toute intersection de fermés est un fermé. On se sert aussi souvent de la notion suivante.

Définition 1.2.6 *(topologie restreinte).*
Soient $(E, \mathcal{T})$ un espace topologique, et $A \subset E$ (pas forcément fermé ou ouvert). La topologie restreinte (de $\mathcal{T}$ à A) est la topologie $\mathcal{T}' = \{O' \subset A\,;\ \text{il existe } O \in \mathcal{T} \text{ tel que } O' = A \cap O\}$.

On vérifie facilement que c'est bien une topologie, et que dans le cas où $\mathcal{T}$ vient d'une distance sur E, alors $\mathcal{T}'$ vient de cette même distance (ou pour couper les cheveux en quatre, de sa restriction à A).

Notons aussi (exercice), que les fermés de A pour la topologie restreinte sont les ensembles de la forme $A \cap G$, où G est un fermé de E.

Passons à la convergence des suites.

Définition 1.2.7
Soit $\{x_n\}$, $n \geq 0$ une suite à valeurs dans E. Soit $\ell \in E$. On dit que la suite $\{x_n\}$ converge vers ℓ, ou que $\lim_{n\to+\infty} x_n = \ell$, quand pour tout voisinage V de ℓ (dans E), il existe $N \geq 0$ tel que $x_n \in V$ pour tout $n \geq N$.

Proposition 1.2.8
Dans un espace métrique (c.à.d., si la topologie sur E est celle qui vient d'une distance dist*), $\lim_{n\to+\infty} x_n = \ell$ si et seulement si pour tout $r > 0$, il existe $N \geq 0$ tel que $x_n \in B(\ell, r)$ pour tout $n \geq N$.*

Autrement dit, il suffit de vérifier la condition pour des voisinages particuliers de ℓ, à savoir les boules ouvertes centrées en ℓ. La vérification est facile. Un sens est vraiment trivial, et voici l'autre. Soit V un voisinage de ℓ. Il existe un ouvert $O \subset V$ contient ℓ, donc aussi $r > 0$ tel que $B(\ell, r) \subset O \subset V$. On choisit N comme dans la proposition (et avec le choix de r), et N marche aussi pour V. $\square$

Notez donc que notre nouvelle définition de convergence d'une suite coïncide avec l'ancienne dans le cas de suites à valeurs réelles (ou à valeurs dans $\mathbb{C}$, muni de la distance $\mathrm{dist}(z_1, z_2) = |z_1 - z_2|$), ou à valeurs dans $\mathbb{R}^n$, si vous vous en souvenez ou comme on le verra plus loin).

La proposition suivante est du même type que les résultats standards de passage à la limite, et est donc fort utile.

Proposition 1.2.9
Soit $(E, \mathcal{T})$ un espace topologique, F un fermé de E, et $\{x_n\}$ une suite à valeurs dans F. Si la suite $\{x_n\}$ converge vers une limite ℓ, alors $\ell \in F$.

Démonstration.
Sinon, ℓ est dans l'ouvert $E \setminus F$, et il existe un voisinage V de x qui est contenu dans $E \setminus F$, c.-à.-d. qui ne rencontre pas F. Mais c'est impossible, parce que par définition d'une limite, $x_n \in V$ pour n assez grand, et que $x_n \in F$ par hypothèse. $\square$

c) Fonctions continues

Définition 1.2.10
Soient $(E_1, \mathcal{T}_1)$ *et* $(E_2, \mathcal{T}_2)$ *deux espaces topologiques,* $f : E_1 \to E_2$ *une application, et* $x \in E_1$*. On dit que* f *est continue en* x *quand, pour tout voisinage* W *de* $f(x)$ *dans* E_2*, il existe un voisinage* V *de* x *dans* E_1 *tel que* $f(V) \subset W$*.*

Noter la similitude avec la définition de convergence d'une suite, que l'on peut voir comme application de $\mathbb{N}$ dans E.

Exercice.
Donner la définition de " $\lim_{x \to x_0} f(x) = \ell$ ".

Proposition 1.2.11
Soient $(E_1, \mathcal{T}_1)$ *et* $(E_2, \mathcal{T}_2)$ *deux espaces topologiques,* $f : E_1 \to E_2$ *une application, et* $x \in E_1$*. Pour que* f *soit continue en* x*, il suffit que pour tout ouvert* W *de* E_2 *qui contient* $f(x)$*, il existe un voisinage* V *de* x *dans* E_1 *tel que* $f(V) \subset W$*. Si* $\mathcal{T}_2$ *vient d'une distance* dist *sur* E_2*, il suffit même que, pour tout* $r > 0$*, il existe un voisinage* V *de* x *dans* E_1 *tel que* $f(V) \subset B(f(x), r)$*.*

La démonstration est la même que pour la proposition précédente. La réciproque est triviale. Et d'ailleurs on peut prendre V ouvert, ou même une boule centrée en x. □

Vous connaissez déjà plein de fonctions continues, typiquement définies sur des intervalles et à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$. [Les fonctions polynômes, les fonctions dérivables, $x^7 \sin(1/x)$, etc.] Inutile de faire semblant de ne pas savoir.

Exemple utile : toute fonction lipschitzienne est continue.
On rappelle que si (E_1, d_1) et (E_2, d_2) sont des espaces métriques, on dit que la fonction $f : E_1 \to E_2$ est lipschitzienne quand il existe $C \geq 0$ telle que

$$(*) \qquad d_2(f(x), f(y)) \leq C\, d_1(x, y) \quad \text{pour tout choix de } x \in E \text{ et } y \in E.$$

La démonstration est facile : fixons $x \in E_1$ et voyons la continuité en x. On se donne un voisinage V de $f(x)$ dans E_2. Par définition, il existe $r_2 > 0$ tel que $B(f(x), r_2) \subset V$ (boule dans E_2). Prenons $r_1 = r_2/C$ si $C > 0$, et $r_1 = 1$ sinon [1], et vérifions que $f(B(x, r_1)) \subset B(f(x), r_2) \subset V$; on en déduira le résultat. Mais si $y \in B(x, r_1)$, alors $d_2(f(x), f(y)) \leq C d_1(x, y) < C r_1 \leq r_2$, ce qu'on voulait. □

1. N'importe que réel strictement positif convient.

Corollaire
Si (E, dist) *est un espace métrique, alors pour tout* $y \in E$*, la fonction* $x \to \mathrm{dist}(x, y)$*, de* E *dans* $\mathbb{R}_+$*, est continue.*

En effet, cette fonction est lipschitzienne avec la constante $C = 1$, par l'inégalité triangulaire : pour $x, x' \in E$, $\mathrm{dist}(x, y) - \mathrm{dist}(x', y) \leq \mathrm{dist}(x, x')$, parce que $\mathrm{dist}(x, y) \leq \mathrm{dist}(x, x') + \mathrm{dist}(x', y)$, et pour l'autre inégalité, il suffit d'échanger x et x'. □

En fait on a un peu mieux : la fonction $(x, y) \to \mathrm{dist}(x, y)$, de $E \times E$ dans $\mathbb{R}_+$, est continue. Pour cet énoncé, il convient de mettre une distance sur $E \times E$. Par exemple, on peut choisir

$$\mathrm{dist}((x_1, x_2), (y_1, y_2)) = \mathrm{dist}(x_1, y_1) + \mathrm{dist}(x_2, y_2)\,.$$

Mais on aurait aussi pu prendre

$$\mathrm{dist}((x_1, x_2), (y_1, y_2)) = \mathrm{Max}(\,\mathrm{dist}(x_1, y_1), \mathrm{dist}(x_2, y_2))\,,$$

et obtenir le même résultat. La démonstration est facile et laissée en exercice.

Exercice.
Démontrer que toute fonction höldérienne $f : E_1 \to E_2$ est continue. Höldérienne signifie qu'il existe un exposant $\alpha > 0$ et une constante $C \geq 0$ telle que

$$(**) \qquad d_2(f(x), f(y)) \leq C\, d_1(x, y)^\alpha \quad \text{pour tout choix de } x \in E \text{ et } y \in E.$$

On dit que $f : E_1 \to E_2$ est continue si elle est continue en tout point $x \in E_1$. Le critère suivant est facile, *mais montre bien que les fonctions continues sont les morphismes d'espaces topologiques (en un sens qui ne sera pas précisé).*

Proposition 1.2.12
Soient $(E_1, \mathcal{T}_1)$ *et* $(E_2, \mathcal{T}_2)$ *deux espaces topologiques,* $f : E_1 \to E_2$ *une application. Les trois propriétés suivantes sont équivalentes :*
1. f *est continue*
2. $f^{-1}(U)$ *est un ouvert (de* E_1*) pour tout ouvert* U *de* E_2
3. $f^{-1}(F)$ *est un fermé (de* E_1*) pour tout fermé* F *de* E_2*.*

La démonstration est facile (chasse aux définitions).
Supposons d'abord f continue, et montrons **2.**. Soit U un ouvert de E_2. Soit $V = f^{-1}(U) = \{x \in E_1\,;\, f(x) \in U\}$. Il suffit de montrer que pour tout $x \in V$,

il existe un voisinage W de x qui est contenu dans V. Mais $f(x) \in U$, qui est ouvert, donc est un voisinage de $f(x)$. La définition de continuité donne un voisinage W de x tel que $f(W) \subset U$. Alors $W \subset V$, par définition de V, $\square$.

Dans l'autre sens, si on a **2.** et si $x \in E_1$, f est continue en x parce que pour tout voisinage W de $f(x)$ dans E_2, W contient un ouvert U qui contient $f(x)$, alors $V = f^{-1}(U)$ est un ouvert qui contient x, donc c'est un voisinage de x dans E_1, et $f(V) \subset U$ parce que $V = f^{-1}(U)$. Donc aussi $f(V) \subset W$, comme demandé.

L'équivalence de **2.** et **3.** est facile. Par exemple, si on suppose **2.** et si F est fermé, on note $U = E_2 \setminus F$, et on note que $f^{-1}(F) = \{x \in E_1\,;\, f(x) \in F\} = \{x \in E_1\,;\, f(x) \notin U\}$ est le complémentaire de $f^{-1}(U)$, qui est ouvert. On fait de même pour "**3.** implique **2.**". $\square$

Deux mots sur la composition.

Proposition 1.2.13
Soient $(E_1, \mathcal{T}_1)$, $(E_2, \mathcal{T}_2)$, $(E_3, \mathcal{T}_3)$ des espaces topologiques, $f : E_1 \to E_2$ et $g : E_2 \to E_3$ deux applications.
- *Si f est continue en x et g est continue au point $f(x)$, alors $g \circ f$ est continue au point x.*
- *Si f et g sont continues, alors $g \circ f$ est continue.*

C'est facile (vérification laissée en exercice). Donnons juste la démonstration de la seconde partie (qui peut aussi être déduite de la première).
Soit U_3 un ouvert de E_3. Alors $U_2 = g^{-1}(U_3)$ est ouvert si g est continue, et donc $U_1 = f^{-1}(U_2)$ est aussi ouvert si f est continue. Mais

$$U_1 = \{x \in E_1\,;\, f(x) \in U_2\} = \{x \in E_1\,;\, g(f(x)) \in U_3\} = (g \circ f)^{-1}(U_3)\,.$$

C'est donc bien que $(g \circ f)^{-1}(U_3)$ est ouvert pour tout ouvert U_3 de E_3. $\square$

d) Intérieur et adhérence

Encore un peu de vocabulaire. Dans la suite on se donne un espace topologique $(E, \mathcal{T})$ et une partie quelconque A de E.

Définition 1.2.14
L'intérieur de $A \subset E$, noté $\mathring{A}$, est l'union de tous les ouverts contenus dans A.

C'est donc un ouvert, et même le plus grand ouvert contenu dans A.

La vérification est facile.
Il peut arriver que $\mathring{A} = \emptyset$, même si A lui-même n'est pas vide (on dit alors que A est d'intérieur vide). Par exemple $A = \{x\}$ est d'intérieur vide dans $\mathbb{R}$.

On utilisera dans la suite la caractérisation facile de l'intérieur de A.

Proposition 1.2.15
L'intérieur de A est l'ensemble des $x \in E$ tels qu'il existe un voisinage de x contenu dans A.

En effet, si $x \in \mathring{A}$, $\mathring{A}$ est un voisinage ouvert de x contenu dans A ; réciproquement, si x possède un voisinage contenu dans A, ce voisinage contient un ouvert O qui contient x, et alors $O \subset \mathring{A}$ par définition de $\mathring{A}$, donc $x \in \mathring{A}$. □

Passons au complémentaire.

Définition 1.2.16
L'adhérence de A, en général notée $\overline{A}$, est l'intersection de tous les fermés qui contiennent A.

C'est donc un fermé, et c'est en fait le plus petit fermé qui contient A. La vérification de cette dernière affirmation est facile.
Parfois on dit aussi la fermeture de A. Le mot adhérence est un peu malheureux, parce que la notion n'est pas exactement la même que celle des points d'adhérence d'une suite, mais tant pis : c'est le vocabulaire utilisé par tout le monde.

Noter que le complémentaire de l'intérieur de A est l'adhérence du complémentaire. [C'est le complémentaire de l'union des ouverts contenus dans A ; c'est donc l'intersection des complémentaires de ces ouverts, qui sont justement les fermés qui contiennent $E \setminus A$].

Quand $\overline{A} = E$, on dit que A est dense.
Par exemple, $\mathbb{Q}$ est dense dans $\mathbb{R}$.

On utilisera aussi la caractérisation suivante par les voisinages, obtenue à partir de celle de la proposition 1.2.15.

Proposition 1.2.17
Pour tout sous-ensemble A de E, $\overline{A}$ est l'ensemble des $x \in E$ tels que tout voisinage de x rencontre A.

En effet, soit $x \in E$. Alors $x \in \overline{A}$ si et seulement si x n'est pas dans l'intérieur de A^c, ce qui signifie qu'il n'y a pas de voisinage de x qui est contenu dans A^c, ou encore que pour tout voisinage V de x, V n'est pas contenu dans dans A^c, ou en langage plus clair, V rencontre A.

Exercice :
Chercher $A \subset \mathbb{R}$ tel que $\overline{A} \neq \overline{\mathring{A}}$ et $B \subset \mathbb{R}$ tel que $\mathring{B} \neq (\mathring{\overline{B}})$.

Définition 1.2.18
La frontière de A est l'ensemble $\partial A = \overline{A} \setminus \mathring{A}$.

C'est donc l'ensemble des points $x \in E$ tels que tout voisinage de x rencontre à la fois A et son complémentaire.

C'est un fermé. Noter que $\partial A = \overline{A} \cap \overline{A^c}$. Donc aussi $\partial A = \partial(A^c)$. Vous pouvez tracer quelques exemples dans le plan pour vérifier qu'on a eu raison d'appeler ceci frontière.

e) Utilisation des suites dans les espaces métriques

L'avantage des espaces métriques par rapport à des espaces topologiques généraux (dont on ne parlera pas ici), c'est que beaucoup de choses peuvent être décidées avec des suites. Voici juste deux exemples ci-dessous ; la propriété de Bolzano-Weierstrass de la prochaine section pour décider de la compacité en sera un autre.

Proposition 1.2.19
Soient (E, dist) un espace métrique et $A \subset E$. L'adhérence de A est l'ensemble des limites de suites à valeurs dans A.

Autrement dit, pour $x \in E$, $x \in \overline{A}$ si et seulement s'il existe une suite $\{x_n\}$ à valeurs dans A, telle que $\lim_{n\to+\infty} x_n = x$.

On sait déjà que la limite d'une suite à valeurs dans $\overline{A}$ est dans $\overline{A}$, par la proposition 1.2.9 et parce que $\overline{A}$ est fermé. Cela donne la partie facile.

Réciproquement, soit $x \in \overline{A}$; pour tout n, $B(x, 2^{-n})$ est un voisinage de x, donc rencontre A. Choisissons $x_n \in A \cap B(x, 2^{-n})$. Alors $\lim_{n\to+\infty} x_n = x$, comme souhaité. $\square$

Proposition 1.2.20
Soient (E_1, dist) un espace métrique, $(E_2, \mathcal{T})$ un espace topologique, $f : E_1 \to E_2$ une application, et $x \in E_1$. Alors f est continue au point x si et seulement si $\lim_{n\to+\infty} f(x_n) = f(x)$ pour toute suite $\{x_n\}$ telle que $\lim_{n\to+\infty} x_n = x$.

Démonstration.
Si f est continue en x et $\{x_n\}$ tend vers x, alors

$$\lim_{n\to+\infty} f(x_n) = f(x)$$

par composition de limites. C'est donc la réciproque qui est intéressante.

Supposons que f n'est pas continue en x. Il y a donc un voisinage V de $f(x)$ tel que l'on n'ait $f(W) \subset V$ pour aucun voisinage W de x. Pour tout $n \geq 0$, on essaie $W = B(x, 2^{-n})$. Comme $f(W)$ n'est pas contenu dans V, il existe $x_n \in B(x, 2^{-n})$ tel que $f(x_n) \notin V$. La suite $\{x_n\}$ converge bien vers x, mais $f(x_n)$ reste hors de V, donc ne tend pas vers $f(x)$, donc la propriété sur les suites est violée aussi. $\square$

Exercice :
Énoncer un critère semblable sur l'existence de $\lim_{y \to x} f(y)$ sous les hypothèses du théorème.

Conclusion de la section :
Tout ceci est surtout (pour l'instant) un vocabulaire pratique pour parler de convergence ; évidemment, il y a beaucoup d'exemples intéressants en dehors de $\mathbb{R}^n$, et de notions complémentaires utiles, mais nous allons nous concentrer dans ce cours sur $\mathbb{R}^n$ et exploiter les résultats déjà connus sur $\mathbb{R}$ (par exemple, le fait que la somme ou le produit de deux fonctions continues est continu(e)).

1.3 TOPOLOGIE DE $\mathbb{R}^n$.

On a déjà vu des choses sur la topologie de $\mathbb{R}$, et on va s'en servir pour étudier $\mathbb{R}^n$.

1.3.1 Topologie de $\mathbb{R}^n$

On munit $\mathbb{R}^n$ de la distance euclidienne usuelle

$$\mathrm{dist}(x,y) = \{\sum_{i=1}^{n} |x_i - y_i|^2\}^{1/2},$$

où on a noté $x = (x_1, \ldots, x_n)$ et $y = (y_1, \ldots, y_n)$. Cela donne aussi une topologie sur $\mathbb{R}^n$. Vérifions tout de suite que l'on aurait pu choisir d'autres distances, sans changer la topologie.

Définition 1.3.1
Soit E un espace vectoriel sur $\mathbb{R}$. Une **norme** *sur E est une application $N : E \to [0, +\infty[$, ayant les propriétés suivantes :*

$$N(\lambda x) = |\lambda| N(x) \quad \textit{pour } x \in E \textit{ et } \lambda \in \mathbb{R}, \tag{1.1}$$

$$N(x+y) \leq N(x) + N(y) \quad \textit{pour tous } x, y \in E, \tag{1.2}$$

(l'inégalité triangulaire), et (pour $x \in E$)

$$N(x) = 0 \textit{ si et seulement si } x = 0. \tag{1.3}$$

Voici quelques exemples de normes et de distances. Pour $x = (x_1, \ldots, x_n)$, on peut poser :

$$||x||_1 = \sum_{j=1}^{n} |x_j|, \tag{1.4}$$

$$||x||_\infty = \mathrm{Max}(|x_1|, |x_2|, \ldots, |x_n|), \tag{1.5}$$

$$||x||_2 = \Big\{\sum_{j=1}^{n} |x_j|^2\Big\}^{1/2}, \tag{1.6}$$

et même, pour $1 \leq p < +\infty$,

$$||x||_p = \Big\{\sum_{j=1}^{n} |x_j|^p\Big\}^{1/p}, \tag{1.7}$$

qui généralise $||x||_1$ et $||x||_2$.

Comparaison des normes
On utilisera les inégalités suivantes qui disent, avec le vocabulaire qui sera défini ci-dessous, que les normes $||x||_p$, $1 \leq p \leq +\infty$, sont équivalentes :

$$||x||_\infty \leq ||x||_p \leq ||x||_1 \leq n||x||_\infty \text{ pour } 1 < p < +\infty \text{ et tout } x \in \mathbb{R}^n\,. \quad (1.8)$$

La première inégalité est facile (ne garder que le plus grand terme dans la somme des $|x_j|^p$), ainsi que la dernière (chaque terme de la somme est inférieur au plus grand, et on somme). Pour la seconde, il s'agit de voir que

$$\sum_{j=1}^{n} |x_j|^p \leq \{\sum_{j=1}^{n} |x_j|\}^p\,.$$

On fait ceci par récurrence sur n. Le cas de $n = 1$ est trivial. Pour $n = 2$, il s'agit de voir que $a^p + b^p \leq (a+b)^p$ pour $a, b \geq 0$. On fixe a, on note que c'est vrai pour $b = 0$, puis on dérive par rapport à b, et il suffit de noter que pour tout b,

$$pb^{p-1} \leq p(a+b)^{p-1}\,.$$

La fonction $\mathbb{R}^+ \ni b \mapsto (a+b)^p - a^p - b^p$ est donc croissante et nulle pour $b = 0$. Elle est donc positive pour tout b positif.
Enfin, pour passer de n à $n+1$, on utilise l'inégalité qu'on vient de démontrer :

$$\begin{aligned}
\textstyle\sum_{j=1}^{n+1} |x_j|^p &= |x_{n+1}|^p + \textstyle\sum_{j=1}^{n} |x_j|^p \\
&= a^b + b^p \\
&\leq (a+b)^p \\
&= \Big\{ |x_{n+1}| + \{\textstyle\sum_{j=1}^{n} |x_j|^p\}^{1/p} \Big\}^p \\
&\leq \Big\{ |x_{n+1}| + \textstyle\sum_{j=1}^{n} |x_j| \Big\}^p \\
&= \Big\{ \textstyle\sum_{j=1}^{n+1} |x_j| \Big\}^p
\end{aligned}$$

par l'hypothèse de récurrence (élevée à la puissance $1/p$). □

Exercice : Montrer que $||x||_p \leq ||x||_q$ pour $1 \leq q < p \leq +\infty$ et $x \in \mathbb{R}^n$.

Proposition 1.3.2
Pour $1 \leq p \leq +\infty$, la fonction $||x||_p$ définie ci-dessus est une norme sur $\mathbb{R}^n$.

On a une définition semblable quand E est un espace vectoriel sur $\mathbb{C}$. Alors N est toujours définie sur E et à valeurs dans $[0, +\infty[$, et la seule chose qui change est qu'on demande (1.1) aussi pour $\lambda \in \mathbb{C}$.

Pour la démonstration de la proposition, seule la propriété (1.2) est non triviale (pour (1.3), on utilise le fait que si une somme de nombres ≥ 0 est nulle, chaque terme est nul).

Quand $p = 1$ et quand $p = +\infty$, c'est une conséquence facile de l'inégalité triangulaire dans $\mathbb{R}$, qu'on laisse également en exercice.

Reste donc (1.2) avec $1 < p < +\infty$. On va juste le vérifier quand $p = 2$. Quand $p \neq 2$, c'est un petit peu délicat (quelques inégalités de convexité à démontrer), et vous le verrez l'an prochain sous le nom d'inégalité de Minkowski.

Pour $p = 2$, donc, on utilise le produit scalaire

$$\langle x \, , \, y \rangle = \sum_{i=1}^{n} x_i \, y_i \, .$$

On doit prouver que

$$||x + y||_2 \leq ||x||_2 + ||y||_2 \, ,$$

ou encore que

$$||x + y||_2^2 \leq ||x||_2^2 + ||y||_2^2 + 2||x||_2 \, ||y||_2 \, ,$$

(puisque tous ces nombres sont positifs).
Mais

$$||x + y||_2^2 = ||x||_2^2 + ||y||_2^2 + 2\langle x, y \rangle$$

en développant les sommes, donc il suffit de voir que

$$|\langle x, y \rangle| \leq ||x||_2 \, ||y||_2 \, ,$$

qui est l'inégalité de Cauchy-Schwarz.

Pour démontrer cette dernière, on rappelle qu'on dit que

$$\lambda \mapsto ||x + \lambda y||_2^2 = ||x||_2^2 + \lambda^2 ||y||_2^2 + 2\lambda \langle x \, , \, y \rangle \, ,$$

est un polynôme de degré 2 en λ, qui reste positif, donc n'a pas deux racines distinctes, ce qui implique qu'il a un discriminant négatif ou nul.

Distance associée à une norme

Maintenant, dès qu'on a une norme N sur un espace vectoriel E (et on dit alors que (E, N) est un espace vectoriel normé), on peut définir une distance sur E en posant $\mathrm{dist}(x, y) = N(x - y)$. Le fait que dist est bien une distance sur E est une conséquence facile des définitions (et on n'a pas vraiment besoin de (1.1), mais seulement du cas où $\lambda = -1$). En plus, la distance est invariante par translation :

$$\mathrm{dist}(x+z, y+z) = \mathrm{dist}(x,y)\,, \forall x,y,z \in E$$

et (1.1) dit même que

$$\mathrm{dist}(\lambda x, \lambda y) = |\lambda|\,\mathrm{dist}(x,y)\,.$$

Bref, cette distance se marie bien aux deux lois interne et externe qui donnent à E sa structure d'espace vectoriel sur $\mathbb{R}$.

Exemples sur $\mathbb{R}^n$.
Le plus souvent, on prendra $N(x) = ||x||_2$, ce qui donne la distance euclidienne vue plus haut. Le plus souvent, on notera juste $||x||$ ou même $|x|$ au lieu de $||x||_2$. Mais on peut prendre n'importe quelle norme $||x||_p$, $1 \le p \le +\infty$, et on obtient une distance dist_p, définie par

$$\mathrm{dist}_p(x,y) = ||x-y||_p\,, \text{ pour } x \in \mathbb{R}^n\,.$$

Noter que les normes $||x||_p$, $1 \le p \le +\infty$, sur $\mathbb{R}^n$ sont (deux à deux) équivalentes, au sens suivant.

Définition 1.3.3
On dit que deux **normes** *N_1 et N_2 sur E sont* **équivalentes** *quand il existe une constante $C \ge 1$ telle que*

$$N_1(x) \le C N_2(x) \quad \textit{et} \quad N_2(x) \le C N_1(x) \quad \textit{pour tout } x \in E. \tag{1.9}$$

Evidemment, C ne doit surtout pas dépendre de x et on n'a pas besoin de prendre la même constante dans chacune des inégalités (la plus grande des deux constantes conviendra pour les deux inégalités). Ici, l'équivalence des diverses normes $||x||_p$ vient de (1.8) (et on peut même prendre $C = n$).

Pour ceux qui veulent en savoir plus, *voici d'autres exemples, dont le but est de montrer qu'il n'y a pas que $\mathbb{R}^n$ comme espace vectoriel muni d'une norme.*
1. *On peut prendre pour E l'ensemble des suites $\{x_n\}$ à valeurs réelles (pour avoir un espace vectoriel réel), telles que $\sum_{n=0}^{+\infty} |x_n| < +\infty$. Il est d'usage de noter ℓ^1 ou $\ell^1(\mathbb{N})$ cet espace.*

La norme la plus naturelle sur ℓ^1 est $||x||_1 = \sum_{n=0}^{+\infty} |x_n|$, mais on pourrait aussi utiliser $||x||_2 = \left\{\sum_{n=0}^{+\infty} |x_n|^2\right\}^{1/2}$, ou $||x||_\infty = \sup_{n\ge 0} |x_n|$.

On peut comparer ces normes, au sens où

$$||x||_\infty \leq ||x||_2 \leq ||x||_1$$

pour tout $x \in \ell^1$. *La première inégalité est toujours aussi facile. pour tout* $\varepsilon > 0$, *on choisit* n *tel que* $|x_n| \geq ||x||_\infty - \varepsilon$, *alors, rien qu'en gardant le terme* $|x_n|$ *dans* $||x||_2$, *on trouve que*

$$||x||_2 \geq |x_n| \geq ||x||_\infty - \varepsilon\,.$$

Pour la seconde, on note qu'on a encore

$$\sum_{n \leq N} |x_n|^2 \leq \Big\{ \sum_{n \leq N} |x_n| \Big\}^2$$

pour tout N *(c'est juste (1.8) sur* $\mathbb{R}^n$*) ; on passe à la limite en* N, *en notant que chacun des deux membres donne une suite croissante, que le membre de droite tend vers* $||x||_1^2$, *et le membre de gauche tend vers* $||x||_2^2$.

Ce qui nous intéresse le plus dans cet exemple, c'est que ces trois normes ne sont plus du tout équivalentes en dimension infinie. En effet, notons $x(N)$ *la suite définie par* $x(N)_n = 1$ *pour* $0 \leq n < N$, *et* $x(N)_n = 0$ *pour* $n \geq N$. *Alors* $||x(N)||_1 = N$, $||x(N)||_2 = \sqrt{N}$, *et* $||x(N)||_\infty = 1$, *ce qui montre assez clairement qu'aucune de ces normes n'est équivalente à une autre.*

2. *On peut aussi prendre pour* E *l'ensemble des fonctions* f *continues sur un intervalle fermé borné* I, *et utiliser l'une des normes*

$$||f||_\infty = \sup\{|f(x)|\,;\, x \in I\} \text{ (la norme de la convergence uniforme),}$$

ou

$$||f||_1 = \int_I |f(x)|\,dx \text{ (la norme intégrale ou } L^1\text{)},$$

ou encore

$$||f||_2 = \Big\{ \int_I |f(x)|^2 dx \Big\}^{1/2} \text{ (la norme } L^2\text{)}\ ,$$

(voir l'an prochain).

Ces normes ne sont pas équivalentes non plus.

Topologie sur un espace normé

Retournons au cas d'un espace vectoriel normé (E, N). Puisque E est muni d'une distance, il est automatiquement muni d'une topologie (où les ouverts sont définis comme au chapitre précédent). La proposition suivante dit que remplacer la norme N par une autre norme équivalente ne modifie pas la topologie :

Proposition 1.3.4
Soient N_1 et N_2 deux normes sur E. Si N_1 et N_2 sont équivalentes, alors elles définissent la même topologie sur E.

Autrement dit, la liste des ouverts est exactement la même dans les deux cas. Notons d_1 la distance associée à N_1 (donc $d_1(x,y) = N_1(x-y)$), et pareil pour d_2 et N_2. Ces **distances** sont **équivalentes** au sens où il existe $C \geq 1$ (et on peut prendre le même que pour (1.9) tel que

$$d_1(x,y) \leq Cd_2(x,y) \text{ et } d_2(x,y) \leq Cd_1(x,y) \text{ pour tout choix de } x,y \in E. \tag{1.10}$$

La vérification est vraiment très facile (écrire $d_1(x,y) = N_1(x-y)$ et $d_2(x,y) = N_2(x-y)$). On va carrément vérifier que si d_1 et d_2 sont deux distances équivalentes sur E, alors elles définissent la même topologie sur E.

Par exemple, soit U un ouvert pour (la topologie associée à) d_1, et montrons que c'est un ouvert pour d_2. Soit donc $x \in U$. Puisque U est ouvert pour d_1, il existe $r > 0$ tel que $B_1(x,r) \subset U$, où l'on note $B_1(x,r) = \{y \in E\,;\, d_1(x,y) < r\}$ (la boule ouverte pour d_1).

Vérifions que $B_2(x, r/C) \subset B_1(x,r)$, où on note $B_2(x,r/C) = \{y \in E\,;\, d_2(x,y) < r/C\}$. C'est facile : si $y \in B_2(x,r/C)$, alors $d_2(x,y) < r/C$, donc $d_1(x,y) \leq Cd_2(x,y) < r$, grâce à (1.10)). Donc on a bien trouvé une petite boule pour d_2, à savoir $B_2(x,r/C)$, qui est centrée en x et contenue dans U. On peut faire cette construction pour tout x dans U et par conséquent U est un ouvert pour d_2.

La réciproque (tout ouvert pour d_2 est ouvert pour d_1) se fait par échange des rôles de d_1 et d_2. La proposition s'en déduit. □

1.3.2 Convergence des suites

Donc on munit $\mathbb{R}^n$ de la topologie associée à la norme euclidienne, ou d'ailleurs de toute autre norme équivalente (en vertu de la proposition ci-dessus). La topologie élémentaire de $\mathbb{R}^n$ sera facile à étudier, au moins dans un premier temps, parce que l'on pourra se ramener à $\mathbb{R}$ en regardant coordonnée par coordonnée.

Proposition 1.3.5
Soit $\{x_k\}$, $k \in \mathbb{N}$, une suite à valeurs dans $\mathbb{R}^n$. Notons x_k^j, $1 \leq j \leq n$, les coordonnées de x_k. Alors :
- *Si la suite $\{x_k\}$ converge (dans $\mathbb{R}^n$) vers une limite x, alors pour tout $j \in \{1,2,\ldots,n\}$, la suite $\{x_k^j\}$ converge vers x^j, la j^{eme} coordonnée de x.*
- *Si pour tout $j \in \{1,2,\ldots,n\}$, la suite $\{x_k^j\}$, $k \in \mathbb{N}$, converge vers une limite x^j, alors la suite $\{x_k\}$ converge vers $x = (x^1,\ldots,x^n)$.*

On a mis les coordonnées en haut pour éviter les conflits de notations ; c'est aussi pour cela qu'on a indicé nos suites par k.

Supposons pour commencer que $\{x_k\}$ converge vers x, fixons j, et montrons que $\{x_k^j\}$ converge vers x^j. On se donne $\varepsilon > 0$; on sait (par la proposition sur les limites de suites dans un espace métrique) qu'il existe $N \geq 0$ tel que pour $k \geq N$, on ait $\mathrm{dist}(x_k, x) \leq \varepsilon$. Alors on a aussi :

$$|x_k^j - x^j| \leq \mathrm{dist}(x_k, x) \leq \varepsilon$$

pour $k \geq N$. Bref, on a vérifié à la main que $\lim_{k\to+\infty} x_k^j = x^j$.

Réciproquement, montrons que si $\lim_{k\to+\infty} x_k^j = x^j$ pour tout $j \leq n$, alors $\lim_{k\to+\infty} x_k = x$, où $x = (x^1, \ldots, x^n)$.

Soit $\varepsilon > 0$. On trouve, pour tout j, un N_j tel que $|x_k^j - x_j| \leq \varepsilon/n$ pour tout $k \geq N_j$. Prenons pour N le plus grand des N_j ; alors pour $k \geq N$, on a $||x_k - x|| \leq \sum_j |x_k^j - x^j| \leq \sum_j \varepsilon/n \leq \varepsilon$, par (1.8) notamment, et comme souhaité. □

On aurait aussi pu démontrer la partie directe en montrant d'abord que la j-ième projection $x \to x^j$, qui va de $\mathbb{R}^n$ dans $\mathbb{R}$, est continue, puis en appliquant la proposition 1.2.20 ci-dessus, qui dit que si $\{x_k\}$ converge vers x, et $f : \mathbb{R}^n \to \mathbb{R}$ est continue en x, alors $\{f(x_k)\}$ converge vers $f(x)$.

On a la même proposition pour la continuité des fonctions, avec la même démonstration (qu'on ne recopiera donc pas) :

Proposition 1.3.6
Soient $(E, \mathcal{T})$ un espace topologique, $f : E \to \mathbb{R}^n$ une application, et x un point de x. Alors f est continue au point x si et seulement si chacune de ses fonctions coordonnées est continue au point x.

Est implicite dans cet énoncé le fait que $\mathbb{R}^n$ et $\mathbb{R}$ sont munis des topologies usuelles (venant par exemple de la distance euclidienne). Les fonctions coordonnées sont les fonctions f_j, $1 \leq j \leq n$, définies par le fait que pour tout $y \in E$, $f(y) = (f_1(y), \ldots, f_n(y))$.

Evidemment, on déduit de la proposition le fait que f est continue (partout) si et seulement si chacune des fonctions coordonnées est continue partout.

1.3.3 Complétude

Définition 1.3.7
On dit qu'un espace métrique (E, dist) est complet si toute suite de Cauchy dans E est convergente.

On "rappelle" que la suite $\{x_k\}$ est de Cauchy si, pour tout $\varepsilon > 0$, il existe $N \geq 0$ tel que $\mathrm{dist}(x_k, x_\ell) \leq \varepsilon$ pour tout $k \geq N$ et tout $\ell \geq N$.

Noter que, bien sûr, la définition d'une suite de Cauchy dépend de la distance autant que de E. Cependant, si on remplace dist par une distance équivalente d_2 (comme en (1.10)), cela ne change pas la liste des suites de Cauchy, ni la topologie, donc (E, dist) est complet si et seulement si (E, d_1) est complet.

Notons aussi que, dans n'importe quel espace métrique, toute suite convergente est de Cauchy. La démonstration est la même que dans $\mathbb{R}$.

On a vu comme premier exemple d'espace complet : $\mathbb{R}$. En voici un autre.

Théorème 1.3.8
L'espace $\mathbb{R}^n$ (muni de la distance euclidienne, ou donc n'importe quelle distance équivalente) est complet.

Démonstration.
Soit $\{x_k\}$ une suite de Cauchy. Pour tout $j \leq n$, on a que $|x_k^j - x_\ell^j| \leq |x_k - x_\ell|$ pour tout choix de k et $\ell \in \mathbb{N}$, donc la suite $\{x_k^j\}$ (à valeurs dans $\mathbb{R}$) est de Cauchy. Comme $\mathbb{R}$ est complet, $\{x_k^j\}$ converge. Comme ceci est vrai pour tout j, la proposition 1.3.5 dit que $\{x_k\}$ aussi est convergente. $\square$

Exemple typique d'espace métrique non complet :
une partie A de E qui n'est pas fermée (comme $\mathbb{Q}$ dans $\mathbb{R}$) n'est pas complète. Bien sûr A est muni de la restriction à A de la distance sur E.

En effet, soit $x \in \overline{A} \setminus A$. On sait que tout point de $\overline{A} \setminus A$ est limite d'une suite $\{x_k\}$ dans A (c'est la proposition 1.2.15).

Cette suite est de Cauchy dans E (puisqu'elle est convergente), donc elle est aussi de Cauchy en tant que suite dans A (la définition ne fait intervenir que les distances entre les x_k). Il ne reste plus qu'à vérifier qu'elle ne converge pas dans A.

Mais si $\{x_k\}$ convergeait dans A, vers une limite x', cela signifierait que

$$\lim_{k \to +\infty} \mathrm{dist}(x', x_k) = 0\,. \tag{1.11}$$

Or on sait déjà que $\lim_{k\to+\infty} \mathrm{dist}(x, x_k) = 0$ (puisque $\{x_k\}$ converge vers x dans E), et puisque $\mathrm{dist}(x, x') \leq \mathrm{dist}(x, x_k) + \mathrm{dist}(x_k, x')$ et que le second membre tend vers 0, il vient $\mathrm{dist}(x, x') \leq 0$ par passage à la limite, donc $x = x'$. [Une contradiction, puisqu'on a supposé que $x \notin A$ et que $x' \in A$.]

On aurait aussi pu noter que (1.11) signifie aussi que $\{x_k\}$ converge vers x' dans E, et utiliser l'unicité de la limite dans E. L'argument ci-dessus re-démontre en fait **l'unicité de la limite dans un espace métrique** E.

Signalons pour finir que la réciproque est vraie quand E est complet :

Proposition 1.3.9
Si (E, dist) est un espace métrique complet et $A \subset E$, alors (A, dist) est complet si et seulement si A est fermé dans E.

On vient de voir que A n'est pas complet quand il n'est pas fermé. Supposons maintenant qu'il est fermé. Soit $\{x_k\}$ une suite de Cauchy dans A. C'est aussi une suite de Cauchy dans E (la distance est la même !), donc elle converge dans E vers une limite x. Mais $x \in \overline{A}$ parce que la suite est à valeurs dans A, donc $x \in A$ puisque A est fermé. Et finalement $\lim_{k \to +\infty} x_k = x$ dans A aussi (vérifier les définitions). $\square$

Evidemment, ceci signifie en particulier que tout fermé de $\mathbb{R}^n$ (muni de la distance euclidienne) est complet.

1.3.4 Compacité

Définition 1.3.10
Soit (E, dist) un espace métrique. On dira que E a la propriété de Bolzano-Weierstrass si de toute suite $\{x_n\}_{n \in \mathbb{N}}$ à valeurs dans E, on peut extraire une sous-suite convergente $\{x_{\phi(k)}\}_{k \in \mathbb{N}}$.

Par léger abus de notation, on dira aussi, de manière équivalente que E est compact (quand il a la propriété de Bolzano-Weierstrass).

Hors cours, pour les mordus.
En fait, la vraie définition générale de "E est compact" (pour un espace topologique) est la suivante. On demande à la fois que E soit séparé (juste une petite précaution), et surtout que E ait la propriété de Borel-Lebesgue sur les recouvrements par des ouverts. Voyons ce que ceci veut dire.

On dit que l'espace topologique $(E, \mathcal{T})$ est séparé si pour tout choix de $x, y \in E$, avec $x \neq y$, il existe un voisinage V de x qui ne contient pas y. Dans un sens, c'est bien le moins qu'on puisse demander (on exige que la topologie sache distinguer x de y). Pour un espace métrique, c'est trivial : prendre $V = B(x, \frac{1}{2}\mathrm{dist}(x, y))$. Mais la topologie grossière sur un ensemble à au moins deux éléments n'est pas séparée.

La propriété de Borel-Lebesgue s'énonce ainsi : pour tout recouvrement de E par des ouverts, il existe un sous-recouvrement de E par un nombre fini de ces ouverts.

De manière plus explicite, si U_i, $i \in I$, est une famille d'ouverts de E (indexée par un ensemble I) telle que $E \subset \cup_{i \in I} U_i$, alors il existe un ensemble fini $I_0 \subset I$ tel que $E \subset \cup_{i \in I_0} U_i$.

On a fait un petit abus de langage ici, en écrivant $E \subset \cup_{i\in I} U_i$ (et pareil pour I_0), alors que visiblement on aurait dû dire $E = \cup_{i\in I} U_i$ (puisque tout se passe dans E). Il y a quand même une excuse : en pratique, E est souvent un sous-espace d'un espace F plus grand, et les ouverts de E sont de la forme $E \cap U_i$, où les U_i sont des ouverts de F. Alors la formulation plus haut est exacte, à la fois en écrivant $E = \cup_i (E \cap U_i)$ et plus simplement $E \subset \cup_i U_i$.

Par exemple, l''intervalle $I = [0, 1]$ (dans $\mathbb{R}$) est compact. Ceci n'est pas trop difficile à vérifier, mais quand même il faut travailler un peu. Mais juste pour avoir une idée, fixez $\varepsilon > 0$ petit, et partez du recouvrement de I par les ouverts $]x - \varepsilon, x + \varepsilon[$, $x \in I$, et voyez à la main quel recouvrement fini on peut prendre.

Il se trouve que (et c'est un théorème), lorsque E est un espace métrique, E est compact si et seulement s'il a la propriété de Bolzano-Weierstrass. Donc la définition ci-dessus n'est pas mensongère, et elle nous permettra de gagner du temps (parce que nous n'aurons pas vraiment besoin d'espaces topologiques non métriques ni de la propriété de Borel-Lebesgue). On va donc abandonner la propriété de Borel-Lebesgue avec un petit regret, et continuer ce cours avec Bolzano-Weierstrass.

Exemples.
On a vu que tout intervalle $[-M, M]$ dans $\mathbb{R}$ est compact (a la propriété de Bolzano-Weierstrass). On va voir bientôt que, plus généralement, toute partie fermée et bornée de $\mathbb{R}^n$ est compacte.

Par contre, $\mathbb{R}$ tout entier n'est pas compact (prendre la suite de terme général $x_k = 2^k$), et $[0, 1[$ non plus (prendre une suite qui tend vers 1).

Théorème 1.3.11
Tout espace métrique compact est complet.

Soit donc (E, d) un espace métrique compact. On veut montrer que (E, d) est complet, donc que toute suite de Cauchy dans E est convergente. Soit $\{x_n\}$ une telle suite. Le plus dur est de deviner une limite potentielle, et ici on l'obtient de la manière suivante. Par le théorème de Bolzano-Weierstrass, on peut extraire une sous-suite de $\{x_n\}$ qui converge. Notons-la $\{x_{n_k}\}$ pour utiliser une autre notation courante. [Donc $n_k = \varphi(k)$ avec nos notations standard.] Soit x la limite ; on va vérifier que

$$\lim_{n\to+\infty} x_n = x\,, \tag{1.12}$$

(et pas seulement pour la sous-suite $\{x_{n_k}\}$). On se donne donc $\varepsilon > 0$.

Puisque $\{x_n\}$ est suite de Cauchy, il existe $N \geq 0$ tel que

$$\operatorname{dist}(x_m, x_n) \leq \varepsilon/2 \quad \text{pour } m, n \geq N. \tag{1.13}$$

Puisque $\lim_{k\to+\infty} x_{n_k} = x$, il existe $K \geq 0$ tel que $\operatorname{dist}(x_{n_k}, x) \leq \varepsilon/2$ pour $k \geq K$. Choisissons $k \geq K$ tel qu'en plus $n_k \geq N$ (c'est vrai pour tout k assez grand). Alors on peut prendre $m = n_k$ dans (1.13) ci-dessus, et on trouve que $\operatorname{dist}(x_{n_k}, x_n) \leq \varepsilon/2$ pour tout $n \geq N$. Et donc aussi

$$\operatorname{dist}(x, x_n) \leq \operatorname{dist}(x, x_{n_k}) + \operatorname{dist}(x_{n_k}, x_n) \leq \varepsilon/2 + \varepsilon/2 = \varepsilon \text{ pour } n \geq N\,.$$

On en déduit bien (1.12), donc aussi que toute suite de Cauchy dans E est convergente, et que (E, d) est complet. $\square$

On a démontré au passage que que si $\{x_n\}$ est une suite de Cauchy dans un espace métrique, et s'il existe une sous-suite $\{x_{\varphi(k)}\}$ qui converge, alors $\{x_n\}$ elle même converge vers la même limite.

Le théorème suivant sera bien utile.

Théorème 1.3.12
Soit K une partie de $\mathbb{R}^n$, que nous munissons de la topologie restreinte (qui est aussi celle qui vient de la restriction à K de la distance euclidienne). Alors K est compact si et seulement si il est à la fois fermé et borné dans $\mathbb{R}^n$.

Commençons par les vérifications faciles de conditions nécessaires.

La condition de fermeture est nécessaire, parce que si K est compact, il est complet, donc aussi fermé à cause des propositions 1.3.11 et 1.3.9. Ou directement, s'il n'est pas fermé, on choisit x dans $\overline{K} \setminus K$, puis une suite $\{x_k\}$ à valeurs dans K et qui converge vers x. Toutes les sous-suites de $\{x_k\}$ convergent donc également vers x (dans $\mathbb{R}^n$), et donc aucune ne converge dans K (parce que $x \notin K$ et par l'argument d'unicité des limites fait avant la proposition 1.3.9). Bref, K ne vérifie pas la propriété de Bolzano-Weierstrass.

Rappelons que dans $\mathbb{R}^n$, on dit que K est borné quand il existe $R > 0$ tel que $K \subset B(0, R)$.

Si tel n'est pas le cas, pour tout $k \geq 0$, il existe $x_k \in K$ tel que $||x_k||_2 \geq 2^k$. Ceci nous donne une suite $\{x_k\}$ dans K, dont il est facile de montrer qu'aucune sous-suite ne converge (elle s'en vont toutes à l'infini). Par exemple, pour montrer que si $y \in \mathbb{R}^n$ et une sous-suite $\{x_{\varphi(k)}\}$ sont donnés, alors $\{x_{\varphi(k)}\}$ ne converge pas vers y, on observe que :
$\operatorname{dist}(x_{\varphi(k)}, y) \geq \operatorname{dist}(x_{\varphi(k)}, y) - \operatorname{dist}(y, 0) \geq 2^{\varphi(k)} - \operatorname{dist}(y, 0) \geq 2^k - \operatorname{dist}(y, 0)\,,$

parce que $\varphi(k) \geq k$; le membre de droite tend vers $+\infty$, donc ne tend pas vers 0, donc y n'est pas limite de la suite extraite.

Cet argument vaut aussi bien dans un espace métrique (où l'on dira que K est borné s'il est vide, ou s'il existe $x \in K$ et $R > 0$ tels que $K \subset B(x, R)$). Donc un espace métrique compact est toujours borné.

On va donner deux démonstrations de la partie intéressante du théorème, à savoir que $K \subset \mathbb{R}^n$ est compact quand il est fermé borné. La deuxième est "hors cours".

Première démonstration (extractions successives).
On commence par une démonstration dont l'idée est de se ramener à $\mathbb{R}$ en extrayant des sous-suites succcessives.

Soit donc $K \subset \mathbb{R}^n$ fermé et borné, et on veut montrer que K a la propriété de Bolzano-Weierstrass. On se donne donc une suite $\{x_k\}$ à valeurs dans K, et on veut extraire une sous-suite qui converge dans K.

Notons que pour chaque choix de coordonnée j, la suite $\{x_k^j\}$ est bornée. [Car si $K \subset B(0, R)$, alors $|x_k^j| \leq R$ pour tout j et tout k.] Ceci restera vrai de toute sous-suite de $\{x_k^j\}$.

Donc on part de la suite $\{x_k\}$, et on commence par regarder la suite bornée $\{x_k^1\}$ à valeurs réelles. Par le théorème de Bolzano-Weierstrass (démontré dans la première section), il existe une sous-suite convergente $\{x^1_{\varphi_1(k)}\}$, associée à une première fonction strictement croissante $\varphi_1 : \mathbb{N} \to \mathbb{N}$. On note x^1 sa limite.

Maintenant on regarde la sous-suite $\{x^2_{\varphi_1(k)}\}_{k\in\mathbb{N}}$ correspondante, mais pour la seconde coordonnée. C'est encore une suite bornée. Donc on peut en extraire une nouvelle sous-suite convergente de limite x^2. Appelons φ_2 l'application strictement croissante correspondante. La suite extraite est donc $\{x^2_{\varphi_1(\varphi_2(k))}\}_{k\in\mathbb{N}}$.

On recommence avec la suite bornée extraite $\{x^3_{\varphi_1(\varphi_2(k))}\}_{k\in\mathbb{N}}$. C'est une suite bornée, et on en extrait une sous-suite convergente $\{x^3_{\varphi_1(\varphi_2(\varphi_3(k)))}\}_{k\in\mathbb{N}}$ qui converge vers un certain x^3.

Et ainsi de suite. A la fin de la construction, on a obtenu une suite extraite $\{x^n_{\psi(k)}\}_{k\in\mathbb{N}}$, avec $\psi = \varphi_1 \circ \ldots \circ \varphi_n$, et qui converge vers un certain x^n.

Vérifions que la suite extraite $\{x_{\psi(k)}\}$ (qui est cette fois à valeurs dans $\mathbb{R}^n$), converge vers x, le vecteur dont la j-ième coordonnée est $x^j = \lim_{k\to+\infty} x^j_{\varphi_1(\varphi_2(\ldots(\varphi_j(k)\ldots))}$ (la j-ième suite extraite de notre argument).

On sait qu'il suffit de procéder coordonnée par coordonnée, et de vérifier que pour chaque $j \in \{1, \ldots, n\}$, on a $\lim_{k\to+\infty} x^j_{\psi(k)} = x^j$. Ou encore, il suffit de vérifier que la suite $\{x^j_{\psi(k)}\}$ est une suite extraite de la suite $\{x^j_{\varphi_1(\varphi_2(\ldots(\varphi_j(k)\ldots))}\}$ citée ci-dessus, et qui converge vers x^j. C'est bien le cas, puisque ψ est la composée de $\varphi_1 \circ \varphi_2 \circ \dots \varphi_j$ et de $\varphi_{j+1} \circ \dots \varphi_n$. En termes moins techniques, on dit juste que la suite $\{x^j_{\psi(k)}\}$ est construite à partir de $\{x^j_{\varphi_1(\varphi_2(\dots(\varphi_j(k)\dots))}\}$ par extractions successives, et qu'une suite extraite de suite extraite est une suite extraite.

En résumant ce qu'on a obtenu coordonnée par coordonnée, la suite $\{x_{\psi(k)}\}$, $k \geq 0$, est convergente. Pour l'instant, la convergence est dans $\mathbb{R}^n$.

Maintenant, la limite x est dans $\overline{K}$, puisque chaque élément est dans K. Donc, $x \in K$ puisque K est fermé. Enfin, dire que $x = \lim_{k\to+\infty} x_{\psi(k)}$ dans $\mathbb{R}^n$, c'est aussi dire que $\lim_{k\to+\infty} ||x_{\psi(k)} - x|| = 0$, donc encore que $x = \lim_{k\to+\infty} x_{\psi(k)}$ dans K. (En fait, le seul risque était que $x \notin K$).

C'est fini, on a trouvé une suite extraite qui converge dans K (et on a utilisé nos deux hypothèses). $\square$

Seconde démonstration (pour les mordus).

On peut aussi regarder attentivement la démonstration donnée dans le cas de $\mathbb{R}$, et s'assurer qu'elle marche encore.

D'abord, on avait vu que la seule chose à démontrer est que toute suite dans K a au moins une valeur d'adhérence. Dès qu'on a cette valeur d'adhérence x, on trouve une sous-suite qui converge vers x exactement comme dans la proposition 1.1.12, qui marche dans un espace métrique quelconque.

Il reste à trouver une valeur d'adhérence, ce qu'on essaie de faire comme au théorème 1.1.11, mais en généralisant un peu la construction. On va le refaire parce que c'est joli. On commence par le lemme suivant.

Lemme 1.3.13

Soit $K \subset \mathbb{R}^n$ une partie bornée. Alors pour tout $k > 0$, il existe une famille finie B_i, $i \in I_k$, de boules de rayon 2^{-k}, telle que

$$K \subset \bigcup_{i\in I_k} B_i\,. \tag{1.14}$$

Evidemment, on aurait pu remplacer 2^{-k} par n'importe quel $\varepsilon > 0$. Plus on prend les rayons petits (donc k grand), plus il semble difficile de trouver les B_i.

Pour info, quand K vérifie la propriété du lemme (pour tout k), on dit qu'il est précompact.

Evidemment, ceci ressemble un peu à la propriété de Borel-Lebesgue, mais en moins fort puisqu'on peut partir d'un recouvrement donné de K par toutes les boules de rayon ε.

Démonstration du lemme.

Puisque K est borné, il existe un entier R tel que $K \subset B(0, R) \subset [-R, R]^n$

(un gros cube centré en 0*). On choisit un entier* N *assez grand pour que* $N > 2^k\sqrt{n}$*, et on note* Z *l'ensemble des points* $y \in [-R, R]^n$ *dont chaque coordonnée est un multiple entier de* $1/N$*. Ainsi* Z *est un ensemble fini, et tout point de* $[-R, R]^n$ *est à distance* $\sqrt{n}/N < 2^{-k}$ *d'un point de* Z*. Alors les boules* $B(y, 2^{-k})$*,* $k \in Z$*, sont en nombre fini, et recouvrent bien* $[-R, R]^n$*, donc aussi* K*.* □

Passons à la construction de boîtes. On se donne une suite $\{x_m\}$ *à valeurs dans* K *(notre ensemble borné). On va construire par récurrence une suite d'ensembles* R_k*,* $k \geq 0$*, avec les deux propriétés suivantes :*

$$R_k \text{ contient } x_m \text{ pour une infinité de valeurs de } m \tag{1.15}$$

$$R_k \text{ est contenu dans une boule de rayon } 2^{-k}, \tag{1.16}$$

et

$$R_k \subset R_{k-1} \text{ pour } k \geq 1. \tag{1.17}$$

On commence par R_0*. On utilise le lemme, qui donne une famille finie* B_i*,* $i \in I_0$*, de boules de rayon* 1 *et qui recouvre* K*. Si aucune ne contenait* x_m *pour une infinité de valeurs de* m*, ce serait vrai aussi pour la réunion de ces boules, ce qui est faux : il y a bien une infinité de valeurs de* m *au départ, et* $x_m \subset K \subset \cup B_i$ *pour chacune.*

Donc on choisit pour R_0 *une boule* B_i *qui contient* x_m *pour une infinité de valeurs de* m*, et on a (1.15) et (1.16) pour* $k = 0$*.*

Soit maintenant $k \geq 1$*, supposons construit* R_{k-1}*, et construisons* R_k*. Par hypothèse de récurrence,* R_k *contient* x_m *pour une infinité de valeurs de* m*. A cause du lemme, on peut recouvrir* K *par un nombre fini de boules* B_i*,* $i \in I_k$*, de rayon* 2^{-k}*. Comme dans le cas où* $k = 0$*, l'intersection de* R_{k-1} *avec au moins une de ces boules contient* x_m *pour une infinité de valeurs de* m*. On choisit* B_i *comme ceci, et on prend* $R_k = R_{k-1} \cap B_i$*.*

Les propriétés (1.16) et (1.17) à l'ordre k *sont alors clairement satisfaites, et ceci termine notre construction des* R_k *par récurrence.*

Reprenons notre recherche de valeurs d'adhérence. Choisissons un point y_k *dans chaque* $K \cap R_k$*. C'est sans doute aussi bien de choisir l'un des éléments de la suite (il y en a plein à cause de (1.15)), mais cela n'est même pas nécessaire. Par contre, c'est utile de savoir que* $K \cap R_k$ *n'est pas vide.*

Notons que pour tout $N \geq 0$ *et tout choix de* $k \geq N$ *et* $l \geq N$*, les deux points* y_k *et* y_l *sont dans* R_l *(à cause de (1.17)), et donc* $\mathrm{dist}(y_k, y_l) \leq 2^{-N+1}$ *(à cause de (1.16)). On en déduit que* $\{y_k\}$ *est une suite de Cauchy. Donc elle converge, puisque* K *est* **complet** *(il est fermé dans* $\mathbb{R}^n$ *qui est complet). Soit* y *la limite. Il ne reste plus qu'à vérifier que* y *est une valeur d'adhérence de notre suite* $\{x_m\}$*.*

On se donne donc $\varepsilon > 0$. Puisque y est la limite des y_k, il existe $k_0 \geq 0$ tel que $\operatorname{dist}(y, y_k) \leq \varepsilon/2$ pour $k \geq k_0$. Choisissons k tel que de plus $2^{-k+1} < \varepsilon/2$. Notons qu'alors $R_k \subset B(y_k, 2^{-k+1}) \subset B(y_k, \varepsilon/2)$ à cause de (1.16), donc aussi $R_k \subset B(y, \varepsilon)$ par l'inégalité triangulaire. Mais R_k contient x_m pour une infinité de valeurs de m, par (1.15), donc $B(y, \varepsilon)$ aussi contient x_m pour une infinité de valeurs de m. Et comme ceci est vrai pour tout $\varepsilon > 0$, y est bien une valeur d'adhérence de la suite $\{x_m\}$.

On a fini de montrer l'existence d'une valeur d'adhérence $y \in E$ de notre suite. Soit on en déduit l'existence d'une sous-suite de $\{x_m\}$ qui converge vers y, comme indiqué plus haut, et on en déduit que $y \in K$ parce que K est fermé, soit on montre directement que toute valeur d'adhérence d'une suite de points de K est dans $\overline{K}$ (exercice), soit encore on utilise le fait que l'on pouvait choisir y_k dans K, donc que $y \in K$. Dans tous les cas, on trouve que $y \in K$, donc la suite $\{x_m\}$ a une valeur d'adhérence dans K. On en déduit le théorème (seconde démonstration).

Notons au passage qu'on a en fait démontré que tout espace métrique précompact et complet a la propriété de Bolzano-Weierstrass. La réciproque est vraie, et même plutôt facile. □

Remarque

L'histoire des fermés bornés qui sont compacts n'est vraie que parce qu'on a supposé que K est un sous-ensemble de $\mathbb{R}^n$. "Compact" a une définition précise, et n'est pas synonyme de "fermé borné". Voici deux exemples.

1. Dans l'espace métrique $[0, 1[$, l'intervalle $[0, 1[$ est une partie fermée bornée, mais n'est pas un compact. Assurez-vous que vous comprenez bien pourquoi, puis lisez la suite.

Ceci se présente plus naturellement quand on se place dans un ouvert U de $\mathbb{R}^n$, et qu'on regarde les parties de U. Dans ce cadre naturel aussi, il serait erroné de croire que les parties fermées bornées de U sont compactes. [Bien sûr, il s'agit des parties qui sont fermées dans U, pour la topologie restreinte, donc des intersections de U avec un fermé de $\mathbb{R}^n$.]

2. *Exemple en dimension infinie (**pour les mordus**).*

Plaçons nous dans l'espace vectoriel normé $E = \ell^1(\mathbb{N})$, l'espace des suites $x = \{x_n\}$ telles que $||x||_1 = \sum_n |x_n|$ est fini, muni de la norme $||x||_1$. C'est un espace métrique, et on peut vérifier qu'il est complet. Mais la boule unité fermée $B = \{x \in \ell^1(\mathbb{N})\,;\, ||x||_1 \leq 1\}$ n'est pas compacte. Montrons pour le vérifier que la suite $\{e_k\}$ dans $\ell^1(\mathbb{N})$, où $(e_k)_n = 0$ si $n \neq k$ et $(e_k)_n = 1$ si $n = k$ ne converge pas, et n'a pas non plus de sous-suite convergente.

On aura besoin de savoir que pour tout j, l'application $\psi_j : E \to \mathbb{R}$ qui à $x \in \ell^1(\mathbb{N})$ associe sa j-ième coordonnée x_j est continue. En fait, *elle est même lipschitzienne avec une constante 1, puisque pour $x, y \in E$, $|\psi_j(x) - \psi_j(y)| = |x_j - y_j| \leq ||x - y||_1 = \sum_k |x_k - y_k|$.*

Revenons à la non compacité de B. Soit donc $\{e_{\varphi(k)}\}$ une sous-suite de notre suite initiale, et supposons pour obtenir une contradiction que $y = \lim_{k\to+\infty} e_{\varphi(k)}$ existe.

D'abord, fixons j. Puisque ψ_j est continue, on en déduit que $\psi_j(y) = \lim_{k\to+\infty} \psi_j(e_{\varphi(k)})$. Mais $\psi_j(e_{\varphi(k)})$, qui est la j-ième coordonnée de $e_{\varphi(k)}$, est nul pour k assez grand (en fait, dès que $\varphi(k) > j$). Donc la limite est nulle, et $y_j = 0$.

On vient de montrer que $y_j = 0$ pour tout j, donc $y = 0$.

D'un autre côté, dans tout espace vectoriel normé, $||x||$ est une fonction continue de x (voir (19) plus bas et la ligne qui le suit), donc $||y||_1 = \lim_{k\to+\infty} ||e_{\varphi(k)}||_1 = 1$, parce que chaque e_n est de norme 1. Ceci contredit bien sûr le fait que $y = 0$, et de cette contradiction on déduit qu'aucune sous-suite $\{e_{\varphi(k)}\}$ ne converge.

Bref, la boule unité fermée de $\ell^1(\mathbb{N})$ n'est pas compacte. En fait c'est toujours comme ça dans les espaces vectoriels normés de dimension infinie : c'est le théorème de Riesz que vous verrez ultérieurement.

1.3.5 Applications

On donne quelques applications spectaculaires du théorème 1.3.12. On commence par la généralisation logique du théorème 1.1.13.

Théorème 1.3.14
Soient (K, dist) un espace métrique compact. [Par exemple, K peut être n'importe quelle partie fermée bornée de $\mathbb{R}^n$.] Soit $f : K \to \mathbb{R}$ une application continue. Alors f est bornée et atteint ses bornes.

Rappelons que dans le cas où K est une partie de $\mathbb{R}^n$, il est muni de la distance euclidienne, ou d'ailleurs de n'importe quelle autre distance équivalente.

Comme précédemment, la conclusion signifie que $m = \inf\{f(x)\,;\, x \in K\}$ et $M = \sup\{f(x)\,;\, x \in K\}$ sont des réels, et qu'il existe $a \in K$ et $b \in K$ tels que $f(a) = m$ et $f(b) = M$.

La démonstration est la même que pour le théorème 1.1.13. □

Théorème 1.3.15
Soit N une norme sur $\mathbb{R}^n$. Alors il existe une constante $C > 0$ telle que

$$C^{-1}\, ||x||_2 \leq N(x) \leq C\, ||x||_2 \quad \textit{pour tout } x \in \mathbb{R}^n \tag{1.18}$$

En bref, N est équivalente à la norme euclidienne. Nous avons vu ce résultat avec les normes $||x||_p$, où $1 \leq p \leq +\infty$ (voir au début de la section les inégaltés (1.8)). Ici, ce qui est nouveau, c'est que N est n'importe quelle norme.

Nous comparons à la norme euclidienne pour simplifier, mais notons que si N_1 et N_2 sont deux normes sur $\mathbb{R}^n$, alors N_2 est équivalente à N_1 (au sens de la définition 1.3.3), simplement parce que N_1 et N_2 sont toutes les deux équivalentes à $||\cdot||_2$.

Quand même, la constante d'équivalence (la constante C qui intervient dans (1.9)) ne dépend pas de x, mais dépend de N_1 et N_2. C'est inévitable : on peut prendre $N_1(x) = ||x||_2$ et $N_2(x) = \lambda\, ||x||_2$ avec un $\lambda > 0$ gigantesque, ou plus sérieusement, en dimension 2, prendre $N_2(x) = |x_1| + \lambda |x_2|$.

Pour résumer l'énoncé de façon un peu lapidaire :

Toutes les normes sur $\mathbb{R}^n$ sont équivalentes.

Du coup, elles définissent des distances équivalentes et la même topologie.

Démonstration (hors programme).

Une partie de l'équivalence est facile à établir. Notons $(e_1, \ldots, e_n)$ la base canonique de $\mathbb{R}^n$, puis posons $M = N(e_1) + \ldots + N(e_n)$. Soit $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$. Alors

$$\begin{aligned} N(x) &= N(\textstyle\sum_{j=1}^n x_j e_j) \\ &\leq \textstyle\sum_{j=1}^n N(x_j e_j) \\ &= \textstyle\sum_{j=1}^n |x_j| N(e_j) \\ &\leq \left[\sup_j |x_j|\right] \textstyle\sum_{j=1}^n N(e_j) \\ &= M\, ||x||_\infty \\ &\leq M\, ||x||_2 \end{aligned} \tag{1.19}$$

en appliquant l'inégalité triangulaire (1.2), puis l'homogénéité de la norme (1.1), puis la définition de $||x||_\infty$ donnée en (1.5) et enfin la première inégalité facile dans (1.8). [On aurait pu faire un peu mieux en appliquant Cauchy-schwarz après la fin de la première ligne, ce qui permettait de prendre $M = [N(e_1)^2 + \ldots + N(e_n)^2]^{1/2}$.

Donc N est dominée par $||x||_2$, et il ne restera plus qu'à voir le contraire.

Vérifions que la fonction N est Lipschitzienne, avec la constante M. Notons que pour $x, y \in \mathbb{R}^n$, $N(x) \leq N(y) + N(x-y)$ par l'inégalité triangulaire, ce qui fait que

$$N(x) - N(y) \leq N(x-y) \leq M\, ||x-y||_2\,. \tag{1.20}$$

Le même raisonnement (en échangeant x et y) donne

$$N(y) - N(x) \leq M||x-y||_2\,,$$

donc N est bien lipschitzienne et par conséquent elle est continue.

Appliquons le théorème 1.3.12 à la restriction de la fonction N à la sphère unité S (pour la distance euclidienne). On trouve que $m = \inf\{N(x)\,;\, x \in S\}$ est fini (ce qui était clair de toute façon, puisque $m \geq 0$), mais surtout qu'il existe $a \in S$ tels $N(a) = m$. En particulier, $m > 0$ par (1.3) et puisque $a \in S$, donc $a \neq 0$. C'était là la partie délicate (tout ce mal pour montrer que $m > 0$).

De par la définition de m, $N(y) \geq m$ pour tout $y \in S$; utilisons ceci pour vérifier que

$$N(x) \geq m\,||x||_2 \quad \textit{pour } x \in \mathbb{R}^n\,. \tag{1.21}$$

C'est clair si $x = 0$, puisqu'alors $||x||_2 = 0$. Sinon, on écrit $x = \lambda y$, avec $\lambda = ||x||_2$ et donc $y = x/||x||_2 \in S$. Et il vient

$$N(x) = N(\lambda y) = |\lambda|\,N(y) \geq |\lambda|\,m = m\,||x||_2\,,$$

comme souhaité.

L'équivalence (1.13) se déduit maintenant de (1.15) et (1.17). □

Théorème 1.3.16
Soient (E, d_1) un espace métrique compact, (F, d_2) un espace métrique, et $f : E \to F$ une application continue. Alors $f(E)$ est compact aussi.

Deux remarques.
D'abord, ce théorème (tout comme le suivant d'ailleurs) est encore vrai pour des espaces topologiques (pas nécessairement métriques), mais on a à peine donné la définition.
Ensuite, pas question de montrer que F tout entier est compact, si f n'est pas surjective.

Pour la démonstration, on veut montrer que $f(E)$ vérifie la propriété de Bolzano-Weierstrass, donc on se donne une suite $\{y_n\}$ dans $f(E)$, et on cherche une suite extraite convergente. Pour tout n, il existe $x_n \in E$ tel que $f(x_n) = y_n$. Puisque E est compact, on peut extraire de la suite $\{x_n\}$ une suite $\{x_{\varphi(k)}\}$ convergente. Par composition de limites, la suite $\{f(x_{\varphi(k)})\}$ converge vers l'image par f de la limite de la suite $\{x_{\varphi(k)}\}$. Mais cette suite n'est autre que la suite extraite $\{y_{\varphi(k)}\}$. On a donc bien trouvé une sous-suite de $\{y_n\}$ qui converge, et $f(E)$ est compact. □

Théorème 1.3.17
Soient (E, d_1) un espace métrique compact, (F, d_1) un espace métrique, et $f : E \to F$ une application continue bijective. Alors f est un homéomorphisme.

Ceci signifie que f est continue bijective (ce qu'on sait déjà), et que f^{-1} est continue.

Soit donc y un point de F, et vérifions la continuité de f^{-1} en y. Supposons en effet que f^{-1} n'est pas continue en y. Il existe donc un voisinage V de $x = f^{-1}(y)$ tel que, pour aucun voisinage W de y, on n'ait $f^{-1}(W) \subset V$. En particulier, si $W = B(y, 2^{-n})$, on trouve que $f^{-1}(W) \not\subset V$, donc il existe $y_n \in B(y, 2^{-n})$ tel que $x_n = f^{-1}(y_n)$ n'est pas dans V.

Puisque E est compact, on peut extraire de $\{x_n\}$ une sous-suite $\{x_{\varphi(k)}\}$ converge vers une limite ℓ. Comme aucun des x_n n'est dans V (et que V contient une boule $B(x,r)$), on trouve que $\ell \neq x$.

Maintenant $\{f(x_{\varphi(k)})\}$ converge vers $f(\ell)$, puisque f est continue en ℓ. Mais $f(x_{\varphi(k)}) = y_{\varphi(k)}$, donc $\{f(x_{\varphi(k)})\}$ est une sous-suite de $\{y_n\}$, qui par choix converge vers y. Bref, $\{f(x_{\varphi(k)})\}$ converge aussi vers y, et $y = f(\ell)$. On savait déjà que $y = f(x)$, et ceci contredit l'injectivité de f. $\square$

Chapitre 2

Calcul différentiel

Avant d'aborder le cas de plusieurs variables, on fait un premier tour de chauffe en rappelant (de manière non exhaustive) ce qui est supposé connu dans le cas de la dimension 1.

2.1 Fonctions d'une variable

2.1.1 Fonctions de classe C^k

On commence par quelques définitions. On dira que f est de classe C^n dans I si elle est n-fois dérivable et si sa n-ème dérivée est continue. Dans le cas $n = 0$, C^0 désigne donc l'ensemble des fonctions continues sur I. On peut aussi considérer le cas où I est un intervalle fermé $[a, b]$. $C^0([a, b])$ est alors l'ensemble des fonctions continues de l'espace topologique $[a, b]$ à valeur dans $\mathbb{R}$. On notera que cela coïncide avec une autre définition peut-être rencontrée les années précédentes disant que f est continue dans $]a, b[$ et que de plus f admet une limite à droite en a et une limite à gauche en b.

2.1.2 Formules de Taylor

On commence par la formule de Taylor avec reste intégral pour les fonctions d'une variable réelle définie sur un intervalle I ouvert.

Théorème 2.1.1
Soit $f : I \to \mathbb{R}$ une fonction de classe $\mathcal{C}^{n+1}$ et on suppose que $0 \in I$. Pour tout $k \leq n$ on a

$$f(s) = f(0) + sf'(0) + \frac{s^2}{2}f''(0) + \ldots + \frac{s^k}{k!}f^{(k)}(0) + \int_0^s \frac{(s-u)^k}{k!}f^{(k+1)}(u)du$$

Preuve : Elle se fait par récurrence.
La formule est vraie pour $k = 0$. Elle dit simplement dans ce cas que, pour $f \in C^1$, on a :

$$f(s) = f(0) + \int_0^s f'(u)\,du\,.$$

Supposons qu'elle le soit pour l'entier $k-1$, avec $k \geq 1$, c'est-à-dire que

$$\int_0^s \frac{(s-u)^{k-1}}{(k-1)!} f^{(k)}(u)du = f(s) - f(0) - sf'(0) - \ldots - \frac{s^{k-1}}{(k-1)!} f^{(k-1)}(0)\,.$$

On intègre par parties. Pour cela on rappelle la formule suivante. Si v et w sont deux fonctions de classe C^1 sur $[a,b]$, alors on a

$$\int_a^b v'(t)w(t)\,dt = -\int_a^b v(t)w'(t)\,dt + v(b)w(b) - v(a)w(a)\,. \tag{2.1}$$

On applique cette formule avec $v(t) = f^{(k)}(t)$, $w(t) = \dfrac{(s-t)^k}{k!}$, $a = 0$ et $b = s$. Ceci donne :

$$\int_0^s \frac{(s-u)^k}{k!} f^{(k+1)}(u)\,du = -\frac{s^k}{k!} f^{(k)}(0) + \int_0^s \frac{(s-u)^{k-1}}{(k-1)!} f^{(k)}(u)\,du\,,$$

et donc avec l'hypothèse de récurrence

$$\int_0^s \frac{(s-u)^k}{k!} f^{(k+1)}(u)du =$$

$$-\frac{s^k}{k!} f^{(k)}(0) + f(s) - f(0) - sf'(0) - \ldots - \frac{s^{k-1}}{(k-1)!} f^{(k-1)}(0)\,.$$

On voit donc que la formule est vraie pour l'entier k. □

Remarque 2.1.2
On utilise cette formule également sous la forme

$$f(s) = f(0) + sf'(0) + \ldots + \frac{s^k}{k!} f^{(k)}(0) + \frac{s^{k+1}}{k!} \int_0^1 (1-u)^k f^{(k+1)}(su)du\,, \tag{2.2}$$

dont on voit qu'elle implique :

$$\left| f(s) - \left(f(0) + sf'(0) + \ldots + \frac{s^k}{k!} f^{(k)}(0) \right) \right| \leq M_s \frac{|s|^{k+1}}{(k+1)!}\,, \tag{2.3}$$

avec

$$M_s = \sup_{u \in [0,s]} |f^{(k+1)}(u)| \,.$$

On note que M_s est fini car une fonction continue (ici $f^{(k)}$) sur un compact (ici $[0,s]$) est bornée. On note aussi que, si $s < 0$, on doit écrire plus correctement :

$$M_s = \sup_{u \in [s,0]} |f^{(k+1)}(u)| \,.$$

Le cas $k = 0$, correspond à la formule des accroissements finis

$$\Big| f(s) - f(0) \Big| \leq \sup_{u \in I} |f'(u)| \, |s| \,, \tag{2.4}$$

Si on n'est pas intéressé par la constante précise, on peut écrire :

$$f(s) = f(0) + sf'(0) + \ldots + \frac{s^k}{k!} f^{(k)}(0) + \ \mathcal{O}(|s|^{k+1}) \,. \tag{2.5}$$

On notera que cette dernière formule implique :

$$f(s) = f(0) + sf'(0) + \ldots + \frac{s^k}{k!} f^{(k)}(0) + \ o(|s|^k) \,, \tag{2.6}$$

formule qui est vraie sous l'hypothèse plus faible que $f \in C^k$. On démontre en fait cette dernière formule en utilisant la formule de Taylor avec reste intégral au cran $(k-1)$.
Une dernière variante de (2.3) est d'écrire qu'il existe $\xi \in [0,s]$ tel que :

$$f(s) = f(0) + sf'(0) + \ldots + \frac{s^k}{k!} f^{(k)}(0) + \frac{s^{k+1}}{(k+1)!} \ f^{(k+1)}(\xi) \,, \tag{2.7}$$

Si cette dernière formule peut paraître séduisante, il ne faut pas oublier que ξ dépend de s et qu'on ne sait rien de cette dépendance par rapport à s. Sans le cas $k = 0$, on retrouve que pour une fonction de classe C^1 on peut écrire :

$$f(s) - f(0) = sf'(\xi) \,. \tag{2.8}$$

Exercice 2.1.3
Comment trouve-t-on (2.2) à partir de la formule du théorème 2.1.1 ?

2.1.3 Extrema

Définition 2.1.4
Soit $f : I \to \mathbb{R}$ une fonction, et s_0 un point de I. On dit que f admet un maximum (resp. minimum) local en s_0 s'il existe un intervalle $]s_0-\alpha, s_0+\alpha[$ dans I pour tout s duquel on ait $f(s) \leq f(s_0)$ (resp. $f(s) \geq f(s_0)$).

Voici d'abord une condition nécessaire d'extremum local.

Proposition 2.1.5
Soit $f : I \to \mathbb{R}$ une fonction dérivable et soit $s_0 \in I$. Si f admet un extremum local en s_0, alors $f'(s_0) = 0$.

Preuve : On prend $s_0 = 0$ pour simplifier. Puisque f est dérivable, il existe une fonction ϵ de limite nulle en 0 telle que

$$f(s) = f(0) + s(f'(0) + \epsilon(s)).$$

Supposons que $f'(0) \neq 0$, par exemple que $f'(0) > 0$. Puisque ϵ tend vers 0 en 0, il existe un intervalle $]-\alpha, \alpha[$ pour tout s duquel $f'(0)+\epsilon(s) > 0$. Mais alors $f(s) > f(0)$ pour $\alpha > s > 0$ et $f(s) < f(0)$ pour $-\alpha < s < 0$, ce qui contredit le fait que 0 est un extremum local pour f. $\square$

Remarque 2.1.6
Attention ! Dans le cas où la fonction f est définie sur un intervalle fermé $[a, b]$, le critère ci-dessus ne vaut que pour les extrema de f dans $]a, b[$. On le voit dans la démonstration ci-dessus puisque l'on doit pouvoir calculer $f(s)$ pour des $s < 0$ et des $s > 0$. On peut aussi penser au cas d'une fonction strictement monotone, par exemple $f : [0, 1] \to \mathbb{R}$, $f(x) = x$. Cette fonction admet un maximum en $x = 1$, alors que $f'(1) = 1$.

La formule de Taylor permet d'énoncer une condition suffisante pour qu'une fonction (suffisamment régulière) admette un extremum local en s_0. Soit $f : I \to \mathbb{R}$ une fonction de classe $\mathcal{C}^3$. On écrit la formule de Taylor (2.2) pour $k = 2$:

$$f(s) = f(0) + sf'(0) + \frac{s^2}{2}f''(0) + \frac{s^3}{2!}\int_0^1 (1-u)^2 f^{(3)}(su)du\,.$$

Il est peut-être utile de rappeler que l'on peut obtenir un résultat moins précis, mais qui nous suffirait ici, en supposant que la fonction f est de classe

$\mathcal{C}^2$ seulement. Sous cette hypothèse on peut en effet montrer, en utilisant le théorème des accroissements finis, que

$$f(s) = f(0) + sf'(0) + \frac{s^2}{2} f''(0) + s^2 \epsilon(s)\,,$$

pour une certaine fonction ϵ de limite nulle en 0. C'est d'aillleurs sous cette forme qu'on utilise la formule de Taylor pour démontrer le résultat suivant.

Proposition 2.1.7
Soit $f : I \to \mathbb{R}$ une fonction de classe $\mathcal{C}^3$, et s_0 un point de I. Si $f'(s_0) = 0$, et si $f''(s_0) \neq 0$, alors la fonction f admet un extremum local en s_0. C'est un maximum si $f''(s_0) < 0$ et un minimum si $f''(s_0) > 0$.

Preuve :
On reprend la formule de Taylor ci-dessus, sachant que $f'(0) = 0$:

$$f(s) - f(0) = \frac{s^2}{2}(f''(0) + s \int_0^1 (1-u)^2 f^{(3)}(su)\, du)\,.$$

Supposons par exemple que $f''(0) > 0$. Puisque la fonction $f^{(3)}$ est continue, elle est bornée sur l'intervalle $[0, s]$ (ou $[s, 0]$ si $s < 0$). Il existe donc une constante $C > 0$ telle que

$$|\int_0^1 (1-u)^2 f^{(3)}(su) du| \leq C\,.$$

Prenant $\alpha > 0$ assez petit, on peut donc affirmer que pour tout $s \in]-\alpha, \alpha[$, on a

$$f''(0) + s \int_0^1 (1-u)^2 f^{(3)}(su) du \geq f''(0) - C\alpha > 0\,.$$

Donc, pour tout $s \in]-\alpha, \alpha[$, on a $f(s) \geq f(0)$, ce qui montre que f admet un minimum local en 0. Le cas $f''(0) < 0$ se traite de la même manière. □

Remarque 2.1.8
On appelle souvent points critiques de f les points où f' s'annule. On vient de voir que les points critiques où la dérivée seconde de f ne s'annule pas sont des extrema de f. Les points critiques où la dérivée seconde de f s'annule sont dits dégénérés, et on peut par exemple utiliser la formule de Taylor à un ordre plus élevé pour connaitre l'allure du graphe de f au voisinage de ces points.

Exercice 2.1.9
Parmi les rectangles d'aire 1*, quel est celui qui a le plus petit périmètre ?*

Remarque 2.1.10
On peut aussi s'interroger sur les extrema d'une fonction de classe C^2 *définie sur un intervalle* $[a, b]$ *fermé. Par exemple* f *admet un minimum local en* a *si* $f'(a) > 0$.

2.2 FONCTIONS DE PLUSIEURS VARIABLES : DIFFERENTIELLES.

2.2.1 Dérivées partielles et différentiabilité

Dans la suite, on se donne un ouvert U de $\mathbb{R}^n$, et une fonction $f : U \to \mathbb{R}^m$. [On comprendra peut-être plus facilement dans ce cadre un peu plus général. En tout cas les difficultés vont venir du domaine de définition, et très peu du domaine d'arrivée).

On rappelle que f est continue en un point si et seulement si chacune de ses fonctions coordonnées f_i, $1 \leq i \leq m$, est continue en ce point. Ca sera pareil pour les diverses notions de dérivabilité ci-dessous.

On va essayer d'appeler i les numéros de coordonnées, et j les numéros des variables (donc $1 \leq j \leq n$).

On commence par la définition la plus facile, destinée aux amateurs de fonctions d'une seule variable.

Définition 2.2.1
Soient $x \in U$ *et* $1 \leq i \leq n$. *On dit que* f *a une dérivée partielle par rapport à* x_j *au point* $x = (x_1, \ldots, x_n)$ *si la fonction d'une seule variable* $t \to f(x_1, \ldots, x_{j-1}, t, x_{j+1}, \ldots, x_n)$ *est dérivable en* $t = x_j$. *Alors on note* $\frac{\partial f}{\partial x_j}(x)$ *la valeur de cette dérivée.*

Bref, on fixe toutes les variables sauf x_j, on regarde la fonction f comme une fonction de la seule variable x_j, et on dérive en x_j. Par exemple, la fonction

$$f(x_1, x_2) = x_1^2 x_2^3$$

a des dérivées partielles par rapport à chacune des deux variables, et

$$\frac{\partial f}{\partial x_1}(x_1, x_2) = 2x_1 x_2^3 \text{ et } \frac{\partial f}{\partial x_2}(x_1, x_2) = 3x_1^2 x_2^2 \,.$$

Evidemment, les notations sont un peu fausses quand $j = 1$ (il n'y a plus de $j - 1$) ou $j = n$ (il n'y a plus de $j + 1$), mais nous espérons que cet abus de langage compréhensible.

On a pris un ouvert U pour être sûr que f soit définie dans un voisinage de x. Il y aurait des adaptations (du genre, "demi-dérivées" ou dérivée à gauche ou à droite), mais oublions ceci en dimension > 1.

Quand f est à valeurs dans $\mathbb{R}^n$, on calcule simplement les dérivées partielles éventuelles coordonnée par coordonnée (c.-à.-d., indépendamment pour chaque f_i).

Il n'y a pas grand-chose à dire de plus sur les dérivées partielles : vous savez tout, puisqu'il s'agit simplement de fonctions d'une seule variable et si vous ne savez pas révisez vos cours antérieurs. Mais malheureusement, la notion ne suffit pas, et la bonne généralisation de la notion de dérivée est la suivante.

Définition 2.2.2
On dit que f est différentiable au point x s'il existe une application linéaire $A : \mathbb{R}^n \to R^m$ telle que

$$\lim_{y \to 0\,,\, y \neq 0} \frac{1}{||y||} \left[f(x+y) - f(x) - A(y) \right] = 0\,. \tag{2.9}$$

Avec des notations plus agréables, ceci veut dire qu'on a le développement limité à l'ordre 1 en 0

$$f(x+y) = f(x) + A(y) + o(||y||)\,. \tag{2.10}$$

[Noter que le cas de $y = 0$, qui était exclu dans la définition de la limite, est trivial de toute façon.] Si on préfère écrire la variable directement, on peut aussi utiliser le développement limité

$$f(y) = f(x) + A(y - x) + o(||y - x||) \tag{2.11}$$

(cette fois au voisinage de x).
Enfin, si on préfère les limites, on peut écrire :

$$\lim_{y \to x\,,\, y \neq x} \frac{1}{||y - x||} \left[f(y) - f(x) - A(y - x) \right] = 0. \tag{2.12}$$

Notez que quand $n = 1$, il s'agit bien de la notion de dérivabilité que vous connaissez. C'est juste que toute application linéaire $A : \mathbb{R} \to \mathbb{R}^m$ est de la forme $y \to ay$ pour un $a \in \mathbb{R}^m$ (prendre $a = A(1)$). En effet, dans ce cadre, (2.9) est équivalent à

$$\lim_{y \to 0, y \neq 0} \frac{f(x+y) - f(x) - A(y)}{y} = 0\,, \tag{2.13}$$

toujours avec $A(y) = ay = f'(x)y$.

Si vous avez fait un peu de fonctions analytiques, la dérivée complexe de $f : \mathbb{C} \to \mathbb{C}$ est du même type. On demande que

$$f'(x) = \lim_{z \to 0\,,\, z \neq 0} \frac{f(x+z) - f(x)}{z}$$

existe, et c'est pareil à (2.9), mais avec une application $\mathbb{C}$-linéaire A, de la forme $z \to f'(x)z$.

On peut sans rien changer généraliser la notion au cas où f est à valeurs dans un espace vectoriel normé F (de dimension infinie). On peut même, sans se fatiguer trop, généraliser au cas où f est définie dans un ouvert de l'espace vectoriel normé E. Mais alors, on demande en plus que A soit continue (ou de manière équivalente bornée), c.-à.-d., qu'il existe $C \geq 0$ tel que

$$||A(y)|| \leq C||y|| \;\; \textit{pour tout } y \in E\,.$$

Quand $E = \mathbb{R}^n$ comme ici, cette condition de continuité est automatique (exercice, ou voir la démonstration de la prochaine proposition). L'application linéaire A est appelée **différentielle de f en x**, et on la notera $Df(x)$. Il nous arrivera de noter $Df(x) \cdot y$, au lieu de $Df(x)(y)$, l'effet de l'application linéaire $Df(x)$ sur le vecteur $y \in \mathbb{R}^n$.

Pour faire des calculs, on fait souvent appel à la matrice de $Df(x)$ dans la base canonique, qu'on appellera **matrice jacobienne de f en x**. Il n'y a pas vraiment une notation standard (et parfois on la note $Df(x)$ comme la différentielle elle même). Nous essayerons dans la suite de la noter M, ou $M(x)$, ou $M_f(x)$.

Donc M est une matrice avec n lignes et m colonnes ; chaque colonne est comme d'habitude l'image du j-ième vecteur de base de $\mathbb{R}^n$ par $Df(x)$. Avec des notations matricielles, (2.11) devient

$$f(y) = f(x) + M(x)(y - x) + o(||y - x||), \tag{2.14}$$

où maintenant x est vu comme une matrice-colonne, et le terme du milieu est une multiplication de matrices.

Quand elle existe, la différentielle et donc sa matrice jacobienne sont faciles à calculer, à cause de la proposition suivante.

Proposition 2.2.3
Supposons que f soit différentiable au point x, et notons $M(x)$ sa matrice jacobienne au point x. Alors f est continue au point x, et f admet une dérivée partielle $\frac{\partial f}{\partial x_j}$ par rapport à chaque variable. De plus $\frac{\partial f}{\partial x_j}(x)$ est le j-ième vecteur colonne de $M(x)$.

Démonstration de la proposition
Démontrons directement la continuité en x en revenant à la définition avec les δ et ε. D'abord, il convient de noter qu'il existe $K \geq 0$ tel que

$$||A(y)|| \leq K\,||y||$$

pour $y \in \mathbb{R}^n$, où l'on a encore posé $A = Df(x)$.
En effet,

$$||A(y)|| = ||\sum_j y_j\, A(e_j)|| \leq \sum_j a_j |y_j| \leq (\sum_j a_j^2)^{1/2}\, ||y||$$

où on a posé $a_j = ||A(e_j)||$ et utilisé l'inégalité de Cauchy-Schwarz. Rappelons qu'on utilise ici la norme euclidienne mais aussi que cela n'a pas grande importance puisque toutes les normes sur l'espace de départ $\mathbb{R}^n$ ou l'espace d'arrivée $\mathbb{R}^m$ sont équivalentes.

Maintenant on se donne $\varepsilon > 0$. On sait qu'il existe $r_0 > 0$ tel que pour $y \in B(0, r_0)$, l'expression dans la limite de (2.9) soit de norme au plus $\frac{\varepsilon}{K+1}$:

$$\frac{1}{||y||}\, ||f(x+y) - f(x) - A(y)|| \leq \frac{\varepsilon}{K+1}\,,\ \forall y \in B(0, r_0) \setminus \{0\}\,,$$

qui est équivalent à :

$$||f(x+y) - f(x) - A(y)|| \leq \frac{\varepsilon}{K+1}\, ||y||\,,\ \forall y \in B(0, r_0)\,.$$

Notez que le cas de $y = 0$ est trivial (bien qu'il ait été exclu de la définition 2.2.2 à cause de la notion de limite).

On choisit $\delta = \mathrm{Min}(1, r_0, \frac{\varepsilon}{K+1})$. Alors, pour $y \in B(0, \delta)$,

$$\begin{aligned} ||f(x+y) - f(x)|| &\leq ||f(x+y) - f(x) - A(y)|| + ||A(y)|| \\ &\leq \tfrac{\varepsilon ||y||}{K+1} + K\,||y|| \\ &\leq \tfrac{\varepsilon}{K+1} + K\delta \\ &\leq \varepsilon\,. \end{aligned}$$

On en déduit bien la continuité de f en x.

Passons à l'existence des dérivées partielles. On va se contenter de prouver l'existence de $\frac{\partial f}{\partial x_1}$ en x, mais bien sûr les autres seraient obtenues de la même manière.

Donc il s'agit de montrer que la fonction $t \to f(x_1 + t, x_2, \ldots, x_n)$ est dérivable en 0, de dérivée $A(e_1)$ (avec les notations ci-dessus). C'est-à-dire, que

$$\lim_{t\to 0\,,\,t\neq 0} \frac{f(x+te_1) - f(x) - A(te_1)}{|t|} = 0\,. \tag{2.15}$$

C'est exactement (2.9), restreint à l'espace vectoriel engendré par e_1. □

Ainsi, si l'on sait que $Df(x)$ existe, la matrice jacobienne est facile à calculer : c'est juste la matrice composée des vecteurs colonnes $\frac{\partial f}{\partial x_j}$. Et le théorème suivant permet justement d'obtenir l'existence de $Df(x)$.

Théorème 2.2.4
Supposons que pour tout y dans un voisinage de x, f a des dérivées partielles $\frac{\partial f}{\partial x_j}$, $1 \leq j \leq n$, au point y, et que de plus chaque $\frac{\partial f}{\partial x_j}(y)$ est continue au point x. Alors f est différentiable en x (et donc sa différentielle se calcule à partir des $\frac{\partial f}{\partial x_j}(x)$ comme à la proposition précédente).

Les notations seront plus difficiles que la démonstration elle-même. Notons $v_j(y) = \frac{\partial f}{\partial x_j}(y)$ pour simplifier. On doit choisir une application linéaire A, et on définit bien sûr A par

$$A(z) = \sum_{1\leq j\leq n} z_j\, v_j(x)$$

pour $z = (z_1, \ldots, z_n) \in \mathbb{R}^n$.
Il nous faut donc évaluer

$$G(z) = f(x+z) - f(x) - A(z) = f(x+z) - f(x) - \sum_{1\leq j\leq n} z_j v_j(x)\,, \tag{2.16}$$

pour z petit.

On se donne $\varepsilon > 0$, et on choisit $r > 0$ tel que les dérivées partielles $\frac{\partial f}{\partial x_j}(y)$ existent partout sur $B(x,r)$, et de plus

$$|v_j(y) - v_j(x)| \leq \varepsilon\,, \text{ pour } y \in B(x,r)\,.$$

On veut obtenir que dans ce cas,

$$||G(z)|| \leq n\varepsilon ||z|| \text{ pour } z \in B(0,r)\,, \tag{2.17}$$

et on en déduira aussitôt le résultat (comparer à (2.9) ou (2.10)).

Soit donc $z \in B(0,r)$ qu'on écrit dans la base (e_j) sous la forme $z = \sum_j z_j e_j$. On va supposer pour commencer que les z_i sont positifs, pour simplifier les notations. Posons $y_0 = x$, $y_1 = x + z_1 e_1$, et ainsi de suite. Donc $y_k = x + \sum_{j \leq k} z_j e_j$ et à la fin $y_n = x + z$.

Notons que pour tout $1 \leq j \leq n$, le segment qui va de y_{j-1} à y_j est un segment parallèle au j-ième axe, et est entièrement contenu dans $B(x,r)$. Donc la dérivée partielle $v_j(y)$ est bien définie sur ce segment.

Considérons l'application g_j définie sur le segment $[0, z_j]$ par

$$g_j(t) = f(y_{j-1} + te_j)\,.$$

On va donc de $f(y_{j-1})$ quand $t = 0$ à $f(y_{j-1} + z_j e_j) = f(y_j)$ quand $t = z - j$. Cette application est dérivable, et sa dérivée est :

$$g_j'(t) = \frac{\partial f}{\partial x_j}(y_{j-1} + te_j) = v_j(y_{j-1} + te_j)\,.$$

[En fait, g_j est à une translation près, la fonction d'une variable qui sert dans la définition de $\frac{\partial f}{\partial x_j}$.]

Appliquons le théorème des accroissements finis à g_j, entre 0 et z_j. On trouve un point $\xi_j \in [0, z_j]$ tel que

$$g_j(z_j) - g_j(0) = z_j v_j(y_{j-1} + \xi_j e_j)\,.$$

De manière équivalente,

$$f(y_j) - f(y_{j-1}) = z_j v_j(\zeta_j)\,,$$

en posant $\zeta_j = y_{j-1} + \xi_j e_j$.

On additionne tout ceci. Comme la somme des $f(y_j) - f(y_{j-1})$ vaut

$$f(y_n) - f(y_0) = f(x+z) - f(x)\,,$$

il vient

$$f(x+z) - f(x) = \sum_j z_j\, v_j(\zeta_j).$$

En comparant avec (2.16) , il vient également

$$\begin{aligned} |G(z)| &= |f(x+z) - f(x) - \textstyle\sum_{1 \leq j \leq n} z_j\, v_j(x)| \\ &= |\textstyle\sum_{1 \leq j \leq n} z_j [v_j(\zeta_j) - v_j(x)]| \\ &\leq ||z|| \textstyle\sum_{1 \leq j \leq n} |v_j(\zeta_j) - v_j(x)| \\ &\leq n\varepsilon ||z||\,, \end{aligned}$$

puisque $|v_j(y) - v_j(x)| \leq \varepsilon$ pour $y \in B(0,r)$ (et les ζ_j sont bien dans $B(0,r)$).

C'est ce qu'on voulait, sauf qu'on a supposé que $z_j \geq 0$ pour tout j pour simplifier. Dans le cas général, on note ε_j le signe de z_j (on fait ce qu'on veut si $z_j = 0$; de toute façon $y_j = y_{j-1}$ et il n'y a rien à estimer), et on peut définir g_j sur $[0, |z_j|]$ par $g_j(t) = f(y_{j-1} + t\varepsilon_j e_j)$. Après on fait les calculs comme avant, et de toute façon les signes qui apparaîtraient sont perdus dans la bataille à la fin.

On pouvait aussi éviter ce problème en définissant g_j sur $[0,1]$ par la formule $g_j(t) = f(y_{j-1} + tz_j e_j)$ et en dérivant une fonction composée. □

Remarque 2.2.5
Sans notre hypothèse de continuité, le théorème est faux. Un exemple instructif est de considérer l'application de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par :

$$f(x,y) = \frac{x^2 y}{x^2 + y^2}\,, \quad \textit{pour } (x,y) \neq 0\,,$$

et

$$f(0,0) = 0\,.$$

Elle a des dérivées partielles partout trivialement, sauf peut-être en $(0,0)$. Comme $f(0,y) = 0$ et $f(x,0) = 0$, les dérivées partielles existent aussi en $(0,0)$. Si f était différentiable en $(0,0)$, la différentielle serait donc nulle, et on aurait $|f(x,y)| = o(||(x,y)||)$. Mais $f(x,x) = x/2$, ce qui ne correspond pas.

Il n'y a rien de magique dans cet exemple : on peut prendre presque n'importe quelle fonction f donnée en coordonnées polaires par

$$f(r\cos\theta, r\sin\theta) = rg(\theta)\,,$$

avec g Dans notre exemple précédent, $g(\theta) = \cos^2\theta \sin\theta$. Il y a automatiquement des dérivées partielles en $(0,0)$ (à savoir $g(0)$ et $g(\pi/2)$), et même dans n'importe quelle direction (voir plus loin la notion de dérivée directionnelle) (la dérivée dans la direction d'angle θ est $g(\theta)$). Mais pour que la fonction soit différentiable en $(0,0)$, il faut encore que toutes ces dérivée radiales viennent d'une même application linéaire. Du coup, c'est assez inattendu que dans le théorème plus haut, la continuité des dérivées partielles force ces dérivées radiales à une certaine cohérence.

Finissons ce paragraphe par quelques définitions standard :

On dit que f **est de classe** C^1 sur U quand les dérivées partielles $\frac{\partial f}{\partial x_j}$ sont définies partout et continues sur U. Il revient au même de dire (à cause de la proposition et du théorème) que f est de classe C^1 si f est différentiable sur U, et que $Df(x)$ est une fonction continue sur U. Cette fonction est en principe à valeurs dans $\mathcal{L}(R^n, R^m)$, qui est un espace vectoriel de dimension finie, donc isomorphe à un R^{mn}, donc muni automatiquement d'une topologie ; ceci dit, la continuité de Df est équivalente à la continuité de la matrice jacobienne $M(x)$, qui se vérifie coordonnée par coordonnée. Finalement, donc, le plus simple est de dire que les dérivées partielles sont continues.

Quand $f : U \to \mathbb{R}$, les dérivées partielles sont à valeurs dans $\mathbb{R}$. On appelle **gradient de** f **en** x, et on note $\nabla f(x)$, le vecteur de $\mathbb{R}^n$ dont les coordonnées sont les $\frac{\partial f}{\partial x_j}(x)$. Dans ces conditions, la différentielle agit de manière simple, puisque

$$Df(x)(y) = \sum_{j=1}^{n} y_j \frac{\partial f}{\partial x_j}(x) = \nabla f(x) \cdot y \tag{2.18}$$

par la proposition 2.2.3 et où la dernière expression est un produit scalaire euclidien.

Noter que le gradient donne la direction le long de laquelle f varie le plus fort, et est orienté dans la direction où f augmente, et sa taille est proportionnelle à la taille de la dérivée dans cette direction. Dans les directions orthogonales à $\nabla f(x)$ (s'il est non nul), la dérivée de f est nulle. Ainsi, dans les bons cas, la ligne ou surface de niveau passant par x est perpendiculaire à $\nabla f(x)$.

Exercice 2.2.6
Vérifier ces affirmations dans le cas de la fonction $f(x, y) = x^2 + y^2 - 1$ *ou dans le cas de la fonction* $f(x, y, z) = x^2 + y^2 + z^2 - 1$.

Toujours quand f est à valeurs dans $\mathbb{R}$, on utilise souvent la divergence de f, définie par

$$\mathrm{Div}(f)(x) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x). \tag{2.19}$$

C'est aussi la trace de la matrice jacobienne $M(x)$.

2.2.2 Dérivée directionnelle

Revenons sur une notion qui a été évoquée dans le paragraphe précédent.

Définition 2.2.7 *Soit $f : U \to \mathbb{R}$ une application, x un point de Ω et u un vecteur non nul de $\mathbb{R}^n$. On appelle dérivée directionnelle de f en x dans la direction de u la dérivée en $s = 0$, si elle existe, de la fonction d'une variable*

$$f_u : s \mapsto f(x + su).$$

On la note alors $\partial_u f(x)$.

Exemple 2.2.8 *Les dérivées partielles $\partial_{x_i} f$ ne sont autres que les dérivées directionnelles de f dans les directions des vecteurs de la base canonique (e_i).*

Proposition 2.2.9 *Soit $f : U \to \mathbb{R}$. Si f est différentiable en x, alors f admet des dérivées directionnelles en x dans toutes les directions u, et*

$$\partial_u f(x) = (Df)(x)u = \nabla f(x) \cdot u\,.$$

Attention la réciproque est fausse, comme le montrent les exemples suivants présentés sous forme d'exercice (voir Remarque 2.2.5 pour le premier) :

Exercice 2.2.10
1. On considère l'application $f : \mathbb{R}^2 \to \mathbb{R}$ définie par

$$f(x, y) = \frac{x^2 y}{x^2 + y^2}$$

pour $(x, y) \neq (0, 0)$ et $f(0, 0) = 0$. Montrer que pour tout $u \in \mathbb{R}^2$ $\partial_u f(0, 0)$ existe, mais que f n'est pas différentiable en $(0, 0)$.
2. Même question pour l'application $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x, y) = \frac{x^3 y}{x^4+y^2}$ pour $(x, y) \neq (0, 0)$ et $f(0, 0) = 0$.

2.2.3 Dérivées des fonctions composées

Proposition 2.2.11
Soient U un ouvert de $\mathbb{R}^n$, V un ouvert de $\mathbb{R}^m$, $f : U \to \mathbb{R}^m$, $g : V \to \mathbb{R}^k$, et $x \in U$. On suppose que f est différentiable en x et que g est différentiable en $f(x)$ (donc $f(x) \in V$). Alors $g \circ f$ est différentiable en x, et sa dérivée (=différentielle) en x est $D(g \circ f)(x) = Dg(f(x)) \circ Df(x)$.

Ainsi $Dg(f(x)) \circ Df(x)$ est l'application linéaire composée (qui va de $\mathbb{R}^n$ dans $\mathbb{R}^k$, comme il se doit).

Au niveau des matrices jacobiennes aussi les choses sont simples :
La matrice jacobienne de $g \circ f$ en x est $M(x)\, N(f(x))$ (produit de la matrice jacobienne de f en x (une matrice $n \times m$) et de la matrice jacobienne de g en $f(x)$ (une matrice $m \times k$)).
En effet, la matrice de $D(g \circ f)(x)$ est la matrice de la composée de $Dg(f(x))$ et de $Df(x)$; c'est donc bien le produit des matrices de $Dg(f(x))$ et de $Df(x)$.

La démonstration de la proposition est facile : il s'agit juste de manipuler des développements limités d'ordre 1. Quand z tend vers $f(x)$, on sait que

$$g(z) = g(f(x)) + Dg(f(x)) \cdot (z - f(x)) + o(||z - f(x)||). \tag{2.20}$$

On applique ceci avec $z = f(y)$ pour y proche de x. C'est légitime, puisque f est continue en x par la proposition 2.2.3, donc z tend vers $f(x)$ quand y tend vers x. Noter que

$$z = f(y) = f(x) + Df(x) \cdot (y - x) + o(||y - x||) \tag{2.21}$$

On remplace et on trouve

$$\begin{aligned} g(f(y)) &= g(z) \\ &= g(f(x)) + Dg(f(x)) \cdot [Df(x) \cdot (y - x) + o(||y - x||)] + o(||z - f(x)||) \\ &= g(f(x)) + Dg(f(x)) \cdot [Df(x) \cdot (y - x)] + o(||y - x||) + o(||z - f(x)||) \end{aligned}$$

parce que

$$||Dg(f(x)) \cdot v|| \leq C\, ||v||\,,$$

Ensuite,

$$||z - f(x)|| \leq C\, ||y - x||$$

près de y, par (2.21), donc

$$o(||z - f(x)||) = o(||y - x||)\,.$$

Donc

$$g(f(y)) = g(f(x)) + [Dg(f(x)) \circ Df(x)] \cdot (y - x) + o(||y - x||)$$

comme souhaité. □

Il est souvent aussi pratique d'écrire ceci au niveau des dérivées partielles. Alors si f, g, et x sont comme dans l'énoncé de la proposition, on obtient

$$\frac{\partial (g \circ f)}{\partial x_j}(x) = \sum_{i=1}^{m} \frac{\partial g}{\partial y_i}(f(x))\, \frac{\partial f_i}{\partial x_j}(x) \tag{2.22}$$

où éventuellement g a le droit d'être à valeurs dans $\mathbb{R}^k$, et alors (2.22) peut être pris coordonnée par coordonnée, et signifie que pour $1 \leq \ell \leq k$,

$$\frac{\partial(g_\ell \circ f)}{\partial x_j}(x) = \sum_{i=1}^{m} \frac{\partial g_\ell}{\partial y_i}(f(x)) \frac{\partial f_i}{\partial x_j}(x). \tag{2.23}$$

Par exemple, on veut exprimer la “même” fonction à la fois en coordonnées cartésiennes et polaires, et calculer les dérivées correspondantes. Par exemple, supposons que la fonction est donnée par la fonction

$$g(x,y) = x^3 + y^3\,,$$

et on veut calculer ses dérivées en coordonnées polaires. [Cela a l'air d'un problème un peu différent de ce qu'on vient de faire, mais c'est souvent comme ceci que cela se présente.]
Donc on veut poser $x = r\cos\theta$ et $y = r\sin\theta$. Ceci fait qu'en fait on s'intéresse à la fonction $G = g \circ f$, où f est définie sur le domaine adéquat $U \subset \mathbb{R}^2$ (et qu'il faudrait choisir en fonction du problème) par $f(r,\theta) = (r\cos\theta, r\sin\theta)$.
Dans notre exemple, donc, G sera donnée par

$$G(r,\theta) = g(r\cos\theta, r\sin\theta) = r^3\cos^3\theta + r^3\sin^3\theta\,,$$

dont on pourrait bien sûr calculer les dérivées partielles directement. Mais en tout cas la proposition dit que

$$\begin{aligned}
\tfrac{\partial G}{\partial r} &= \tfrac{\partial g}{\partial x}\tfrac{\partial x}{\partial r} + \tfrac{\partial g}{\partial y}\tfrac{\partial y}{\partial r} \\
&= \tfrac{\partial g}{\partial x}\cos\theta + \tfrac{\partial g}{\partial y}\sin\theta \\
&= 3x^2\cos\theta + 3y^2\sin\theta \\
&= 3r^2\cos^3\theta + 3r^2\sin^3\theta,
\end{aligned}$$

où l'on a noté un peu abusivement x et y les deux coordonnées de f (ce sont des fonctions de r et θ).
Evidemment, le résultat final concorde avec ce qu'on aurait pu calculer directement plus vite, et aussi on voit que les calculs ci-dessus, une fois tout réécrit, ne sont autres que des calculs de dérivées de fonctions composées d'une seule variable à la fois. De même,

$$\begin{aligned}
\tfrac{\partial G}{\partial \theta} &= \tfrac{\partial g}{\partial x}\tfrac{\partial x}{\partial \theta} + \tfrac{\partial g}{\partial y}\tfrac{\partial y}{\partial \theta} \\
&= -\tfrac{\partial g}{\partial x}\, r\sin\theta + \tfrac{\partial g}{\partial y}\, r\cos\theta \\
&= -3x^2 r\sin\theta + 3y^2 r\cos\theta \\
&= -3r^3\cos^2\theta\sin\theta + 3r^3\cos\theta\sin^2\theta\,,
\end{aligned}$$

ce qui est à nouveau sans surprise quant au résultat final.
Vous avez certainement fait ce genre de calcul en physique, surtout si vous

avez vu de la thermodynamique, ou d'ailleurs le plus souvent on ne note pas différemment les fonctions g et G (ce qui est un peu dangereux, mais pas trop si l'on note consciencieusement les arguments de la fonction), et est motivé par le fait qu'après tout c'est la même fonction (température, ou autre), juste vue comme exprimée à partir d'autres variables).

2.2.4 DÉRIVÉES D'ORDRES SUPÉRIEURS.

Juste quelques mots sur cette question et d'ailleurs on va se concentrer sur des dérivées d'ordre 2.

Si U est un ouvert de $\mathbb{R}^n$ et $f : U \to \mathbb{R}$ est, disons, de classe C^1 pour commencer, on a n applications continues $g_j = \frac{\partial f}{\partial x_j}$, définies sur U et à valeurs dans $\mathbb{R}$. On peut se demander si ces fonctions sont à leur tour dérivables (et ainsi de suite).

Ainsi, on dit que f a des dérivées partielles d'ordre 2 si f a des dérivées partielles continues (donc elle est différentiable), et si chacune de ces dérivées a elle-même des dérivées partielles (en tout point). On notera $\frac{\partial^2 f}{\partial x_k \partial x_j}$ la dérivée partielle de $\frac{\partial f}{\partial x_j}$ par rapport à k, sauf quand $j = k$, où l'on note directement $\frac{\partial^2 f}{\partial x_k^2}$.

On dit que f est de classe C^2 sur U si de plus toutes les différentielles d'ordre deux qu'on vient de définir sont continues sur U.

On dit que f est de classe C^k si l'on peut continuer à dériver ainsi jusqu'à l'ordre k compris, et que les dérivées partielles $\frac{\partial^k f}{\partial x_{j_1} \cdots \partial x_{j_k}}$ sont continues. Enfin, on dit que f est de classe C^∞ s'il est de classe C^k pour tout k.

Proposition 2.2.12
Soient U un ouvert de $\mathbb{R}^n$, V un ouvert de $\mathbb{R}^m$, $f : U \to \mathbb{R}^m$, $g : V \to \mathbb{R}^k$, et $x \in U$. On suppose que f est de classe C^k sur U et que g est de classe C^k sur V. Alors $g \circ f$ est de classe C^k sur U.

La proposition est admise. Cette proposition permet en particulier de montrer qu'une fonction est de classe C^k une fois connue cette propriété pour des fonctions plus simples. Par exemple, si f est une fonction de $\mathbb{R}$ dans $\mathbb{R}$ de classe C^k, la fonction $(x, y) \mapsto f(x^2 + y^4)$ est de classe C^k.

2.2.5 Formule de Taylor

On va se contenter de $k = 2$. Par une démonstration semblable à celle du théorème 2.2.4 (considérer la fonction $[0, 1] \ni t \mapsto g(t) = f((1 - t)x + ty)$ et utiliser la formule de Taylor à une variable en 0), on montre que si f a des dérivées partielles d'ordre 2 dans un voisinage de x, et qui sont continues au point x, on a le développement limité d'ordre 2 au voisinage de x :

$$f(y) = f(x)+\sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x)(y_j-x_j)+\frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_k}(x)(y_j-x_j)(y_k-x_k)+o(||x-y||^2)$$

ou de manière équivalente,

$$f(x+z) = f(x)+\sum_{j=1}^{n} z_j \frac{\partial f}{\partial x_j}(x)+\frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} z_j z_k \frac{\partial^2 f}{\partial x_j \partial x_k}(x)+o(||z||^2)\,, \quad (2.24)$$

quand z tend vers 0.
On peut aussi déduire de la démonstration que (toujours si les dérivées partielles secondes sont continues en x) que :

Proposition 2.2.13
Si f est de classe C^2, alors :

$$\frac{\partial^2 f}{\partial x_j \partial x_k}(x) = \frac{\partial^2 f}{\partial x_k \partial x_j}(x)\,, \quad (2.25)$$

Ceci tombe bien, parce que du coup il est à peu près inutile de se souvenir si $\frac{\partial^2 f}{\partial x_k \partial x_j}$ est la dérivée partielle de $\frac{\partial f}{\partial x_j}$ par rapport à k, ou le contraire.
Du coup aussi, vous pouvez encore écrire (2.24) avec moins de termes :

$$f(x+z) = f(x)+\sum_{j=1}^{n} z_j \frac{\partial f}{\partial x_j}(x)+\frac{1}{2}\sum_{j=1}^{n} z_j^2 \frac{\partial^2 f}{\partial x_j^2}(x)+\sum_{1\le j<k\le n} z_j z_k \frac{\partial^2 f}{\partial x_j \partial x_k}(x)+o(||z||^2)\,. \quad (2.26)$$

Noter encore que, puisqu'il s'agit de calculer des dérivées partielles successives, vous savez calculer les dérivées successives pour une fonction composée, même si les calculs sont parfois compliqués, comme on le voit dans le paragraphe suivant.

Enfin pour étendre la propriété mentionnée (2.25) pour $k = 2$, l'ordre dans lequel on prend les dérivées partielles d'ordre inférieur ou égal à k d'une fonction de classe C^k ne compte pas. Ainsi par exemple pour $k = 3$ et $n = 2$ on a :

$$\partial^3_{xxy} f = \partial^3_{xyx} f = \partial_{yxx} f\,.$$

Remarque 2.2.14
Il y a d'autres notations utilisées pour les dérivées partielles. Mentionnons $\partial_x f$ *au lieu de* $\dfrac{\partial f}{\partial x}$, $\partial^2_{xy} f$ *au lieu de* $\dfrac{\partial^2 f}{\partial x \partial y}$, $\partial^2_{ij} f$ *au lieu de* $\dfrac{\partial^2 f}{\partial x_i \partial x_j}$.

2.2.6 Le Laplacien en coordonnées polaires

Il s'agit de réécrire pour une fonction u de classe C^2 sur $\mathbb{R}^2$, la fonction $(\Delta u)(x,y) = \partial^2_{xx} u + \partial^2_{yy} u$ en utilisant les coordonnées polaires. Plus précisément, pour $(x,y) \neq (0,0)$, il existe un unique couple (r,θ) dans $]0,+\infty[\times[0,2\pi[$ tel que

$$\left\{ \begin{array}{l} x = r\cos\theta \\ y = r\sin\theta \end{array} \right.$$

Pour $f \in C^2$, notons g la fonction définie sur $]0,+\infty[\times[0,2\pi[$ par

$$g(r,\theta) = f(x,y) = f(r\cos\theta, r\sin\theta)\,.$$

On voit d'abord facilement que

$$\left\{ \begin{array}{l} \partial_r g(r,\theta) = \cos\theta\,\partial_x f(x,y) + \sin\theta\,\partial_y f(x,y) \\ \partial_\theta g(r,\theta) = -r\sin\theta\,\partial_x f(x,y) + r\cos\theta\,\partial_y f(x,y). \end{array} \right. \tag{2.27}$$

Ensuite, en formant les combinaisons linéaires adéquates, on obtient

$$\left\{ \begin{array}{l} \partial_x f(x,y) = \cos\theta\,\partial_r g(r,\theta) - \dfrac{\sin\theta}{r}\,\partial_\theta g(r,\theta) \\ \partial_y f(x,y) = \sin\theta\,\partial_r g(r,\theta) + \dfrac{\cos\theta}{r}\,\partial_\theta g(r,\theta). \end{array} \right. \tag{2.28}$$

On peut alors démontrer le résultat suivant.

Lemme 2.2.15
Soit u *une fonction de classe* $\mathcal{C}^2$, *et* $v(r,\theta) = u(r\cos\theta, r\sin\theta)$. *On a*

$$\Delta u(r\cos\theta, r\sin\theta) = \partial^2_{rr} v(r,\theta) + \frac{1}{r}\partial_r v(r,\theta) + \frac{1}{r^2}\partial^2_{\theta\theta} v(r,\theta)$$

Preuve : Posons $x = r\cos\theta$ et $y = r\sin\theta$. Soit aussi $w(r,\theta) = \partial_x u(x,y)$. On a d'abord, en utilisant (2.28) pour $f(x,y) = \partial_x u(x,y) = w(r,\theta)$,

$$\partial^2_{xx} u(x,y) = \partial_x(\partial_x u(x,y)) = \cos\theta\,\partial_r w(r,\theta) - \frac{\sin\theta}{r}\,\partial_\theta w(r,\theta).$$

On a ensuite, en utilisant (2.28) pour $f(x,y) = u(x,y) = v(r,\theta)$,

$$\begin{cases} \partial_r w(r,\theta) = \partial_r(\partial_x u(x,y)) = \partial_r(\cos\theta\,\partial_r v(r,\theta) - \dfrac{\sin\theta}{r}\,\partial_\theta v(r,\theta)) \\ \partial_\theta w(r,\theta) = \partial_\theta(\partial_x u(x,y)) = \partial_\theta(\cos\theta\,\partial_r v(r,\theta) - \dfrac{\sin\theta}{r}\,\partial_\theta v(r,\theta)), \end{cases}$$

On a donc finalement

$$\partial^2_{xx} u(x,y) = \cos^2\theta\,\partial^2_{rr} v(r,\theta) + \frac{\sin^2\theta}{r}\partial_r v(r,\theta) + \frac{\sin^2\theta}{r^2}\partial^2_{\theta\theta} v(r,\theta).$$

Un calcul identique donne

$$\partial^2_{yy} u(x,y) = \sin^2\theta\,\partial^2_{rr} v(r,\theta) + \frac{\cos^2\theta}{r}\partial_r v(r,\theta) + \frac{\cos^2\theta}{r^2}\partial^2_{\theta\theta} v(r,\theta),$$

et l'on obtient le lemme en ajoutant ces deux dernières égalités. □

On peut par exemple utiliser cette expression pour déterminer toutes les fonctions harmoniques qui sont invariantes par rotation. Il s'agit des fonctions u de classe $\mathcal{C}^2$ telles que $\Delta u(x,y) = 0$, et pour lesquelles, notant $v(r,\theta) = u(r\cos\theta, r\sin\theta)$, on a $\partial_\theta v(r,\theta) = 0$ (v ne dépend pas de θ).

Avec le Lemme 2.2.15, on doit alors avoir

$$0 = \Delta u(x,y) = \partial^2_{rr} v(r,\theta) + \frac{1}{r}\partial_r v(r,\theta)\,.$$

On peut remarquer que cette équation s'écrit

$$0 = \partial_r(r\partial_r v(r,\theta))\,,$$

et donc que ses solutions sont

$$v(r,\theta) = C_1\log(r) + C_2\,. \tag{2.29}$$

Bien entendu, il faut ici supposer que $r \neq 0$.

2.3 RECHERCHE D'EXTREMA LOCAUX

2.3.1 Conditions nécessaires et suffisantes

On se donne un ouvert U de $\mathbb{R}^n$ et une fonction $f : U \to \mathbb{R}$. On recherche des points de U où $f(x)$ est localement maximale ou minimale. Voyons ce que ceci veut dire.

Définition 2.3.1
On dit que f admet un minimum local (respectivement un maximum local) en $x \in U$ s'il existe un voisinage V de x tel que $f(y) \geq f(x)$ (respectivement, $f(y) \leq f(x)$) pour tout $y \in V$.

Un maximum (global) serait quand $f(y) \leq f(x)$ pour tout $y \in U$. Quand f a des dérivées partielles, la recherche des extrema (minima et maxima) locaux est plus facile, à cause de ce qui suit.

Proposition 2.3.2
Si x est un extremum local de f dans U, et si la dérivée partielle $\frac{\partial f}{\partial x_j}(x)$ existe, alors $\frac{\partial f}{\partial x_j} = 0$.

Pour le démontrer, on suppose que f a un extremum local, on se donne V comme ci-dessus, et on se donne $r > 0$ tel que $B(x, r) \subset V$. Ensuite on définit g sur $I =]-r, r[$ par $g(t) = f(x + te_j)$, où e_j est le j-ième vecteur de base. Alors g admet un extremum en 0 (de la même nature que x lui-même). De plus, g est dérivable en 0, avec $g'(0) = \frac{\partial f}{\partial x_j}$ (c'est la définition de $\frac{\partial f}{\partial x_j}$). Il reste donc à voir que $g'(0) = 0$.

Supposons que non ; supposons que g a un maximum local en 0 et que $g'(0) > 0$, et obtenons une contradiction. Les autres cas seraient pareils.

On a le développement limité suivant de g au voisinage de 0 :

$$g(t) = g(0) + tg'(0) + t\varepsilon(t)\,,$$

avec $\lim_{t\to 0} \varepsilon(t) = 0$.
Pour $t > 0$ assez petit, $|\varepsilon(t)| < g'(0)/2$, et donc $g(t) - g(0) = t[g'(0) + \varepsilon(t)] > 0$, ce qui contredit la maximalité de $g(0)$. □

Proposition 2.3.3
Si x est un minimum local de f dans U, et si f a des dérivées partielles jusqu'à l'ordre 2 qui sont continues en x, alors pour tout vecteur $\xi \in \mathbb{R}^n$,

$$\sum_{j=1}^{n} \sum_{k=1}^{n} \xi_j \xi_k \frac{\partial^2 f}{\partial x_j \partial x_k}(x) \geq 0. \tag{2.30}$$

Sinon, on choisit $\xi \in \mathbb{R}^n$ tel que le membre de gauche de (2.24) (appelons-le $Q(\xi)$) soit < 0, on applique le développement limité (2.24) avec $z = t\xi$, t petit, en tenant compte du fait que l'on sait par le lemme 2.3.2 que les termes linéaires sont nuls, et on obtient que

$$0 \leq f(x) - f(x + t\xi) \leq t^2 Q(\xi) + o(t^2)\,,$$

une contradiction. □

Bien sûr on a le même résultat pour un maximum local, avec cette fois la condition nécessaire $Q(\xi) \leq 0$. Il suffit en effet d'appliquer le lemme à $-f$.

Remarque 2.3.4
On sait qu'une matrice symétrique a toutes ses valeurs propres réelles. Si on désigne par Hess f(x) la matrice constituée des dérivées secondes de f au point x, de sorte que

$$Q(\xi) = \langle \xi\,,\, \mathit{Hess}\, f(x)\, \xi \rangle\,,$$

la condition précédente peut se réécrire sous la forme :
Les valeurs propres de la matrice Hess $f\,(x)$ sont positives ou nulles.
Il suffit en effet de prendre dans la condition un vecteur propre de Hess $f(x)$.

Comme dans le cas de la proposition 2.1.7, la proposition peut être complétée par la proposition suivante

Proposition 2.3.5
Si f a des dérivées partielles jusqu'à l'ordre 2 qui sont continues en x, et si pour tout vecteur $\xi \in \mathbb{R}^n \setminus \{0\}$,

$$\sum_{j=1}^{n} \sum_{k=1}^{n} \xi_j \xi_k \, \frac{\partial^2 f}{\partial x_j \partial x_k}(x) > 0\,, \tag{2.31}$$

alors x est un minimum local de f.

Remarque 2.3.6
La condition précédente peut se réécrire sous la forme :
Les valeurs propres de la matrice Hess $f\,(x)$ sont positives.

Une condition nécessaire

On reprend la situation de la proposition 2.3.3. Alors on a, en un point minimum de f :

$$\frac{\partial^2 f}{\partial x_j^2}(x) \geq 0\,,\, \forall j \in \{1, \cdots, n\}\,. \tag{2.32}$$

Il suffit en effet de regarder la variation dans la direction x_j. Mais cette condition, même satisfaite pour tout j n'est pas suffisante !

2.3.2 A la recherche d'extrema sur un exemple

Parlons maintenant de la recherche d'extrema. On va le faire sur un exemple si simple que d'ailleurs on pourrait deviner les résultats sans calculer, mais le processus est bien général.

On va toujours prendre la même fonction f, définie par

$$f(x,y) = 3x^2y^3, \tag{2.33}$$

mais éventuellement sur des domaines de définition différents. Noter qu'en tout cas f est de classe C^1, avec les dérivées partielles

$$\frac{\partial f}{\partial x}(x,y) = 6xy^3 \quad \text{et} \quad \frac{\partial f}{\partial y}(x,y) = 9x^2y^2. \tag{2.34}$$

On commence avec le domaine $A = \overline{B}(0,1)$. Disons qu'on cherche à déterminer les bornes inférieure et supérieure de f (c'est-à-dire, des valeurs de $f(x,y)$ pour $(x,y) \in A$), et les points de A où ces bornes sont atteintes.

Notons

$$m = \inf_{(x,y)\in A} f(x,y) \text{ et } M = \sup_{(x,y)\in A} f(x,y)$$

ces bornes.

Remarquons qu'il y a au moins un point $(x_0,y_0) \in A$ tel que $f(x_0,y_0) = m$, et un point $(x_1,y_1) \in A$ tel que $f(x_1,y_1) = M$, puisque la fonction f est continue sur le compact A, donc atteint ses bornes.

On commence à chercher (x_0,y_0) et (x_1,y_1) dans l'intérieur de A. La proposition 2.3.2 dit qu'alors ce point (appelons-le $z = (x,y)$) est un point critique, donc (par (2.34)) devrait satisfaire $6xy^3 = 9x^2y^2 = 0$, donc $x = 0$ ou $y = 0$. Mais alors $f(z) = 0$.

On se rend compte aisément que $m < 0$ et $M > 0$. On en conclut donc qu'il n'y a pas d'extremum global à l'intérieur de A.

Si l'on était seulement à la recherche d'extrema locaux, on pourrait essayer de poursuivre sans se fatiguer en appliquant la proposition 2.3.3. En effet, f est de classe C^2, avec les dérivées partielles

$$\frac{\partial^2 f}{\partial x^2}(x,y) = 6y^3\,,\ \frac{\partial^2 f}{\partial x\partial y}(x,y) = 18xy^2 \quad \text{et} \quad \frac{\partial^2 f}{\partial y^2}(x,y) = 18x^2y. \tag{2.35}$$

Donc dans le cas où $y = 0$, $Q(\xi) = 0$ pour tout ξ et le lemme 2.3.3 ne donne aucune information. Quand $x = 0$ et $y \neq 0$, $Q(\xi) = 6y^3\xi_1^2$, et le lemme dit que

$z = (x, y)$ ne peut pas être un maximum local quand $y > 0$, ni un minimum local quand $y < 0$.

Et en fait, le mieux dans ce cas est de regarder directement ce qui se passe.

Quand $x = 0$ et $y > 0$, on voit que les valeurs de f autour de z sont ≥ 0, donc on a un minimum local. Quand $x = 0$ et $y < 0$, les valeurs de f autour de z sont ≤ 0, donc on a un maximum local.

Quand $y = 0$, on voit aisément que f prend des valeurs strictement positives et des valeurs strictement négatives dans tout voisinage de z, donc z n'est pas un extremum local.

Revenons à notre quête d'extrema globaux.
On a vu qu'il n'y en a pas dans la boule ouverte, donc qu'ils sont tous à la frontière.

De plus, on sait qu'il y a forcément au moins un minimum et un maximum globaux dans A (donc sur la frontière). On en déduit également que :

$$m = \inf_{(x,y)\in\partial A} f(x, y) \text{ et } M = \sup_{(x,y)\in\partial A} f(x, y)\,.$$

Ici, on a de la chance, le bord est le cercle unité, et se paramètre aisément. On note que tout point du bord peut être écrit $(x, y) = (\cos\theta, \sin\theta)$, et qu'alors $f(x, y) = 3\cos^2\theta\,\sin^3\theta$. On cherche maintenant les extrema de cette nouvelle fonction de θ, sachant qu'ils donneront les extrema de f sur ∂A, puis donc sur A. [Petite vérification à faire, en disant que l'application $\theta \to (\cos\theta, \sin\theta)$ est surjective, donc que les deux bornes supérieures ou inférieures sont les mêmes.]

Il reste à dériver $g(\theta) = 3\cos^2(\theta)\sin^3(\theta)$; on trouve

$$g'(\theta) = -6\cos(\theta)\sin^4(\theta) + 9\cos^3(\theta)\sin^2(\theta)\,,$$

qui s'annule quand

$$2\cos(\theta)\sin^4(\theta) = 3\cos^3(\theta)\sin^2(\theta)\,.$$

C'est le cas quand $\cos\theta = 0$ ou $\sin\theta = 0$ (ce qui d'ailleurs correspond aux points sur les axes qu'on avait déjà repérés), mais comme alors on trouve $f(x, y) = g(\theta) = 0$ comme plus haut, cette solution ne donne pas d'extremum. Donc on peut supposer que $\cos\theta$ et $\sin\theta$ sont non nuls, on peut simplifier, et il reste

$$2\sin^2\theta = 3\cos^2\theta\,.$$

On n'a pas vraiment besoin d'essayer de calculer θ, puisqu'il nous faut seulement (x, y). On note que sur le cercle, la condition se réécrit

$$2y^2 = 3x^2 \,,$$

qui conduit aux équations

$$\sqrt{2}y = \pm\sqrt{3}x$$

qui ont quatre solutions.
En examinant les signes de f sur le cercle et en utilisant la parité de f (par $x \mapsto -x$) et l'imparité de f (par $y \mapsto -y$), on obtient que nécessairement f a deux maxima globaux en $(\pm\sqrt{2/5}, \sqrt{3/5})$ et deux minima globaux en $(\pm\sqrt{2/5}, -\sqrt{3/5})$. On trouve alors que

$$M = -m = \frac{18}{25}\sqrt{3/5}\,.$$

La situation se compliquerait un peu si l'on n'avait pas de paramétrage sympathique de ∂A. Supposons cependant qu'on ait pris un domaine A dont le bord ∂A est une courbe qui a une tangente partout, et qu'on soit comme plus haut en train de se demander quels points z de ∂A peuvent être des extrema locaux. Soit z un tel point, et soit τ un vecteur unitaire de la tangente à ∂A en z. En faisant un développement limité de f au voisinage de A et en se restreignant à ∂A, on démontre un peu comme dans la proposition 2.3.2 qu'il faut que

$$Df(z) \cdot \tau = 0\,.$$

A nouveau, ceci donne une indication supplémentaire sur qui peut bien être un point z. Et vous verrez sans doute plus tard des moyens de trouver une tangente à ∂A sans nécessairement en avoir un paramétrage.

2.4 Applications

2.4.1 L'équation de Laplace

On revient au Laplacien (voir Sous-Section 2.2.6) et à l'équation de Laplace (on se limite au cas de deux variables) associée :

$$\Delta u(x, y) = f(x, y), \tag{2.36}$$

où Δ désigne l'opérateur aux dérivées partielles $\Delta = \partial^2_{xx} + \partial^2_{yy}$, et f est une fonction continue donnée.

Cette équation est très importante à la fois en physique et en mathématiques.

On rappelle qu'une fonction harmonique dans un ouvert Ω, une fonction de classe C^2 dans Ω telles que $\Delta u = 0$ dans Ω. Comme exemple de fonctions harmoniques on peut citer les fonctions

$$(x, y) \mapsto c\,,\ (x, y) \mapsto \alpha x + \beta y\,,\ (x, y) \mapsto \gamma xy + \delta(x^2 - y^2)\,.$$

Rappelons aussi qu'on a vu que quand Ω ne contient pas $(0, 0)$ $(x, y) \mapsto \log(x^2 + y^2)$ est une fonction harmonique.

Principe du Maximum

Théorème 2.4.1
Soit D un ouvert borné et convexe de $\mathbb{R}^2$. Soit u fonction harmonique dans D et continue dans $\bar{D}$. Alors le maximum de u dans $\bar{D}$ est atteint sur le bord de D.

Preuve :
Si, pour $\epsilon > 0$, on pose $v_\epsilon(x, y) = u(x, y) + \epsilon(x^2 + y^2)$, on a

$$\Delta v_\epsilon(x, y) = \Delta u(x, y) + \epsilon\Delta(x^2 + y^2) = 0 + 4\epsilon > 0.$$

D'un autre côté, si (x_ϵ, y_ϵ) est un maximum pour v_ϵ dans D, on a, par exemple en utilisant la formule de Taylor,

$$\Delta v_\epsilon(x_\epsilon, y_\epsilon) = \partial^2_{xx} v_\epsilon(x_\epsilon, y_\epsilon) + \partial^2_{yy} v_\epsilon(x_\epsilon, y_\epsilon) \leq 0,$$

ce qui est absurde.
Or, puisque v_ϵ est continue, elle doit atteindre son maximum sur le compact $\bar{D}$. Elle l'atteint donc en un point (x_ϵ, y_ϵ) du bord ∂D. On a donc, pour tout (x, y),

$$u(x, y) \leq v_\epsilon(x, y) \leq v_\epsilon(x_\epsilon, y_\epsilon) = u(x_\epsilon, y_\epsilon) + \epsilon(x_\epsilon^2 + y_\epsilon^2) \leq \max_{\partial D} u + \epsilon r^2,$$

où $r > 0$ est choisi pour que le disque centré à l'origine de rayon r contienne D. Ceci étant vrai pour tout $\epsilon > 0$, on obtient, pour tout $(x, y) \in \bar{D}$,

$$u(x, y) \leq \max_{(x,y)\in\partial D} u(x, y).$$

□

Encore une fois, ce théorème dit que la fonction harmonique u, supposée aussi continue sur le compact $\bar{D}$, atteint son maximum au moins un en point

du bord de D.
Ceci n'interdit pas que u atteigne aussi son maximum en un point de D. Par exemple si u est constante.
On peut déduire du principe du maximum l'unicité de la solution du problème de Dirichlet pour le Laplacien.

Proposition 2.4.2
Soit D un ouvert borné et convexe de $\mathbb{R}^2$. Soit f une fonction continue sur $\bar{D}$, et g une fonction continue sur ∂D. Le problème de Dirichlet

$$\begin{cases} \Delta u(x,y) = f(x,y), & (x,y) \in D \\ u(x,y) = g(x,y), & (x,y) \in \partial D \end{cases} \tag{2.37}$$

admet au plus une solution $\mathcal{C}^2$.

Preuve : Si u_1 et u_2 sont solutions de (2.37), $w = u_1 - u_2$ vérifie $\Delta w = 0$ et est nulle sur ∂D. Par le principe du maximum (et celui du minimum) on obtient $w = 0$ sur D. □

2.4.2 Extrema liés

Le théorème principal est le suivant :

Théorème 2.4.3
Soit D un domaine ouvert de $\mathbb{R}^n$ et $f, g_1, \ldots, g_k$ $(k \leq n-1)$ des applications de classe C^1 de D dans $\mathbb{R}$. On pose

$$A := \{x \in D \,,\, g_1(x) = \ldots = g_k(x) = 0\} \,.$$

Si la restriction de f à A présente un extremum au point $a = (a_1, \cdots, a_n)$ de A, et si les vecteurs $\nabla g_1(a) \,, \ldots \,, \nabla g_k(a)$ sont linéairement indépendants, alors il existe k réels $\lambda_1, \cdots, \lambda_k$ tels que :

$$\nabla f(a) = \sum_{j=1}^{k} \lambda_j \nabla g_j(a) \,.$$

On note que ce théorème ne donne que des conditions nécessaires pour avoir un extremum. Les λ_j sont appelés les multiplicateurs de Lagrange.

Idée de la démonstration
On va admettre que l'on peut se ramener au cas où :

$$g_\ell(x) = x_\ell - \phi_\ell(x'') \,,\, \ell = 1, \cdots, k \,,$$

où on a noté $x = (x', x'')$, $x' = (x_1, \cdots, x_k)$, $x'' = (x_{k+1}, \cdots, x_n)$ et où les ϕ_ℓ sont de classe C^1.
Cette propriété résulte d'un théorème qui ne figure pas au programme de ce cours (théorème d'inversion locale).

Pour simplifier, on traite d'abord le cas $k = 2$, $n = 3$.
On doit alors regarder la fonction $x_3 \mapsto F(x_3) := f(\phi_1(x_3), \phi_2(x_3), x_3)$ et chercher les extrema de F. Exprimons la condition nécessaire présentée à la proposition 2.3.2. La condition au pont (a_1, a_2, a_3) s'exprime sous la forme $F'(a_3) = 0$, soit

$$\partial_{x_1} f(a)\, \phi_1'(a_3) + \partial_{x_2} f(a)\, \phi_2'(a_3) + \partial_{x_3} f(a) = 0\,.$$

Ceci exprime l'orthogonalité dans $\mathbb{R}^3$ de $\nabla f(a)$ avec le vecteur $(\phi_1'(a_3), \phi_2'(a_3), 1)$. Autrement dit le vecteur $\nabla f(a)$ est dans le plan orthogonal à $(\phi_1'(a_3), \phi_2'(a_3), 1)$ dont une base est donnée par les vecteurs $\nabla g_1(a)$ et $\nabla g_2(a)$. (Dériver par rapport à x_3 les identités $g_1(\phi_1(x_3), \phi_2(x_3), x_3) = 0$ et $g_2(\phi_1(x_3), \phi_2(x_3), x_3) = 0$).

Pour compléter, on traite aussi le cas $k = 1$, $n = 3$.
On doit alors regarder la fonction $(x_2, x_3) \mapsto G(x_2, x_3) := f(\phi(x_2, x_3), x_2, x_3)$ et chercher les extrema de G. Exprimons la condition nécessaire présentée à la proposition 2.3.2. La condition s'exprime sous la forme $\nabla G(a_2, a_3) = 0$, soient deux équations :

$$\begin{aligned} \partial_{x_1} f(a)\, \partial_{x_2}\phi(a_2, a_3) + \partial_{x_2} f(a) &= 0\,, \\ \partial_{x_1} f(a)\, \partial_{x_3}\phi(a_2, a_3) + \partial_{x_3} f(a) &= 0\,. \end{aligned}$$

Ceci exprime l'orthogonalité dans $\mathbb{R}^3$ de $\nabla f(a)$ avec les vecteurs

$$\vec{v_2} := (\partial_{x_2}\phi(a_2, a_3), 1, 0) \text{ et } \vec{v_3} := (\partial_{x_2}\phi(a_2, a_3), 0, 1)\,.$$

Autrement dit le vecteur $\nabla f(a)$ est orthogonal au plan engendré par $\vec{v_2}$ et $\vec{v_3}$. Mais $\nabla g(a)$ est aussi non nul et orthogonal à ce plan. On en déduit (écrire cette fois-ci $g(\phi(x_2, x_3), x_2, x_3) = 0$ et prendre les dérivées partielles par rapport à x_2 et x_3) que $\nabla f(a)$ et $\nabla g(a)$ sont parallèles : $\nabla f(a) = \lambda\, \nabla g(a)$.
Le cas général nécessite une bonne maîtrise des notations et un minimum de géométrie euclidenne dans $\mathbb{R}^n$.

Une application.
Parmi les parallélépipèdes de surface S fixée, lesquels ont un volume maximal ?

Si x, y, z désignent les longueurs des côtés du parallèlépipède, la surface est $2(xy + yz + xz)$ et le volume xyz.

Avant de traiter le problème on peut observer qu'il y a des parallélépipèdes de surface S de volume arbitrairement petit (prendre $z_n = \frac{1}{n}$, $y_n = 1$ et $x_n = (\frac{S}{2} - \frac{1}{n})/(1 + \frac{1}{n})$).
Le problème est de trouver le maximum atteint par le volume xyz en ayant la contrainte

$$2(xy + yz + xz) = S\,.$$

Bien sûr on peut d'abord utiliser l'équation pour obtenir

$$z = \frac{\frac{S}{2} - xy}{x + y}\,.$$

On a alors à maximiser

$$(x, y) \mapsto V_S(x, y) = xy\frac{\frac{S}{2} - xy}{x + y}\,.$$

On va d'abord écrire que l'on a comme condition nécessaire :

$$\partial_x V_S(x, y) = 0\,,\ \partial_y V_S(x, y) = 0\,,$$

avec $x > 0$ et $y > 0$.
On trouve finalement que $x = y = \sqrt{\frac{S}{6}}$, ce qui implique aussi $z = \sqrt{\frac{S}{6}}$, et on trouve (attention, il faudrait vérifier que cet extremum est un maximum), qu'à surface fixée, le parallélépipède de volume maximal est le cube.

Voyons maintenant comment on peut appliquer directement le théorème 2.4.3. On est dans le cas $n = 3$, $k = 1$ (voir plus haut). On est donc dans le cas où

$$f(x, y, z) = xyz\,,\ g(x, y, z) = 2(xy + yz + xz) - S\,.$$

Calculons les gradients, on a :

$$\nabla f = \begin{pmatrix} yz \\ xz \\ xy \end{pmatrix}\,,\ \nabla g = \begin{pmatrix} 2(y + z) \\ 2(x + z) \\ 2(x + y) \end{pmatrix}$$

D'après le théorème, si (x, y, z) est un extremum, il existe λ tel que

$$\nabla f(x, y, z) = \lambda \nabla g(x, y, z)\,,\ g(x, y, z) = 0\,.$$

On trouve les équations :

$$\begin{aligned} yz &= 2\lambda(y+z)\,, \\ xz &= 2\lambda(x+z)\,, \\ xy &= 2\lambda(x+y)\,. \end{aligned}$$

En soustrayant la deuxième équation à la première, on trouve :

$$(y-x)z = 2\lambda(y-x)\,,$$

soit encore

$$(x-y)(z-2\lambda) = 0\,.$$

En procédant de même avec la deuxième et la troisième, puis entre la première et la troisième, on obtient

$$(z-y)(x-2\lambda) = 0 \text{ et } (x-z)(y-2\lambda) = 0\,.$$

(attention, ces trois équations ne sont pas équivalentes aux trois premières et il ne faut pas oublier que l'on cherche une solution dans $g^{-1}(0)$).

On trouve immédiatement une première solution $x = y = z$. On peut vérifier qu'il n'y a pas d'autre solution telles que $g(x,y,z) = 0$ et pour cette solution, on trouve $6x^2 = S$.

Remarque 2.4.4
On a un peu triché dans l'analyse ci-dessus. On a juste trouvé un candidat pour un extremum local. Il n'est en effet pas exclu que les extrema soient au bord. On minimise en effet sur $\{x > 0\,,\, y > 0\,,\, z > 0\,, g(x,y,z) = 0\}$ *et il faut donc vérifier ce qui se passe au bord.*

Remarque 2.4.5
Par la même méthode, on peut répondre à une question très voisine :
Parmi les parallélépipèdes de volume V *fixé, lesquels ont une surface minimale ?*

L'exercice suivant correspond au cas $n = 3$ et $k = 2$.

Exercice 2.4.6
Etudier les extrema de $f(x,y,z) = xyz$ *avec les contraintes :*
$g_1(x,y,z) = x^2 + y^2 + z^2 - 1$ *et* $g_2(x,y,z) = x + y + z - 1$.

Corrigé
Il y a 6 solutions $(1,0,0)$, $(0,1,0)$, $(0,0,1)$, $(-\frac{1}{3}, \frac{2}{3}, \frac{2}{3})$, $(\frac{2}{3}, -\frac{1}{3}, \frac{2}{3})$ et $(\frac{2}{3}, \frac{2}{3}, -\frac{1}{3})$.

2.5 Exercices supplémentaires

1. Déterminer les points critiques de chacune des applications suivantes et donner leur développement limité à l'ordre 2 en chacun de leurs points critiques. Existe-t-il des extrema locaux ? globaux ?
   1. $f(x, y, z) = x^2/2 + xyz - y + z$. 2. $f(x, y) = x^4 + y^4 - 4(x - y)^2$.

2. Etudier les points critiques des fonctions suivantes :

   a) $f_1(x, y) = x^2 - y^3$ b) $f_2(x, y) = x + y + x^2 - xy + y^2 + 1$
   c) $f_3(x, y) = xy + yz + xz + xyz$ d) $f_4(x, y, z) = x^2 + z^2 + x^2y$
   e) $f_5(x, y) = x^2 + 2y^2 + \pi \cos x \cos y$

3. Soit f_1 et f_2 les deux fonctions de $\mathbb{R}^2$ dans $\mathbb{R}$ définies par $f_1(x, y) = y + 2 - (x + 1)^2$ et $f_2(x, y) = y - 2 + (x - 1)^2$. On pose $f = f_1 f_2$. Déterminer les points critiques de f, et déterminer pour chacun d'eux s'il s'agit d'un extremum.

4. Soit $f(x, y) = x^3 + 3x^2y - 15x - 12y$. Trouver les extrema locaux de f. Lesquels sont des minima, lesquels sont des maxima ?

5. Soit a un réel positif. Quels sont les extrema locaux de $f : (x, y) \mapsto f(x, y) = x^2y^2 + x^2 + y^2 + 2axy$?

Exercice 2.5.1 *Montrer que la fonction $u(x, y) = (1 - x^2 - y^2)/(1 - 2x + x^2 + y^2)$ est harmonique dans le disque $x^2 + y^2 < 1$. Le principe du maximum est-il vérifié dans $D = \{x^2 + y^2 < 1\}$ pour cette fonction ?*

Corrigé : u n'est pas définie en $(1,0)$, et ne se polonge pas par continuité en ce point puisque

$$\lim_{x\to 1^-} u(x,0) = \lim_{x\to 1^-} \frac{1-x^2}{(1-x)^2} = \lim_{x\to 1^-} \frac{1+x}{1-x} = +\infty.$$

On a

$$\partial_x u(x,y) = 2\frac{(x-1)^2 - y-^2}{((x-1)^2 + y-^2)^2}, \partial_y u(x,y) = 4\frac{y(x-1)}{((x-1)^2 + y-^2)^2},$$

et donc

$$\partial^2_{xx} u(x,y) = -4\frac{-1+3x-3x^2+3y^2+x^3-3xy^2}{(1-2x+x^2+y^2)^3} = -\partial^2_{yy} u(x,y),$$

ce qui montre que u est harmonique. On peut voir également que u n'a pas de point critique dans D, mais de toutes façons, on a bien

$$u(x,y) \leq \max_{(x,y)\in\partial D} u(x,y) = +\infty.$$

Exercice 2.5.2
Enoncer et démontrer un principe du minimum.

Corrigé : Si u est harmonique dans D et continue sur $\bar{D}$, $v = -u$ aussi. Le principe du maximum dit que, pour tout $(x,y) \in \bar{D}$,

$$v(x,y) \leq \max_{(x,y)\in\partial D} v(x,y),$$

c'est-à-dire

$$u(x,y) \geq -\max_{(x,y)\in\partial D} -u(x,y) = \min_{(x,y)\in\partial D} u(x,y).$$