Fonctions continues d'une variable réelle : Compléments.
Notes de cours.

Table des matières

Ces notes abordent trois notions : compacité, continuité uniforme (et ses liens étroits avec la compacité) et convexité.

1 Compacité et continuité

Rappelons qu'une partie K d'un espace topologique X est dite **séquentiellement compacte** si toute suite à valeurs dans K admet une valeur d'adhérence qui appartient à K. Si X est un espace métrique, alors cette définition coïncide avec la définition plus générale de partie compacte d'un espace topologique (théorème de Bolzano-Weierstrass).

La compacité a de nombreuses applications, que nous ne mentionnerons pas toutes ici. Notamment, nous ne parlerons pas :

- ▷ Des applications à la continuité uniforme (théorème de Heine, d'Arzelà-Ascoli), qui seront traitée en Section 2.
- ▷ Des applications à l'analyse fonctionnelle (compacité en dimension infinie, opérateurs compacts), hors le théorème d'Arzelà-Ascoli déjà mentionné.
- ▷ Des applications à l'algèbre (décomposition d'Iwasawa, sous-groupes compacts).
- ▷ De la notion de *convergence uniforme* (ou $\mathcal{C}^1$, ou $\mathcal{C}^k$) *sur tout compact*, précieuse mais mieux à sa place dans des chapitres sur les séries de fonctions ou les intégrales à paramètres.

1.1 Images directes

La propriété la plus importante des espaces compacts est

Théorème 1.
Soient K un espace métrique compact, Y un espace métrique et $f : K \to Y$ une application continue. Alors $f(K)$ est compact.

Ce théorème a ceci de spécifique que beaucoup de propriétés topologiques se comportent bien vis-à-vis des images inverses ; par exemple, l'image inverse d'un ouvert par une application continue est un ouvert. Cette propriété (comme la connexité) se comporte bien par image directe.

Démonstration.
Soit $(y_n)_{n\in\mathbb{N}}$ une suite à valeurs dans $f(K)$. Par définition, il existe une suite $(x_n)_{n\in\mathbb{N}}$ à valeurs dans K telle que $fx_n) = y_n$ pour tout n. Comme K est compact, il existe une sous-suite extraite $(x_{n_k})_{k\in\mathbb{N}}$ convergeant vers un point $x \in K$. Par continuité, $(f(x_{n_k}))_{k\in\mathbb{N}}$ converge vers $f(x) \in f(K)$. Mais $f(x_{n_k}) = y_{n_k}$ pour tout k, donc la sous-suite extraite $(y_{n_k})_{k\in\mathbb{N}}$ converge vers $f(x) \in f(K)$.

Toute suite à valeurs dans $f(K)$ admet une valeur d'adhérence dans $f(K)$, donc $f(K)$ est compact. □

Le corollaire suivant est l'application la plus fréquente de ce théorème. *Il est utilisable dans la grande majorité des sujets d'écrit d'analyse, et ses hypothèses sont à vérifier systématiquement.*

Corollaire 1.1.
Soient K un espace métrique compact et $f : K \to \mathbb{R}$ une application continue. Alors f est bornée et atteint ses bornes.

Démonstration.
Par le théorème précédent, $f(K)$ est compact. Or les compacts de $\mathbb{R}$ sont les fermés bornés, donc $f(K)$ est borné, donc f est bornée. De plus, $f(K)$ est fermé, donc $\sup f(K)$ et $\inf f(K)$ (qui sont bien définis car f est bornée) appartient à $f(K)$, donc il existe $x, y \in K$ tels que $f(x) = \max_K f = \sup_K f$ et $f(y) = \min_K f = \inf_K f$. □

On peut aussi affiner légèrement le théorème des valeurs intermédiaires.

Corollaire 1.2.
Soient $I \subset \mathbb{R}$ un segment et $f : I \to \mathbb{R}$ une application continue. Alors $f(I)$ est un segment.

Démonstration.
$f(I)$ est un intervalle par le théorème des valeurs intermédiaires, et $f(I)$ est fermé et borné dans $\mathbb{R}$ car compact, donc $f(I)$ est un segment. $\square$

Exercice 1.3.
Soit I un intervalle ouvert ou semi-ouvert et $f : I \to \mathbb{R}$ une application continue. Alors $f(I)$ peut être ouvert, semi-ouvert ou fermé. Illustrez chacune des 6 possibilités.

1.2 Parties compactes de $\mathbb{R}^n$

Pour pouvoir utiliser le théorème précédent, il reste à savoir ce qu'est un compact. Dans $\mathbb{R}$, nous savons déjà que ce sont les fermés bornés.

Proposition 1.4.
Soit K un compact d'un espace métrique. Alors K est fermé et borné.

La démonstration en est la même que dans le cas réel. La réciproque est en général fausse.

Proposition 1.5.
Soit K un compact d'un espace métrique. Soit $F \subset K$ un fermé. Alors F est compact.

Démonstration.
Toute suite à valeurs dans F est aussi à valeurs dans K. Par compacité de K, elle admet une valeur d'adhérence dans K. Comme F est fermé, cette valeur d'adhérence est dans F. $\square$

Proposition 1.6.
Soient K_1, K_2 deux compacts. Alors $K_1 \times K_2$ est compact.

Remarquons que l'on a défini la compacité dans le cadre d'espaces métriques. Ainsi, dans cet énoncé, K_1 et K_2 sont des compacts de deux espaces métriques (X_1, d_1) et (X_2, d_2) ; l'ensemble $K_1 \times K_2$ est un compact de $X_1 \times X_2$, qui est métrique par exemple pour la distance $\max\{d_1, d_2\}$.

Démonstration.
Soit $(x_n, y_n)_{n\in\mathbb{N}}$ une suite à valeurs dans $K_1 \times K_2$. Par compacité, il existe une sous-suite extraite $(x_{n_k})_{k\in\mathbb{N}}$ convergeant vers $x \in K_1$. Par compacité, on peut ré-extraire une sous-suite $(y_{n_{k_\ell}})_{\ell\in\mathbb{N}}$ de $(y_{n_k})_{k\in\mathbb{N}}$ convergeant vers $y \in K_2$. Mais alors $(x_{n_{k_\ell}}, y_{n_{k_\ell}})_{\ell\in\mathbb{N}}$ converge vers $(x, y) \in K_1 \times K_2$. $\square$

Remarque 1.7.
Par récurrence, tout produit fini de compacts est compact. Plus généralement, le théorème de Tychonoff affirme que tout produit de compact est compact. Ce théorème utilise l'axiome du choix. De plus, une utilisation effective de ce théorème nécessite de comprendre la topologie produit pour des produits infinis, ce qui peut être délicat.

Grâce à ces deux propositions, on peut caractériser les compacts des espaces vectoriels réels (ou complexes) de dimension finie.

Théorème 2 (Heine-Borel)**.**
Soit $n \geq 0$. Les compacts de $(\mathbb{R}^n, \|\cdot\|_\infty)$ (ou $\mathbb{C}^n$) sont les fermés bornés.

Démonstration.
On sait déjà que les compacts sont fermés bornés ; il reste à montrer que les fermés bornés sont compacts. Soit K un fermé borné de $\mathbb{R}^n$.

Comme K est borné pour $\|\cdot\|_\infty$, il existe $M \geq 0$ tel que $K \subset [-M, M]^n$. Or chaque segment $[-M, M]$ est compact, un produit de compacts est compact, et $[-M, M]^n \subset \mathbb{R}^n$ est bien muni de la topologie produit, donc $[-M, M]^n$ est compact. De plus, K est fermé, et un fermé d'un compact est compact, donc K est compact. □

En dimension infinie, la situation est très différente : la plupart des fermés bornés ne sont alors pas compacts. Par exemple, dans tout espace vectoriel réel normé de dimension infinie, la boule unité fermée n'est pas compacte ! Un autre exemple est la suite de fonctions $f_n : x \mapsto x^n$ dans $\mathcal{C}([0, 1], \mathbb{R})$, qui est bornée mais n'admet pas de sous-suite convergente dans $\mathcal{C}([0, 1], \mathbb{R})$ (en particulier, son adhérence est un fermé borné non compact).

Nous renvoyons la lectrice à la Sous-section 2.5 pour des exemples de parties fermées dans des espaces vectoriels de dimension infinie.

1.3 Application 1 : Équivalence de normes en dimension finie

Théorème 3 (Équivalence de normes)**.**
Soit E un espace vectoriel réel ou complexe de dimension finie. Soient N, N' deux normes sur E. Alors il existe $C \geq 1$ tel que, pour tout $x \in E$,

$$\frac{1}{C} N(x) \leq N'(x) \leq C N(x).$$

On dit aussi que N et N' sont équivalentes.

Démonstration.
Soit $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E. Soit $\|\cdot\|_\infty$ la norme du supremum des coordonnées. La notion d'équivalence étant une relation d'équivalence sur les normes, par transitivité, il suffit de montrer que toute norme N est équivalente à $\|\cdot\|_\infty$.

Soit $x = \sum_{i=1}^n X_i e_i \in E$. Alors $N(x) \leq \sum_{i=1}^n |X_i| N(e_i)$ par l'inégalité triangulaire et l'homogénéité de N. Donc

$$N(x) \leq n \max_{1 \leq i \leq n} \{|X_i| N(e_i)\} \leq n \max_{1 \leq i \leq n} \{|X_i|\} \cdot \max_{1 \leq i \leq n} \{N(e_i)\} = n \max_{1 \leq i \leq n} \{N(e_i)\} \|x\|_\infty .$$

D'une part, on a majoré $N(x)$ par $C \|x\|_\infty$ pour une certaine constante C, ce qui est la moitié du résultat voulu. D'autre part, $N(x - y) \leq C \|x - y\|_\infty$ pour tous x, $y \in E$, donc N est lipschitzienne, et en particulier continue, pour la norme $\|\cdot\|_\infty$.

Soit $\mathbb{S}_E := \{x \in E : \|x\|_\infty = 1\}$. Alors $\mathbb{S}_E$ est fermé borné pour $\|\cdot\|_\infty$, donc compact pour la topologie engendrée par cette norme. Donc N atteint son minimum sur $\mathbb{S}_E$. Or $N(x) > 0$ pour tout $x \in \mathbb{S}_E$, donc ce minimum est strictement positif. Il existe donc une constance $c > 0$ telle que $N(x) \geq c$ pour tout $x \in \mathbb{S}_E$.

Finalement, pour tout $x \in E \setminus \{0\}$, on a $N(x) = \|x\|_\infty N(x / \|x\|_\infty) \geq c \|x\|_\infty$, et $N(0) \geq c \|0\|_\infty$. Ceci fournit la deuxième inégalité souhaitée. □

Remarque 1.8.
Ce théorème est faux en dimension infinie. Nous verrons plus tard le contre-exemple des normes $\|\cdot\|_1$, $\|\cdot\|_2$ et $\|\cdot\|_\infty$ sur $\mathcal{C}([0, 1], \mathbb{R})$.

Remarque 1.9.
Le théorème de Heine-Borel avait été énoncé pour la norme $\|\cdot\|_\infty$ sur $\mathbb{R}^n$. D'après ce théorème, il est valable pour toute norme sur tout $\mathbb{R}$-espace vectoriel de dimension finie. On l'énonce donc plus fréquemment sous la forme "Les compacts de $\mathbb{R}^n$ sont les fermés bornés", sans préciser de norme, car la notion de partie bornée ne dépend pas de la norme en dimension finie.

Exemple 1.10.
Les groupes $O(n)$ et $SO(n)$ sont fermés et bornés dans $\mathcal{M}_n(\mathbb{R})$, donc compacts.

1.4 Application 2 : Théorème de Stone-Weirstrass

Le théorème de Stone-Weierstrass joue un rôle important en analyse. En particulier son application aux polynômes trigonométriques jouera un rôle clef en analyse de Fourier.

1.4.1 Sous-algèbres et théorème de Stone-Weierstrass

Définition 1.11.
*Une **algèbre** est un espace vectoriel A muni d'une opération de multiplication $* : A \times A \to A$ qui soit bilinéaire et associative.*

*Une **sous-algèbre** est un sous-espace vectoriel B stable par $*$: pour tous $x,\ y \in B$, on a $x * y \in B$.*

Les exemples clefs sont les espaces de fonctions $\mathcal{C}(X, \mathbb{R})$, $\mathcal{C}^k(I, \mathbb{R})$... qui sont des $\mathbb{R}$-espaces vectoriels. On peut multiplier les fonctions point par point, ce qui fournit l'opération $*$, qui est bien bilinéaire et associative.

Si X est un ouvert de $\mathbb{R}^n$, alors l'espace des fonctions polynômiales, ou des polynômes trigonométriques, sont des sous-algèbres de $\mathcal{C}(X, \mathbb{R})$ (ou $\mathcal{C}(X, \mathbb{C})$ suivant que l'on autorise les valeurs complexes ou non).

Théorème 4 (Stone-Weierstrass)**.**
Soit K un espace métrique compact. Soit A une sous-algèbre de $\mathcal{C}(K, \mathbb{R})$. Supposons que
- ▷ *A contient la fonction $\mathbf{1}$ constante égale à 1 ;*
- ▷ *A sépare les points : pour tous $x \neq y \in K$, il existe $f \in A$ telle que $f(x) \neq f(y)$.*

Alors A est dense dans $\mathcal{C}(K, \mathbb{R})$: pour toute $f \in \mathcal{C}(K, \mathbb{R})$, il existe une suite $(g_n)_{n\in\mathbb{N}}$ de fonctions de A convergeant uniformément vers f.

La démonstration est assez longue, et sera omise.

1.4.2 Applications

Les deux exemples principaux sont :
- ▷ La sous-algèbre des fonctions polynômiales sur un segment $[a, b]$.
- ▷ La sous-algèbre des polynômes trigonométrique sur le cercle $[0, 1]_{0\sim1}$.

Pour le premier : les fonction polynômiales sont continues sur $[a, b]$, et l'ensemble des fonctions polynômiales contient $\mathbf{1}$, est stable par addition, multiplication par un réel et multiplication point par point. De plus, les fonctions polynômiales séparent les points (il suffit de considérer $f(x) = x$). Donc les fonctions polynômiales sont denses dans $\mathcal{C}([a, b], \mathbb{R})$.

Pour le second : les polynômes trigonométriques sont continus sur $[0, 1]_{0\sim1}$, et l'ensemble des polynômes trigonométriques contient $\mathbf{1} = e^{0x}$, est stable par addition, multiplication par un réel et multiplication point par point. De plus, polynômes trigonométriques séparent les points (il suffit de considérer $f(x) = \cos(2\pi x)$ et $g(x) = \sin(2\pi x)$, en se rappelant que l'on a identifié 0 et 1).

Donc les polynômes trigonométriques sont denses dans $\mathcal{C}([0,1]_{0\sim 1}, \mathbb{R})$. Autrement dit, les polynômes trigonométriques sont denses dans l'espace des fonctions continues 1-périodiques sur $\mathbb{R}$.

Le théorème de Stone-Weierstrass s'applique à des situations plus diverses. On peut ainsi montrer aussi facilement que l'ensemble des fonctions polynômiales à 2 variables est dense dans $\mathcal{C}([a,b] \times [c,d], \mathbb{R})$.

1.5 Application 3 : Coercivité et ellipsoïde de John

1.5.1 Coercivité

Une application importante de la compacité est la suivante.

Définition 1.12.
Soit $f : \mathbb{R}^n \to \mathbb{R}$ et $\ell \in \mathbb{R}$. On dit que $\lim_{x\to\infty} f(x) = \ell$ si, pour tout $\varepsilon > 0$, il existe K tel que $|f(x) - \ell| \leq \varepsilon$ pour tout x tel que $\|x\| \geq K$.

La définition des limites infinies est laissée au lecteur.

Proposition 1.13 ([Gou, Chapitre 1.3, Exercice 3])**.**
Soit $f : \mathbb{R}^n \to \mathbb{R}$ continue telle que $\lim_{x\to\infty} f(x) = +\infty$. Alors f est bornée inférieurement et atteint son minimum.

Démonstration.
Soit $x_0 \in \mathbb{R}^n$. Comme $\lim_{x\to\infty} f(x) = +\infty$, il existe $M \geq 0$ tel que $f(x) \geq f(x_0)$ dès que $\|x\| \geq M$. Mais alors $\inf_{\mathbb{R}^n} f = \inf_{\overline{B}(0,M)} f$. Or $\overline{B}(0,M)$ est compact par le théorème de Heine-Borel, donc f atteint son infimum sur cette boule, donc il existe $x \in \overline{B}(0,M)$ tel que $f(x) = \min_{\mathbb{R}^n} f$. □

Cette idée est flexible, en ce qu'elle s'adapte à de nombreuses situations : fonctions définies sur des ouverts de $\mathbb{R}^n$ (dans ce cas, f doit tendre vers $+\infty$ quand on se rapproche du bord de l'ouvert ou de l'infini), voire en dimension infinie (mais, dans ce cas, f doit être grande en-dehors de compacts ; on parle de fonctions coercives).

Exemple 1.14.
Soit ABC un triangle dans le plan euclidien. Alors il existe un point G qui minimise la somme des carrés des distances $f(M) := MA^2 + MB^2 + MC^2$. En effet, f est continue et est supérieure à $M \mapsto MA^2$, donc tend vers $+\infty$ en l'infini.

Le minimiseur de f est unique : c'est le centre de gravité du triangle. Cela se démontre par exemple en dérivant f et en montrant qu'elle a un unique point critique, qui est le centre de gravité du triangle. On peut aussi passer en coordonnées et factoriser astucieusement f.

Exemple 1.15.
Soit ABC un triangle dans le plan euclidien. Alors il existe un point F qui minimise la somme des distances $f(M) := MA + MB + MC$. En effet, f est continue et est supérieure à $M \mapsto MA$, donc tend vers $+\infty$ en l'infini.

Le minimiseur de f est unique : c'est le point de Fermat du triangle. L'analyse en est plus délicate, car f n'est pas dérivable partout, donc l'analyse de ses points critiques (s'ils existent !) n'est pas suffisante. Nous renvoyons la lectrice intéressée vers [Ska, Exercice 8.10].

Exemple 1.16.
Soit $\|\cdot\|$ une norme sur $\mathcal{C}([a,b], \mathbb{R})$. Soit $f : [a,b] \to \mathbb{R}$ une fonction continue. Soit $n \geq 0$. Alors la fonction $P \mapsto \|f - P\|$ est continue sur $\mathbb{R}_n[X]$, de par l'inégalité triangulaire et l'équivalence des

normes sur $\mathbb{R}_n[X]$ en dimension finie. De plus, elle tend vers $+\infty$ en l'infini. Elle atteint donc son minimum.

Autrement dit, à degré fixé, il existe un polynôme qui approche f "le mieux possible" pour la norme $\|\cdot\|$. Ce polynôme n'est pas nécessairement unique. Dans cet argument, $\mathbb{R}_n[X]$ peut être remplacé par n'importe quel sous-espace de $\mathcal{C}([a,b],\mathbb{R})$ de dimension finie.

1.5.2 Ellipsoïde de John

Théorème 5 ([Ska, Exercice 7.23])**.**
Soit K un compact d'intérieur non vide de $\mathbb{R}^n$. Il existe un unique ellipsoïde contenant K de volume minimal, appelé ellipsoïde de John de K.

Nous ne détaillerons pas ce développement, qui serait trop long à ce point du cours (il faudrait traiter les formes quadratiques, les ellipsoïdes...). Nous en donnons les grandes lignes, et renvoyons à la référence pour les détails.

- ▷ L'ensemble des ellipsoïdes de $\mathbb{R}^n$ peut être paramétré par un nombre fini N de réels. Par exemple, un ellipsoïde est uniquement déterminé par une équation de la forme $X^tAX+BX = 1$, où A est une matrice symétrique définie positive et B une forme linéaire, donc $N = n(n+3)/2$ convient.
- ▷ L'ensemble des ellipsoïdes contenant K est non vide, et fermé non seulement dans l'espace des ellipsoïdes, mais aussi dans $\mathbb{R}^N$ (on utilise ici la compacité de K). Notons F l'ensemble de ces paramètres.
- ▷ La fonction Vol définie sur F tend vers l'infini en l'infini, car K est d'intérieur non vide.

On peut donc adapter le résultat précédent à la fonction volume sur F, ce qui démontre l'existence d'un minimiseur.

L'unicité se montre à l'aide d'une inégalité de convexité (inégalité de type arithmético-géométrique).

1.6 Application 4 : Intersections de compacts et théorème de Dini

Nous montrons ici qu'une intersection décroissante de compacts non vides est non vide. Ce théorème a de nombreuses applications de technicités diverses, comme par exemple le théorème de Baire dans les espaces localement compacts, ou la définition de parties fractales [1] de $\mathbb{R}^n$. Nous en donnons une, le théorème de Dini.

1.6.1 Intersections de compacts

Le théorème suivant est le cœur de cette partie.

Théorème 6.
Soit X un espace métrique et $(K_n)_{n\geq 0}$ une suite de compacts non vides de X qui est décroissante pour l'inclusion (i.e. $K_{n+1} \subset K_n$ pour tout n).

Alors $\bigcap_{n\geq 0} K_n$ est non vide.

Démonstration.
Les K_n étant non vides, choisissons un point x_n dans chacun d'entre eux. Par décroissance, $x_n \in K_0$ pour tout n, et la suite $(x_n)_{n\in\mathbb{N}}$ a un point d'adhérence $x \in K_0$ par compacité de ce dernier.

1. Ensembles de Julia, systèmes de fonctions itérées...

Soit $N \in \mathbb{N}$. Alors, par décroissance encore, $x_n \in K_N$ pour tout $n \geq N$, donc l'ensemble des valeurs d'adhérence de la suite est dans K_N (ce dernier ensemble étant fermé), donc en particulier $x \in K_N$.

Ceci étant vrai pour tout $N \geq 0$, on obtient finalement $x \in \bigcap_{n \geq 0} K_n$, et ce dernier ensemble est par conséquent non vide. □

1.6.2 Théorème de Dini

Le théorème de Dini éclaire de lien entre convergence simple et convergence uniforme dans le cadre des suites monotones de fonctions.

Théorème 7 (Théorème de Dini [FGN·Ana3, Exercice 2.30])**.**
Soit K un compact et $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions dans $\mathcal{C}(K,\mathbb{R})$. Supposons que :
- ▷ *$(f_n)_{n\in\mathbb{N}}$ est monotone, i.e. $(f_n(x))_{n\in\mathbb{N}}$ est croissante pour tout $x \in K$, ou décroissante pour tout $x \in K$.*
- ▷ *$(f_n)_{n\in\mathbb{N}}$ converge simplement vers une fonction $f \in \mathcal{C}(K,\mathbb{R})$.*

Alors $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f.

Démonstration.
Sans perte de généralité, supposons que $(f_n)_{n\in\mathbb{N}}$ est croissante ; en particulier, $f_n \leq f$ pour tout n. Soit $\varepsilon > 0$. Notre but est de montrer que $f_n > f - \varepsilon$ pour tout n suffisamment grand.

Posons $K_n := \{x \in K : \ f_n(x) \leq f(x) - \varepsilon\}$ pour tout n. Ces ensembles sont compacts car fermés dans le compact K. Ils sont aussi décroissants pour l'inclusion : si $x \in K_{n+1}$, alors $f_n(x) \leq f_{n+1}(x) \leq f(x) - \varepsilon$, donc $x \in K_n$.

Ainsi, si tous les K_n sont non vides, alors leur intersection est non vide. Or un point de leur intersection est un point x tel que $f_n(x) \leq f(x) - \varepsilon$ pour tout n, et donc tel que $(f_n(x))_{n\in\mathbb{N}}$ ne converge pas vers $f(x)$. Cela contredit la convergence simple de $(f_n)_{n\in\mathbb{N}}$ vers f. Donc il existe N tel que $K_N = \emptyset$, et par décroissance $K_n = \emptyset$ pour tout $n \geq N$. Par définition des K_n, cela revient à dire que $f_n > f - \varepsilon$ pour tout $n \geq N$, ce que l'on voulait démontrer. □

1.7 Application 5 : Compacité et distances

1.7.1 Distance à un compact

Définition 1.17 (Distance entre parties)**.**
La distance entre deux parties A et B d'un espace métrique X est définie par

$$d(A,B) = \inf\{d(a,b) : \ a \in A, \ b \in B\}.$$

Contrairement à ce que la notation suggère, il ne s'agit pas d'une distance sur l'espace des parties de X : elle n'est pas définie positive (deux parties ayant un point en commun sont à distance nulle), et ne satisfait pas l'inégalité triangulaire. Elle est cependant positive et symétrique.

Le lien avec la compacité est le suivant :

Proposition 1.18 ([Gou, Chapitre 1.3, Exercice 3])**.**
Soient F un fermé et K un compact de $\mathbb{R}^n$, tous deux non vides. Alors l'infimum est un minimum : il existe $x \in K$ et $y \in F$ tels que $d(K,F) = d(x,y)$.

Démonstration.
Si F est borné, alors F est compact, et $K \times F$ aussi. La fonction distance restreinte à $K \times F$ est continue, donc elle atteint son minimum en une paire (x,y) qui convient.

Si F n'est pas borné, alors la fonction distance restreinte à $K \times F$ est continue et tend vers $+\infty$ en $+\infty$. Plus précisément, soit m une borne de K. Étant donné $M \geq 0$, si $y \in F$ est tel que $\|y\| \geq M + m$, alors $d(x, y) \geq M$ pour tout $x \in K$. En particulier, fixons $x_0 \in K$ et $y_0 \in F$. Si $\|y\| \geq d(x_0, y_0) + m + 1$, alors $d(x, y) \geq d(x_0, y_0) + 1$ pour tout $x \in K$. Par conséquent,

$$d(K, F) = \inf\{d(x, y) : \ x \in K, \ y \in F \cap \overline{B}(0, d(x_0, y_0) + m + 1)\}.$$

On s'est ramené au cas de deux ensembles compacts, pour lequel le minimum est atteint. □

Exercice 1.19.
Montrez à l'aide d'un exemple que l'infimum n'est pas toujours un minimum quand on considère la distance entre deux fermés.

1.7.2 Distance entre compacts

On peut aussi définir une vraie distance entre compacts d'un espace métrique.

Définition 1.20 (Épaississement)**.**
Soient A une partie d'un espace métrique X et $\varepsilon > 0$. Le ε-voisinage de A est

$$B(A, \varepsilon) := \{x \in X : \ d(\{x\}, A) \leq \varepsilon\}.$$

Remarque 1.21.
Un singleton $\{x\}$ étant compact, si A est un fermé de $\mathbb{R}^n$, alors l'infimum dans la définition de $d(\{x\}, A)$ est atteint. Dans ce cas,

$$B(A, \varepsilon) := \{x \in X : \ \exists y \in A, d(x, y) \leq \varepsilon\}.$$

Définition 1.22 (Distance de Hausdorff)**.**
Pour tous compacts K_1, K_2 non vides d'un espace métrique X, on note

$$d_H(K_1, K_2) := \inf\{\varepsilon > 0 : \ K_1 \subset B(K_2, \varepsilon) \text{ et } K_2 \subset B(K_1, \varepsilon)\}$$

*la **distance de Hausdorff** entre K_1 et K_2.*

Proposition 1.23 ([RW·L2, Exercice 11 p. 512])**.**
La distance de Hausdorff sur X est une distance sur l'espace des parties compactes non vides de X.

Démonstration.
Soient $x_0 \in K_1$ et $y_0 \in K_2$. Soit $m_1 := \max\{d(x, x') : \ x, x' \in K_1\}$, et de même pour K_2. Alors tout point de K_1 est à distance au plus $d(x_0, y_0) + m_1$ de y_0, donc $K_1 \subset B(K_2, d(x_0, y_0) + m_1)$. De même, $K_2 \subset B(K_1, d(x_0, y_0) + m_2)$. Par conséquent, $d_H(K_1, K_2)$ est fini (et même inférieur à $d(x_0, y_0) + \max\{m_1, m_2\}$).

Il est alors évident que d_H est positive et symétrique.

Montrons l'inégalité triangulaire. Soient K_1, K_2, K_3 trois compacts, et $\varepsilon > 0$. Soient $x_1 \in K_1$. Il existe $x_2 \in K_2$ tel que $d(x_1, x_2) \leq d_H(K_1, K_2) + \varepsilon$. Il existe $x_3 \in K_3$ tel que $d(x_2, x_3) \leq d_H(K_2, K_3) + \varepsilon$. Par l'inégalité triangulaire usuelle, $d(x_1, x_3) \leq d_H(K_1, K_2) + d_H(K_2, K_3) + 2\varepsilon$, donc $K_1 \subset B(K_3, d_H(K_1, K_2) + d_H(K_2, K_3) + 2\varepsilon)$.

En échangeant K_1 et K_3, on trouve de même $K_3 \subset B(K_1, d_H(K_1, K_2) + d_H(K_2, K_3) + 2\varepsilon)$. Par conséquent, $d_H(K_1, K_3) \leq d_H(K_1, K_2) + d_H(K_2, K_3) + 2\varepsilon$. Le paramètre ε étant arbitraire, on en déduit l'inégalité triangulaire pour d_H.

Il reste à montrer que d_H est définie positive. Soient K_1, K_2 deux compacts à distance de Hausdorff nulle. Soit $x \in K_1$. Alors $x \in B(K_1, 1/n)$ pour tout $n \geq 1$, donc il existe une suite $(y_n)_{n \geq 1}$ à valeurs dans K_2 telle que $d(x, y_n) \leq 2/n$ pour tout $n \geq 1$. En particulier, $x \in \overline{K_2} = K_2$. Ceci étant vrai pour tout x, on obtient $K_1 \subset K_2$. En échangeant le rôle de K_1 et K_2, on montre finalement que $K_1 = K_2$. $\square$

Remarque 1.24.
La fonctionnelle d_H est définie entre toutes les parties bornées non vides d'un espace métrique ; la compacité n'est utilisée ici que pour montrer que d_H est définie positive.

1.8 Application 6 : Transformations d'un compact

Il y aurait beaucoup à dire sur les transformations d'un compact, mais un grand nombre de résultats demande des notions sophistiquées (compacts invariants, minimalité) ou des outils eux aussi sophistiqués (théorie de Baire, théorie de la mesure). Nous avons sélectionné deux résultats plus élémentaires.

Dilatations

Proposition 1.25 ([FGN·Ana3, Exercice 2.3][Gou, Chapitre 1.3, Exercice 5]).
Soit K un compact et $f : K \to K$ une dilatation, i.e ; une application telle que

$$d(f(x), f(y)) \geq d(x, y) \quad \forall x, y \in K.$$

Alors f est une isométrie bijective.

Démonstration.
Montrons dans un premier temps que tout point a est valeur d'adhérence de la suite $(f^n(a))_{n \in \mathbb{N}}$. En effet, soit b une valeur d'adhérence de $(f^n(a))_{n \in \mathbb{N}}$. Il existe donc une sous-suite extraite $(f^{n_k}(a))_{k \geq 0}$ convergeant vers b ; quitte à extraire encore une sous-suite, on peut supposer que $(n_{k+1} - n_k)_{k \in \mathbb{N}}$ est strictement croissante.

Mais alors $d(f^{n_{k+1}}(a), f^{n_k}(a)) \geq d(f^{n_{k+1} - n_k}(a), a)$ par dilatation, et le membre de gauche tend vers $d(b, b) = 0$. Donc le membre de droite tend aussi vers 0 : on a trouvé une sous-suite extraite le long de laquelle $(f^n(a))_{n \in \mathbb{N}}$ converge vers a.

On ne va pas appliquer ce qui précède à f, mais à un autre système. Posons $K' = K \times K$, qui est compact, et $g(x, y) = (f(x), f(y))$. Alors g est une dilatation pour la distance $d_\infty((x, y), (x', y')) := \max\{d(x, y), d(x', y')\}$. Par conséquent, pour tous $a, b \in K$, on peut trouver une sous-suite extraite le long de laquelle $(g^n(a, b))_{n \in \mathbb{N}} = (f^n(a), f^n(b))_{n \in \mathbb{N}}$ converge vers (a, b). La suite $(d(f^n(a), f^n(b)))_{n \in \mathbb{N}}$ est donc croissante, vaut $d(a, b)$ en 0 et converge vers $d(a, b)$ le long d'une sous-suite ; elle est donc constante égale à $d(a, b)$. Autrement dit, $d(f(a), f(b)) = d(a, b)$ pour tous a, b, donc f est une isométrie.

Il reste à montrer que f est bijective. Mais f est injective (comme toute isométrie). De plus, pour tout $a \in K$, le point a est adhérent à $(f^n(a))_{n \in \mathbb{N}}$, donc $a \in \overline{f(K)}$. Or $f(K)$ est compact, donc fermé, donc $a \in f(K)$. Par conséquent, f est aussi surjective. $\square$

Contractions faibles

Proposition 1.26 ([FGN·Ana3, Exercice 2.5][Gou, Chapitre 1.3, Exercice 4][Ska, Exercice 3.8]).
Soit K un compact et $f : K \to K$ une contraction faible, i.e ; une application telle que

$$d(f(x), f(y)) < d(x, y) \quad \forall x \neq y \in K.$$

Alors f a un unique point fixe α, et pour tout $x \in K$, la suite $(f^n(x))_{n \geq 0}$ converge vers α.

Démonstration.
Remarquons que f est 1-lipschitzienne, donc continue. Montrons d'abord l'existence et l'unicité d'un point fixe. Soit $g(x) := d(x, f(x))$ pour $x \in K$. Alors g est continue, donc atteint son minimum en un point α. Si $f(\alpha) \neq \alpha$, alors $g(f(\alpha)) = d(f(\alpha), f^2(\alpha)) < d(\alpha, f(\alpha)) = g(\alpha)$, ce qui contredit la minimalité de $g(\alpha)$. Donc $f(\alpha) = \alpha$.

Soit β un point fixe. Si $\beta \neq \alpha$, alors $d(\alpha, \beta) = d(f(\alpha), f(\beta)) < d(\alpha, \beta)$, ce qui est absurde. Donc α est l'unique point fixe de f.

Soit $x \in K$. Soit β un point d'adhérence de la suite $(f^n(x))_{n\in\mathbb{N}}$, et supposons $\beta \neq \alpha$. Remarquons que la suite $(d(\alpha, f^n(x)))_{n\in\mathbb{N}}$ est positive décroissante, donc convergente ; et elle converge vers $d(\alpha, \beta)$ le long d'une sous-suite, donc elle converge vers $d(\alpha, \beta)$. Mais $d(\alpha, f(\beta)) < d(\alpha, \beta)$, donc $d(\alpha, f(x)) < d(\alpha, \beta)$ pour tout x dans un voisinage de β, et en particulier il existe n grand tel que $d(\alpha, f^n(x)) < d(\alpha, \beta)$. Cela contredit le fait que la suite $(d(\alpha, f^n(x)))_{n\in\mathbb{N}}$ décroît vers $d(\alpha, \beta)$. L'hypothèse de départ est donc absurde, donc α est le seul point d'adhérence de $(f^n(x))_{n\in\mathbb{N}}$, donc cette suite converge vers α. □

Remarque 1.27.
D'après le théorème du point fixe de Banach, dans un espace complet, les conclusions de cette proposition sont valides s'il existe $\lambda < 1$ tel que $d(f(x), f(y)) \leq \lambda d(x, y)$ pour tous x, y. Autrement dit, le taux de contraction doit être uniforme ; on parle parfois de ***contractions fortes****.*

Comme cette proposition le montre, on peut affaiblir l'hypothèse de contraction forte au prix de la compacité.

Exercice 1.28.
Construisez un contre-exemple aux conclusions de la proposition quand K est fermé dans $\mathbb{R}^n$, mais pas compact.

2 Continuité uniforme

Dans cette partie, (X, d_X) et (Y, d_Y) sont deux espaces métriques.

2.1 Définition

La continuité uniforme, définie dans des espaces métriques [2] est une propriété plus restrictive que la continuité.

Définition 2.1 (Continuité uniforme)**.**
Une fonction $f : X \to Y$ est dite ***uniformément continue*** *si*

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall x_1, x_2 \in X,\ d_X(x_1, x_2) \leq \delta \Rightarrow d_Y(f(x_1), f(x_2)) \leq \varepsilon.$$

Cette définition ressemble à celle de la continuité ; elle n'en diffère que par une interversion de quantificateurs. Dans la définition métrique de la continuité, le paramètre δ peut dépendre de x_1 ; ce n'est pas le cas pour la continuité uniforme, pour laquelle le paramètre δ est uniforme en x_1.

Exercice 2.2.
Soit $f : X \to Y$ une fonction uniformément continue et $(u_n)_{n\in\mathbb{N}}$ une suite de Cauchy dans X. Montrez que $(f(u_n))_{n\in\mathbb{N}}$ est une suite de Cauchy dans Y.

2. Ou, plus généralement, dans des *espaces uniformes*.

2.2 Module de continuité

Formellement, dans la définition de la continuité uniforme, le paramètre δ ne dépend que de ε : on peut écrire $\delta = \delta(\varepsilon)$. La proposition suivante formalise cette observation. Nous renvoyons aussi la lectrice à [Ska, Exercice 7.4].

Proposition 2.3.
Soit $f : X \to Y$. Alors f est uniformément continue. si et seulement s'il existe $\delta_0 > 0$ et une fonction $\omega_f : [0, \delta_0) \to [0, +\infty)$ continue en 0, croissante, telle que $\omega_f(0) = 0$ et, pour tous x_1, $x_2 \in X$ à distance au plus δ_0,

$$d_Y(f(x_1), f(x_2)) \leq \omega_f(d_X(x_1, x_2)).$$

Si de plus $X = \mathbb{R}^n$, alors ω_f est définie et continue sur $\mathbb{R}_+$ (ou, autrement dit, $\delta_0 = +\infty$ convient).

*On appelle une telle fonction ω_f un **module de continuité** de f.*

En particulier, pour des fonctions réelles d'une variable réelle uniformément continues, pour tous x, y à distance au plus δ_0,

$$|f(x) - f(y)| \leq \omega_f(|x - y|).$$

Démonstration.
La réciproque étant comparativement facile (il suffit d'utiliser la continuité de ω_f en 0), concentrons-nous sur le sens direct. Il suffit de poser

$$\omega_f(t) := \sup_{\substack{x_1, x_2 \in X \\ d(x_1, x_2) \leq t}} d_Y(f(x_1), f(x_2)).$$

Alors ω_f est croissante, et $\omega_f(0) = 0$. De plus, soit δ_0 tel que $\forall x_1, x_2 \in X,\ d_X(x_1, x_2) \leq \delta_0 \Rightarrow d_Y(f(x_1, f(x_2)) \leq 1$. Alors $\omega_f(\delta_0) \leq 1$, et en particulier ω_f est finie sur $[0, \delta_0)$. Par définition de la continuité uniforme, $\lim_0 \omega_f = 0$, ce qui donne la continuité en 0.

Il reste à montrer les propriétés additionnelles pour les fonctions définies sur $\mathbb{R}$. Soient $0 \leq s \leq t$ et x_1, x_2 à distance au plus t. Soit $x_3 \in [x_1, x_2]$ à distance au plus s de x_1 et au plus $t - s$ de x_2. Alors

$$d(f(x_1), f(x_2)) \leq d(f(x_1), f(x_3)) + d(f(x_2), f(x_3)) \leq \omega_f(s) + \omega_f(t - s).$$

En prenant le supremum sur x_1 et x_2, on obtient finalement $\omega_f(t) \leq \omega_f(s) + \omega_f(t-s)$. En particulier, en itérant cette inégalité,

$$\omega_f(t) \leq \lceil t/\delta_0 \rceil \omega_f(\delta_0) < +\infty,$$

donc ω_f est bien finie sur $\mathbb{R}_+$. Enfin, $\omega_f(t) \leq \omega_f(t + h) \leq \omega_f(t) + \omega_f(h) \to_{h \to 0} \omega_f(t)$, donc ω_f est bien continue à droite partout. De même, $\omega_f(t) - \omega_f(h) \leq \omega_f(t - h) \leq \omega_f(t)$, donc ω_f est bien continue à gauche partout. $\square$

Exemple 2.4.
Une fonction f est lipschitzienne si et seulement s'il existe $L \geq 0$ tel que la fonction $t \mapsto Lt$ soit un module de continuité de f. En particulier, toute fonction de classe $\mathcal{C}^1$ sur un segment est uniformément continue.

Une fonction f est θ-höldérienne si et seulement s'il existe $L \geq 0$ tel que la fonction $t \mapsto Lt^\theta$ soit un module de continuité de f.

Exemple 2.5.
La fonction $f : x \to x^2$ définie sur $\mathbb{R}$ est continue mais pas uniformément continue. En effet, procédons par l'absurde choisissons $\varepsilon = 1$ dans la définition de la continuité uniforme. Alors, pour tout $\delta > 0$, on a $(x+\delta)^2 - x^2 > 2\delta x$. En particulier, les points $\frac{1}{2\delta}$ et $\frac{1}{2\delta} + \delta$ sont à distance δ, mais les valeurs correspondantes de f sont à distance au moins 1.

Plus généralement, si f est une fonction uniformément continue sur $\mathbb{R}$, alors il existe a, $b \geq 0$ tels que $|f(x)| \leq a|x| + b$ pour tout $x \in \mathbb{R}$ [Rom, Exercice 6.12] [Ska, Exercice 7.3].

2.3 Théorème de Heine

Comme mentionné, les fonctions de classe $\mathcal{C}^1$ sur un segment sont uniformément continues. Le théorème de Heine affirme que, sur un segment, la continuité suffit.

Théorème 8 (Théorème de Heine)**.**
Soit X un espace (séquentiellement) compact, Y un espace métrique et $f : X \to Y$. Alors f est continue si et seulement si elle est uniformément continue.

En particulier, pour tous $a < b$, toute fonction $f : [a,b] \to \mathbb{R}$ est uniformément continue.

Démonstration.
On utilise le module de continuité. Soit

$$\omega_f(t) := \sup_{\substack{x_1, x_2 \in X \\ d(x_1,x_2) \leq t}} d_Y(f(x_1), f(x_2)).$$

On sait que ω_f est croissante et $\omega_f(0) = 0$. Il suffit de montrer que $\omega_f < +\infty$ sur un voisinage de 0 et que $\lim_0 \omega_f = 0$.

Comme X est compact, f est bornée, donc $\omega_f(t) < +\infty$ pour tout t.

Soit $\ell := \lim_0 \omega_f$. On procède par l'absurde. Supposons que $\ell > 0$. Alors, pour tout $n \geq 1$, il existe x_n, $y_n \in X$ tels que $d_X(x_n, y_n) \leq 1/n$ et $d_Y(f(x_n), f(y_n)) \geq \ell/2$.

Soit $(x,y) \in X^2$ un point d'adhérence de (x_n, y_n). Alors $d_X(x,y) = \lim_{n\to+\infty} d(x_n, y_n) = 0$, donc $x = y$, et $d_y(f(x), f(y)) \geq \ell/2 > 0$, ce qui est absurde. □

Exercice 2.6.
Réécrire la preuve ci-dessus en utilisant la définition en ε-δ de la continuité uniforme.

Le théorème de Heine sur $\mathbb{R}$ admet une généralisation importante :

Proposition 2.7 ([Ska, Exercice 7.2])**.**
Soit $f : \mathbb{R} \to \mathbb{R}$ une application continue admettant des limites finies en $\pm\infty$. Alors f est uniformément continue.

Démonstration.
Soit $\varepsilon > 0$. Comme f admet des limites finies en $\pm\infty$, la famille $(f(x))_{x\in\mathbb{R}}$ satisfait le critère de Cauchy au voisinage de $\pm\infty$. En particulier, il existe $M \geq 0$ tel que, pour tous x, $y \geq M$, on ait $|f(x) - f(y)| \leq \varepsilon$, et de même pour x, $y \leq -M$.

Posons $K := [-M-2, M+2]$. Alors, $f_{|K}$ étant uniformément continue par le théorème de Heine, il existe $\delta \leq 1$ tel que, pour tous x, $y \in K$ tels que $|x - y| \leq \delta$, on ait $|f(x) - f(y)| \leq \varepsilon$.

Soient x, $y \in \mathbb{R}$ tels que $|x - y| \leq \delta$.

- ▷ Si $x \in [-M-1, M+1]$, alors $y \in K$ (car $\delta \leq 1$), donc $|f(x) - f(y)| \leq \varepsilon$ (car $[x - y| \leq \delta$).

▷ Si $x \geq M+1$, alors $y \geq M$ (car $\delta \leq 1$), donc donc $|f(x)-f(y)| \leq \varepsilon$.
▷ De même $x \leq -M-1$.

Dans tous les cas, $|f(x)-f(y)| \leq \varepsilon$, ce qu'il fallait démontrer. □

Remarque 2.8.
En particulier, on peut trouver facilement des fonctions uniformément continues sur $\mathbb{R}$ dont la dérivée est non bornée ; par exemple, $f(x) = \frac{\cos(x^4)}{1+x^2}$.

2.4 Application 1 : Prolongements de fonctions continues

Soit $A \subset X$ et $f : X \to Y$. À quelle condition f peut-elle être étendue en une fonction continue de X dans Y ? Nécessairement, f doit être continue, mais cette condition est-elle suffisante ? Cela dépend de l'ensemble A en question.

Proposition 2.9 ([Ska, Exercice 7.7]).
Soit $A \subset \mathbb{R}$ et $f : A \to \mathbb{R}$ continue. Si A est fermé, alors il existe $g : \mathbb{R} \to \mathbb{R}$ continue telle que $g_{|A} = f$.

Démonstration.
Comme A est fermé dans $\mathbb{R}$, le complémentaire de A est une union finie d'intervalles ouverts. On étend f :
▷ par une constante sur les éventuels intervalles non bornés (il y en a au plus 2) ;
▷ de façon affine sur chaque intervalle borné.

On vérifie alors que g est bien continue en tout point x :
▷ si $(x, x+1) \cap A$ est non vide, alors il existe un voisinage de x sur lequel on peut contrôler g directement grâce à f ;
▷ sinon, f est affine sur $[x, x+1)$, donc continue ;
▷ de même sur $(x-1, x)$.

□

Plus utile est l'extension de fonctions définies sur un ensemble dense. Cette extension utilise de façon cruciale la notion de suite de Cauchy.

Proposition 2.10.
Soit $A \subset \mathbb{R}$ et $f : A \to \mathbb{R}$ uniformément continue. Si A est dense dans $\mathbb{R}$, alors il existe une unique fonction $g : \mathbb{R} \to \mathbb{R}$ continue telle que $g_{|A} = f$.

Démonstration.
Soit $x \in \mathbb{R}$. Soit $(u_n^x)_{n\in\mathbb{N}}$ une suite à valeurs dans A convergeant vers x. Alors $(u_n^x)_{n\in\mathbb{N}}$ est de Cauchy et f est uniformément continue, donc $(f(u_n^x))_{n\in\mathbb{N}}$ est de Cauchy. Comme $\mathbb{R}$ est complet, $(f(u_n^x))_{n\in\mathbb{N}}$ converge. Soit $g(x)$ sa limite.

Montrons que g est uniformément continue. Soit ω_f définie sur $[0, \delta_0)$ un module de continuité de f. Soient $x \neq y \in \mathbb{R}$ tels que $|x-y| < \delta_0/3$. Soient $(u_n^x)_{n\in\mathbb{N}}$, $(u_n^y)_{n\in\mathbb{N}}$ les suites utilisées pour définir $g(x)$ et $g(y)$. Il existe $N \geq 0$ tel que $|u_n^x - x|$, $|u_n^y - y| < \delta_0/3$ pour tout $n \geq N$. Mais alors, pour tout $n \geq N$,

$$\begin{aligned} |g(x)-g(y)| &\leq |g(x)-f(u_n^x)| + |f(u_n^x)-f(u_n^y)| + |f(u_n^y)-g(y)| \\ &\leq |g(x)-f(u_n^x)| + \omega_f(|u_n^x-u_n^y|) + |f(u_n^y)-g(y)|. \end{aligned}$$

De plus, $\omega_f(|u_n^x - u_n^x|) \leq \omega_f(2|x-y|)$ pour tout n suffisamment grand, et $|g(x)-f(u_n^x)|$ et $|g(y)-f(u_n^y)|$ convergent vers 0. Par conséquent,

$$|g(x)-g(y)| \leq \omega_f(2|x-y|).$$

Donc $t \mapsto \omega_f(2t)$ est un module de continuité pour g, donc g est uniformément continue.

L'unicité provient du fait que deux fonctions continues coïncidant sur un ensemble dense sont égales. □

Remarque 2.11.
La démonstration ci-dessus, contrairement à celle du prolongement de fonctions définies sur des fermées, utilise peut les propriétés des réels. Elle est en fait vraie pour deux espaces métriques X et Y tels que Y soit complet.

Cette version plus générale est particulièrement utile en analyse fonctionnelle : si X est un espace vectoriel normé et Y est un espace de Banach, on peut étendre de façon unique toute application linéaire bornée définie sur un ensemble dense de X à valeurs dans Y.

Remarque 2.12.
L'hypothèse d'uniforme continuité est cruciale. Par exemple, la fonction signe définie sur $\mathbb{R}^*$ est continue, mais ne s'étend pas en une fonction continue sur $\mathbb{R}$.

2.5 Application 2 : Théorème d'Arzelà-Ascoli

Les compacts de $\mathbb{R}^n$ sont les fermés bornés. Comme évoqué précedemment, ce résultat est faux en dimension infinie : les boules fermées sont trop grosses, par exemple. Le théorème d'Arzelà-Ascoli fourni une réponse dans les espaces de fonctions continues.

Ce résultat est suffisamment technique pour qu'il soit dangereux de l'utiliser en développement. Il permet cependant d'éclairer la notion de compacité en dimension infinie.

Définition 2.13.
Une famille de fonctions $(f_i)_{i\in I}$ entre deux espaces métriques est dite ***uniformément équicontinue***[3] *si, pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que, pour tous x, y, si $d(x,y) \leq \delta$, alors $d(f_i(x), f_i(y)) \leq \varepsilon$ pour tout $i \in I$.*

De façon équivalente, une famille uniformément équicontinue est une famille de fonctions uniformément continues ayant un même module de continuité. Par exemple, l'ensemble des fonctions 1-lipschitziennes sur $\mathbb{R}$ est uniformément équicontinu.

Théorème 9 (Arzelà-Ascoli [FGN·Ana3, Exercice 2.34])**.**
Une partie $K \subset \mathcal{C}([a,b],\mathbb{R})$ est compacte si et seulement si K est fermée, bornée, et uniformément équicontinue.

Démonstration.
Si K est compacte, alors K est fermée et bornée. De plus, l'application de $K \times [a,b] \to \mathbb{R}$ qui à (f,x) associe $f(x)$ est continue, donc uniformément continue par le théorème de Heine. Elle admet donc un module de continuité ω, tel que $|f(x) - g(y)| \leq \omega(\|f-g\|_\infty + |x-y|)$ pour tous f, $g \in K$ et x, $y \in [a,b]$ suffisamment proches. En particulier, en prenant $f = g$, pour tout $f \in K$ et tous x, $y \in [a,b]$ suffisamment proches, on a $|f(x) - f(y)| \leq \omega(|x-y|)$.

Soit $K \subset \mathcal{C}([a,b],\mathbb{R})$ fermée, bornée et uniformément continue ; notons ω un module de continuité. Soit $(f_n)_{n\in\mathbb{N}}$ une suite à valeurs dans K. On se donne aussi une suite $(x_n)_{n\in\mathbb{N}}$ dense dans $[a,b]$.

On va construire par récurrence, via un argument diagonal, une sous-suite convergente de $(f_n)_{n\in\mathbb{N}}$.

Pour tout $N \geq 1$, posons $\delta_N := \max\{|x - x_i| : \ x \in [a,b], \ 0 \leq i < N\}$. Alors deux fonctions f, $g \in K$ coïncidant sur $(x_i)_{0\leq i<n}$ sont à distance au plus $2\omega(\delta_N)$ l'une de l'autre. De plus, par densité, $\lim_{N\to+\infty} \delta_N = 0$.

On pose :

3. Un terme moins élégant mais plus adéquat pourrait être celui de famille uniformément uniformément continue...

- ▷ $f_n^{(0)} = f_n$ pour tout $n \geq 0$;
- ▷ Soit $N \geq 0$. L'ensemble $\{f_n^{(N)}(x_N) : \ n \in \mathbb{N}\}$ est borné car K est borné. On peut donc extraire une sous-suite $(f_n^{(N+1)})_{n\in\mathbb{N}}$ de $(f_n^{(N)})_{n\in\mathbb{N}}$ telle que $f_n^{(N+1)}(x_N)$ converge.
- ▷ En particulier, il existe n_0 tel que, pour tous n, $m \geq n_0$ suffisamment grand, on a $|f_n^{(N+1)}(x_k) - f_m^{(N+1)}(x_k)| \leq 3\omega(\delta_N)$. On ne garde que les termes de la suite $n \geq n_0$, avec $n \geq 1$.

Soit alors $g_N = f_0^{(N)} =: f_{n_N}$. La suite n_N est strictement croissante (car, à chaque étape, on n'a gardé que des indices $n \geq 1$). De plus, par construction, pour tout $n, m \geq N \geq 0$, on a $\|g_n - g_m\|_\infty \leq 3\omega(\delta_N)$. Par conséquent, la suite $(g_n)_{n\in\mathbb{N}}$ est de Cauchy, donc converge vers $f \in \mathcal{C}([a,b],\mathbb{R})$. Or K est fermé, donc $f \in K$. □

Exemple 2.14.
L'ensemble des fonctions $f : [0,1] \to \mathbb{R}$ qui sont 1-lipschitziennes et telles que $f(0) = 1$ est compact dans $\mathcal{C}([0,1],\mathbb{R})$.

2.6 Application 3 : Sommes de Riemann

Une autre application importante de la continuité uniforme est la convergence des sommes de Riemann. Nous admettons ici que l'intégrale d'une fonction continue sur un segment est bien définie et a les propriétés attendues (linéarité, positivité). Nous utiliserons le lemme suivant, dont la démonstration est laissée en exercice :

Lemme 2.15.
Soit I un segment, f une fonction réelle continue sur I et $c \leq d$ deux réels tels que

$$c \leq f \leq d.$$

Alors

$$c \leq \frac{1}{|I|} \int_I f(t)\ \mathrm{d}t \leq d.$$

Remarque 2.16.
Ce lemme a une généralisation importante, liée à l'inégalité de Jensen : Si C est un convexe borné d'un espace de Banach, μ une mesure de probabilité sur un espace X et $f : X \to C$, alors $\int f \,\mathrm{d}\mu \in C$. Le lemme précédent et le cas particulier de $C = [c,d]$ et μ la mesure uniforme sur I.

D'un point de vue plus élémentaire, si on se donne des points dans un convexe munis de poids positifs, leur barycentre est encore dans le convexe.

Proposition 2.17.
Soit $[a,b]$ un segment et f une fonction réelle continue sur $[a,b]$. Soit ω_f un module de continuité de f. Alors, pour tout n suffisamment grand,

$$\left| \frac{b-a}{n} \sum_{k=0}^{n-1} f\left(a + k\frac{b-a}{n}\right) - \int_a^b f(t)\ \mathrm{d}t \right| \leq (b-a)\omega_f\left(\frac{b-a}{n}\right)$$

et

$$\left| \frac{b-a}{n} \sum_{k=0}^{n-1} f\left(a + \left(k+\frac{1}{2}\right)\frac{b-a}{n}\right) - \int_a^b f(t)\ \mathrm{d}t \right| \leq (b-a)\omega_f\left(\frac{b-a}{2n}\right).$$

En particulier, ces sommes de Riemann convergent vers l'intégrale de f quand n tend vers $+\infty$.

Démonstration.
Soit δ_0 tel que ω_f soit défini sur $[0, \delta_0)$. Alors $\frac{b-a}{n} < \delta_0$ pour tout n suffisamment grand, et dans ce cas, pour tout $0 \leq k < n$ et $t \in [a + k\frac{b-a}{n}, a + (k+1)\frac{b-a}{n}]$,

$$f\left(a + k\frac{b-a}{n}\right) - \omega_f\left(\frac{b-a}{n}\right) \leq f(t) \leq f\left(a + k\frac{b-a}{n}\right) + \omega_f\left(\frac{b-a}{n}\right).$$

D'après le Lemme 2.15,

$$f\left(a + k\frac{b-a}{n}\right) - \omega_f\left(\frac{b-a}{n}\right) \leq \int_{a+k\frac{b-a}{n}}^{a+(k+1)\frac{b-a}{n}} f(t)\,\mathrm{d}t \leq f\left(a + k\frac{b-a}{n}\right) + \omega_f\left(\frac{b-a}{n}\right).$$

En sommant ces inégalités puis en multipliant par $\frac{b-a}{n}$, on obtient finalement la première inégalité. La seconde s'obtient de façon similaire. □

En particulier, si f est θ-höldérienne, alors il existe $C \geq 0$ tel que

$$\left|\frac{b-a}{n}\sum_{k=0}^{n-1} f\left(a + k\frac{b-a}{n}\right) - \int_a^b f(t)\,\mathrm{d}t\right| \leq Cn^{-\theta}.$$

Dans le cas de fonctions de classe $\mathcal{C}^1$ ou $\mathcal{C}^2$, ces bornes d'erreurs peuvent être significativement améliorées, et mener à un développement plus riche pour l'oral d'agrégation. Cet aspect sera vu plus en détail dans le chapitre sur l'intégration.

2.7 Application 4 : Produit de convolution

Une autre application concerne la convergence des produits de convolution, là encore car la continuité uniforme permet de contrôler les valeurs d'une fonction évaluée en des points proches.

Proposition 2.18.
Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction uniformément continue, et ω_f un module de continuité de f.
Soit g une fonction continue par morceaux sur $\mathbb{R}$, nulle en dehors d'un segment $[a, b]$, et d'intégrale 1*. Posons, pour tous x réel et $\varepsilon > 0$,*

$$f * g_\varepsilon(x) := \varepsilon^{-1}\int_{\mathbb{R}} f(t)g(\varepsilon^{-1}(x-t))\,\mathrm{d}t.$$

Alors $\|f * g_\varepsilon - f\|_\infty \leq \omega_f(\varepsilon \max\{|a|, |b|\})$ pour tout ε suffisamment petit. En particulier, $f * g_\varepsilon$ converge vers f uniformément quand ε tend vers 0*.*

Démonstration.
Soient x un réel et $\varepsilon > 0$. La fonction $t \mapsto g(\varepsilon^{-1}(x-t))$ est nulle si $\varepsilon^{-1}(x-t) \notin [a, b]$, donc si $t \notin [x - \varepsilon b, x - \varepsilon a]$. Par conséquent,

$$f * g_\varepsilon(x) := \varepsilon^{-1}\int_{x-\varepsilon b}^{x-\varepsilon a} f(t)g(\varepsilon^{-1}(x-t))\,\mathrm{d}t.$$

Sur cet intervalle, $|f - f(x)| \leq \omega(\varepsilon \max\{|a|, |b|\})$, et $t \mapsto \varepsilon^{-1}g(\varepsilon^{-1}(x-t))$ est d'intégrale 1. En adaptant le Lemme 2.15,

$$|f * g_\varepsilon(x) - f(x)| \leq \omega_f(\varepsilon \max\{|a|, |b|\}).$$

Ceci étant vrai pour tout $x \in \mathbb{R}$, le résultat suit immédiatement. □

On peut utiliser cette proposition pour montrer :

Proposition 2.19.
Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction continue.

Soit g une fonction continue par morceaux sur $\mathbb{R}$, nulle en dehors d'un segment $[a, b]$, et d'intégrale 1. Posons, pour tous x réel et $\varepsilon > 0$,

$$f * g_\varepsilon(x) := \varepsilon^{-1} \int_{\mathbb{R}} f(t) g(\varepsilon^{-1}(x-t)) \, \mathrm{d}t.$$

Alors $f * g_\varepsilon$ converge vers f uniformément sur tout compact quand ε tend vers 0.

Démonstration.
Il suffit de démontrer la proposition pour les compacts de la forme $[-M, M]$ avec $M > 0$. Soit $[-M, M]$ un tel segment.

Par le théorème de Heine, la restriction de f à $[-M-1, M+1]$ est uniformément continue ; notons ω_M un de ses modules de continuité. De plus, si $\varepsilon < \max\{|a|, |b|\}^{-1}$, alors $t \mapsto g(\varepsilon^{-1}(x-t))$ est nulle en-dehors de $[x-1, x+1] \subset [-M-1, M+1]$. Par conséquent, pour tout $x \in [-M, M]$,

$$f * g_\varepsilon(x) = \varepsilon^{-1} \int_{-M-1}^{M+1} f(t) g(\varepsilon^{-1}(x-t)) \, \mathrm{d}t.$$

Le calcul qui s'ensuit est le même que dans le cas précédent, et donne pour tout ε suffisamment petit

$$\sup_{x \in [-M,M]} |f * g_\varepsilon(x) - f(x)| \leq \omega_M(\varepsilon \max\{|a|, |b|\}),$$

ce que l'on voulait démontrer. $\square$

L'un des intérêt du produit de convolution est que, si f est continue et g de classe $\mathcal{C}^k$, alors $f * g_\varepsilon$ est de classe $\mathcal{C}^k$ pour tout $\varepsilon > 0$. On peut ainsi approcher uniformément (respectivement, uniformément sur tout compact) une fonction uniformément continue (respectivement, continue) par une suite de fonctions $\mathcal{C}^k$, et en particulier [4] de fonctions $\mathcal{C}^\infty$.

Il est possible d'approcher uniformément sur $\mathbb{R}$ n'importe quelle fonction continue par une suite de fonctions $\mathcal{C}^\infty$, mais cela demande un peu plus de travail et, de préférence, quelques outils supplémentaires [5].

3 Convexité

3.1 Convexité de parties

Nous partirons de :

Définition 3.1 (Convexité)**.**
Une partie C d'un espace vectoriel [6] réel (ou complexe) est dite ***convexe*** *si, pour tous $x, y \in C$ et tout $t \in [0, 1]$, on a $(1-t)x + ty \in C$.*

On note parfois $[x, y] = \{(1-t)x + ty : \ t \in [0, 1]\}$. Avec cette notation, une partie C est convexe si et seulement si $[x, y] \subset C$ pour tous $x, y \in C$.

Une première proposition utile est :

4. En utilisant des fonctions $\mathcal{C}^\infty$ à support compact, cf. la feuille d'exercices.
5. Tels que des partitions lisses de l'unité.
6. Ou affine.

Proposition 3.2.
Une intersection quelconque de convexes est convexe.

Démonstration.
Soit $C = \bigcap_{i \in \mathcal{I}} C_i$ une intersection de convexes. Soient x, $y \in C$. On veut montrer que $[x, y] \subset C$.
Soit $i \in \mathcal{I}$. Alors x, $y \in C_i$; par convexité de C_i, on a $[x, y] \subset C_i$. Donc $[x, y] \subset C_i$ pour tout $i \in \mathcal{I}$, donc $[x, y] \subset C$. $\square$

Il y a beaucoup à dire sur les convexes de $\mathbb{R}^n$. Par exemple, étant donné un convexe compact K de $\mathbb{R}^n$:

- ▷ Si K est d'intérieur non vide, étant donné un point intérieur x et $v \neq 0$, il existe un unique $r(v) > 0$ tel que $x + r(v)v \in \partial K$ (où ∂K est la frontière de K dans $\mathbb{R}^n$).
- ▷ La fonction $v \mapsto r(v)$ ainsi définie est continue (et lipschitzienne).

Un convexe compact d'intérieur non vide peut donc toujours être paramétré polairement par une fonction continue.

3.2 Enveloppe convexe et barycentres

Comme une intersection quelconque de convexes est convexe, on peut définir des "convexes minimaux" vérifiants des propriétés héritées par inclusion. En particulier,

Définition 3.3 (Enveloppe convexe)**.**
*Soit A une partie d'un espace vectoriel réel (ou complexe). L'**enveloppe convexe** de A est l'intersection de tous les convexes contenant A. C'est le plus petit convexe contenant A. On le note $Conv(A)$.*

Une notion liée est celle de barycentre :

Définition 3.4 (Barycentre)**.**
*Soit E un espace vectoriel réel (ou complexe). Soient $n \geq 1$ et $(a_1, \dots, a_n) \in E^n$. Soient $(t_1, \dots, t_n) \in [0, 1]^n$ tels que $\sum_{i=1}^n t_i = 1$. Le **barycentre** des points a_i pondérés par les t_i est le point $b = \sum_{i=1}^n t_i a_i$. On dit que b est un barycentre des points a_i.*

Remarque 3.5.
Le barycentre est une notion affine et non vectorielle ; cette propriété utilise crucialement le fait que la somme des poids vaut 1.

Proposition 3.6.
Soit C un convexe et b un barycentre de points de C. Alors $b \in C$.

Démonstration.
Il faut montrer que, pour tout $n \geq 1$, tout $(a_1, \dots, a_n) \in C^n$ et tout $(t_1, \dots, t_n) \in [0, 1]^n$ tels que $\sum_{i=1}^n t_i = 1$, le barycentre appartient encore à C. Sans perte de généralité, on peut supposer de plus les t_i non nuls. On procède par récurrence sur n.
Pour $n = 1$, il n'y a rien à montrer (le barycentre d'un point est lui-même).
Supposons la propriété vraie au rang $n \geq 1$. Soit $b = \sum_{i=1}^{n+1} t_i a_i$ avec les a_i, t_i vérifiant la propriété souhaitée. Alors

$$b = \sum_{i=1}^n t_i a_i + t_{n+1} a_{n+1} = (1 - t_{n+1}) \sum_{i=1}^n \frac{t_i}{1 - t_{n+1}} a_i + t_{n+1} a_{n+1}.$$

Par hypothèse de récurrence, $\sum_{i=1}^n \frac{t_i}{1-t_{n+1}} a_i \in C$. Par convexité, $b \in C$, ce qu'il fallait montrer. $\square$

On peut alors caractériser les enveloppes convexes !

Proposition 3.7.
Soit A une partie d'un espace vectoriel réel (ou complexe). Alors $Conv(A)$ est l'ensemble des barycentres de points de A.

Démonstration.
Soit C un convexe contenant A. Alors C contient tous les barycentres de points de C, et en particulier tous les barycentres de points de A. Donc les barycentres de points de A appartiennent à tous les convexes contenant A, et par conséquent à leur intersection $\mathrm{Conv}(A)$.

Il reste à montrer que l'ensemble des barycentres de points de A est un convexe contenant A. Le fait qu'il contienne A est trivial (un point est le barycentre de lui-même) ; il reste à montrer qu'il est convexe.

Soient b, b' deux barycentres de points de A et $t \in [0,1]$. Écrivons $b = \sum_{i=1}^m t_i a_i$ et $b' = \sum_{i=1}^n t'_i a'_i$ avec les contraintes évidentes. Alors

$$(1-t)b + tb' = \sum_{i=1}^m (1-t)t_i a_i + \sum_{i=1}^n t t'_i a'_i.$$

On vérifie que $\sum_{i=1}^m (1-t)t_i + \sum_{i=1}^n t t'_i = 1$, et donc $(1-t)b + tb'$ est encore un barycentre [7], ce qui termine cette démonstration. □

Cette caractérisation permet de décrire les enveloppes convexes de façon moyennement efficace. Pour montrer qu'un point x appartient à $\mathrm{Conv}(A)$, il suffit de l'écrire comme barycentre de points de A. Pour montrer qu'un point x n'appartient pas à $\mathrm{Conv}(A)$, il suffit de trouver un convexe contenant x mais pas A. En ce qui concerne le dernier point, dans certains cas (par exemple, quand A est ouvert), le théorème de Hahn-Banach géométrique affirme que des demi-espaces suffisent.

En pratique, la recherche d'algorithmes efficaces pour décrire l'enveloppe convexe d'un ensemble de points est non triviale.

Exercice 3.8.
Montrez que, si A est ouvert, alors $Conv(A)$ est ouverte. Donnez un exemple de partie fermée dont l'enveloppe convexe est ouverte.

Remarque 3.9.
*Il existe aussi une notion d'**enveloppe convexe fermée** d'une partie A, qui est l'intersection des convexes fermés contenant A.*

3.3 Convexité de fonctions

Passons aux fonctions convexes, en commençant par le cas réel.

Définition 3.10 (Fonction convexe)**.**
*Soit I un intervalle réel et f une fonction réelle définie sur I. On dit que f est **convexe** si son épigraphe $\{(x,t) : x \in I, t \geq f(x)\}$ est une partie convexe de $\mathbb{R}^2$. On dit que f est **concave** si $-f$ est convexe.*

De façon équivalente, f est convexe si et seulement si, pour tous x, $y \in I$ et $t \in [0,1]$,

$$f((1-t)x + ty) \geq (1-t)f(x) + tf(y).$$

7. Plus généralement, un barycentre de barycentres est un barycentre.

Exercice 3.11.
Démontrez que ces deux formulations de la convexité sont bien équivalentes.

Remarquons que la proposition selon laquelle une intersection de convexes est convexe implique immédiatemment :

Proposition 3.12.
Le maximum de deux fonctions convexes sur un même intervalle est convexe.

De plus, la notion de convexité se comporte bien vis-à-vis des barycentres. La proposition suivante se montre par récurrence, de façon similaire à ce qui a été fait pour montrer que les barycentre d'un convexe appartiennent au convexe.

Proposition 3.13.
Soit f une fonction convexe sur un intervalle I. Soient $n \geq 1$ et $(a_1, \ldots, a_n) \in I^n$. Soient $(t_1, \ldots, t_n) \in [0,1]^n$ tels que $\sum_{i=1}^n t_i = 1$. Alors

$$f\left(\sum_{i=1}^n t_i a_i\right) \leq \sum_{i=1}^n t_i f(a_i). \tag{1}$$

Cette dernière proposition se généralise par l'*inégalité de Jensen* :

Théorème 10 (Inégalité de Jensen)**.**
Soit f une fonction convexe bornée inférieurement sur un intervalle I, et X une variable aléatoire à valeurs dans I. Alors

$$f(\mathbb{E}(X)) \leq \mathbb{E}(f(X)).$$

Exemple 3.14.
Supposons que X est à valeur dans $(a_1, \ldots, a_n)$ et $\mathbb{P}(X = a_i) = t_i$. Alors on retrouve l'Équation (1).

Exemple 3.15.
Si f est la valeur absolue, on retrouve l'inégalité $|\mathbb{E}(X)| \leq \mathbb{E}(|X|)$, un excellent moyen mnémotechnique pour retenir le sens de l'inégalité.

3.4 Caractérisations de la convexité

La notion de convexité s'applique à toute fonction définie sur un intervalle. Comme nous le verrons bientôt, elle s'applique essentiellement à des fonctions continues. Dans le cas de fonctions dérivables ou deux fois dérivables, on dispose de caractérisations de la convexité.

Nous commençons par le critère de convexité de fonctions dérivables. Pour cela, nous utiliserons le lemme très utile suivant portant sur les taux d'accroissements :

Lemme 3.16.
Soit f une fonction convexe sur un intervalle I. Soit $x_0 \in I$. Alors la fonction $x \mapsto \frac{f(x)-f(x_0)}{x-x_0}$ est croissante[8] *sur $I \setminus \{x_0\}$.*

Démonstration.
Soient $x < y$ deux points de $I \setminus \{x_0\}$. Si $x_0 < x$, on écrit x comme barycentre de x_0 et y :

$$x = \frac{y-x}{y-x_0} x_0 + \frac{x-x_0}{y-x_0} y.$$

8. Attention : elle est bien croissante sur cet ensemble entier, pas séparément sur chacune de ses composantes connexes.

Par inégalité de convexité,

$$f(x) \leq \frac{y-x}{y-x_0} f(x_0) + \frac{x-x_0}{y-x_0} f(y),$$

et donc

$$\frac{f(x)-f(x_0)}{x-x_0} \leq \frac{f(y)-f(x_0)}{y-x_0},$$

ce qu'il fallait montrer. Les cas $x_0 \in (x, y)$ et $x_0 < y$ se traitent de même. □

Il s'ensuit :

Proposition 3.17.
Soit f une fonction dérivable sur I. Alors f est convexe si et seulement si f' est croissante.

Démonstration.
Supposons dans un premier temps que f est convexe, et montrons que f' est croissante. Soient $x < y$ deux points de I. Soit $\varepsilon > 0$. Il existe $h \in (0, y-x]$ tel que $\frac{f(x+h)-f(x)}{h} \geq f'(x) - \varepsilon$ et $\frac{f(y)-f(y-h)}{h} \leq f'(y) + \varepsilon$.
Mais la fonction $z \mapsto \frac{f(z)-f(x)}{z-x}$ est croissante par le Lemme 3.16 et $x + h \leq y$, donc

$$f'(x) - \varepsilon \leq \frac{f(x+h)-f(x)}{h} \leq \frac{f(y)-f(x)}{y-x}.$$

De même, la fonction $z \mapsto \frac{f(z)-f(y)}{z-y} = \frac{f(y)-f(z)}{y-z}$ est croissante et $x \leq y - h$, donc

$$\frac{f(y)-f(x)}{y-x} \leq \frac{f(y)-f(y-h)}{h} \leq f'(y) + \varepsilon.$$

On obtient finalement $f'(x) - \varepsilon \leq f'(y) + \varepsilon$. Le paramètre ε étant arbitraire, $f'(x) \leq f'(y)$, ce que l'on voulait démontrer.
Supposons maintenant que f' est croissante. Soient $x < y$ deux points de I. Posons $g(t) := f((1-t)x + ty) - (1-t)f(x) - tf(y)$ pour tout $t \in [0, 1]$. Cette fonction est continue, dérivable, et notre but est de montrer que g est négative.
Mais $g(0) = g(1) = 0$, donc, par le théorème de Rolle, il existe $c \in (0, 1)$ tel que $g'(c) = 0$. De plus, f' est croissante donc g' est croissante, donc $g' \leq 0$ sur $[0, c]$ et $g' \geq 0$ sur $[c, 1]$.

x	0		c		1
$g'(x)$		$-$	0	$+$	
$g(x)$	0	$\searrow$	g(c)	$\nearrow$	0

On déduit de ce tableau de variations que g est négative sur $[0, 1]$, ce qu'il fallait montrer. □

Proposition 3.18.
Soit $f \in \mathcal{C}^1(I, \mathbb{R})$ telle que f' soit dérivable. Alors f est convexe si et seulement si f'' est positive.

Démonstration.
En effet, f' est croissante si et seulement si f'' est positive, et l'on applique ensuite la proposition précédente. □

Remarque 3.19.
La définition de la convexité et ces deux caractérisations s'appliquent à des fonctions de différentes régularités. Si f est de classe $\mathcal{C}^2$, le critère de dérivée seconde positive est souvent plus facile à manipuler. Ceci dit, ce critère ne s'applique pas[9] *aux fonctions continues. Le bon critère à utiliser dépend donc de la régularité des fonctions considérées.*

Souvent, une propriété des fonctions convexes admet plusieurs démonstrations, relativement aisées pour des fonctions de classe $\mathcal{C}^2$ et plus subtiles pour des fonctions quelconques.

Remarque 3.20.
Ces caractérisations permettent aussi de démontrer des propriétés a priori peu évidentes. Par exemple, si $(f_n)_{n\in\mathbb{N}}$ est une suite de fonctions de classe $\mathcal{C}^2$ convexes convergeant uniformément vers une fonction f de classe $\mathcal{C}^2$, alors $f'' \geq 0$ – autrement dit, la positivité de la dérivée seconde passe aux limites uniformes (et non seulement aux limites $\mathcal{C}^2$).

3.5 Convexité en dimension supérieure

La convexité s'applique à des fonctions définies sur n'importe quel espace vectoriel[10] réel (ou complexe).

Définition 3.21 (Fonction convexe)**.**
*Soit C un convexe et f une fonction réelle définie sur C. On dit que f est **convexe** si son épigraphe $\{(x,t) :\ x \in C, t \geq f(x)\}$ est une partie convexe de $C \times \mathbb{R}$. On dit que f est **concave** si $-f$ est convexe.*

De façon équivalente, f est convexe si et seulement si, pour tous x, $y \in C$ et $t \in [0,1]$,

$$f((1-t)x+ty) \geq (1-t)f(x)+tf(y).$$

La caractérisation des fonctions convexes de classe $\mathcal{C}^2$ s'adapte en dimension supérieure.

Définition 3.22 (Matrice hessienne)**.**
*Soit U un ouvert de $\mathbb{R}^n$ et $f \in \mathcal{C}^2(U,\mathbb{R})$. La **matrice hessienne** de f en un point $x \in U$ est la matrice $n \times n$ d ses dérivées partielles d'ordre 2 :*

$$\nabla^2 f(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right)_{1\leq i,j\leq n}.$$

Par le lemme de Schwarz, cette matrice est symétrique.

Proposition 3.23.
Soit C un ouvert convexe de $\mathbb{R}^n$ et $f \in \mathcal{C}^2(C,\mathbb{R})$. La fonction f est convexe si et seulement si, pour tout $x \in U$, la matrice $\nabla^2 f(x)$ est positive[11]*.*

Démonstration.
On utilise la caractérisation obtenue en dimension 1. Soient x, $y \in C$. Posons $g(t) := f((1-t)x+ty)$ pour tout $t \in [0,1]$. Alors f satisfait l'inégalité de convexité entre x et y si et seulement si g est convexe sur $[0,1]$. Or g est de classe $\mathcal{C}^2$, et est donc convexe si et seulement si $g'' \geq 0$. Un calcul donne pour tout $t \in [0,1]$:

$$g''(t) = \langle (y-x), \nabla^2 f((1-t)x+ty)(y-x)\rangle.$$

9. Du moins, sans un minimum de complications.
10. Comme pour les notions de convexité de partie ou de barycentre, elle s'applique en fait à des espaces affines.
11. Au sens que toutes ses valeurs propres sont positives ou nulles.

Si $\nabla^2 f(x)$ est positive pour tout $x \in C$, alors f est convexe. La réciproque demande encore un peu de travail. Soit $x \in C$. La matrice $\nabla^2 f(x)$ étant positive si et seulement si $\langle v, \nabla^2 f(x) v\rangle \geq 0$ pour tout $v \in \mathbb{R}^n$, fixons $v \in \mathbb{R}^n$ non nul. Comme C est ouvert, il existe $\varepsilon > 0$ tel que $B(x, \varepsilon) \subset C$. Posons alors $y := x + \varepsilon \frac{v}{\|v\|}$ et $t = 0$. Le calcul précédent implique alors que

$$g''(0) = \frac{\varepsilon^2}{\|v\|^2} \langle v, \nabla^2 f(x) v\rangle \geq 0,$$

et donc $\langle v, \nabla^2 f(x) v\rangle \geq 0$, ce qu'il fallait démontrer. □

3.6 Application 1 : Continuité, monotonie et limites

3.6.1 Continuité sur les ouverts convexes

Les fonctions convexes sont automatiquement continues (à une subtilité près).

Proposition 3.24.
Soit f une fonction convexe sur un intervalle I. Alors f est continue à l'intérieur de I.

Démonstration.
Soit x un point intérieur à I ; montrons que f est continue en x. Comme x est intérieur, il existe $y < x$ dans I. Alors :

▷ Pour tout $z \in [y, x]$: on sait que $z = \frac{x-z}{x-y} y + \frac{z-y}{x-y} x$. Par inégalité de convexité,

$$f(z) \leq \frac{x-z}{x-y} f(y) + \frac{z-y}{x-y} f(x).$$

▷ Pour tout $z \geq x$: on sait que $x = \frac{x-y}{z-y} y + \frac{z-x}{z-y} z$. Par inégalité de convexité,

$$f(x) \leq \frac{z-x}{z-y} f(y) + \frac{x-y}{z-y} f(z),$$
$$f(z) \geq \frac{z-y}{x-y} f(x) - \frac{z-x}{x-y} f(y).$$

De même, soit $y' > x$ dans I. Alors on obtient deux autres bornes (une borne inférieure sur $[y, x]$, et une borne supérieure sur $[x, y']$. En combinant ces quatres bornes, pour tout $z \in [y, x]$,

$$\frac{z-y'}{x-y'} f(x) - \frac{z-x}{x-y'} f(y') \leq f(z) \leq \frac{x-z}{x-y} f(y) + \frac{z-y}{x-y} f(x),$$

et en particulier $\lim_{z \to x^-} f(z) = f(x)$. De même, on montre que $\lim_{z \to x^+} f(z) = f(x)$, et donc que f est continue en x. □

Cette proposition reste valable en dimension supérieure !

Proposition 3.25 ([FGN·Ana3, Exercice 1.25])**.**
Soit f une fonction convexe sur un convexe $C \subset \mathbb{R}^n$. Alors f est continue à l'intérieur de C.

Démonstration.
Supposons dans un premier temps f bornée supérieurement par M. Soit x un point intérieur de C et $\varepsilon > 0$ tel que $\overline{B}(x, \varepsilon) \subset C$. Alors, par le même calcul qu'en dimension 1, pour tout $z \in \overline{B}(x, \varepsilon)$,

$$\frac{\varepsilon - \|z-x\|}{\varepsilon} f(x) - \frac{\|z-x\|}{\varepsilon} M \leq f(z) \leq \frac{\|z-x\|}{\varepsilon} M + \frac{\varepsilon - \|z-x\|}{\varepsilon} f(x).$$

Le problème suivant consiste à éliminer l'hypothèse que f est bornée supérieurement. Soit x un point intérieur à C. Soient $(a_i)_{1\leq i\leq n+1}$ des points de C tels que x appartienne à l'enveloppe convexe C' des a_i. Alors, par l'inégalité de convexité sur les barycentre, pour tout $y \in C'$,

$$f(y) \leq \max_{1\leq i\leq n+1} f(a_i).$$

Donc f est continue à l'intérieur de C', et en particulier f est continue en x.

Finalement, f est bien continue à l'intérieur de C. □

3.6.2 Comportement au bord d'un intervalle

La continuité au bord est moins évidente. Par exemple, la fonction $f : [0, 1] \to \mathbb{R}$ valant 0 sur $(0, 1)$ et 1 sur $\{0, 1\}$ est convexe, continue à l'intérieur de $[0, 1]$, mais pas continue au bord de l'intervalle. Le comportement au bord va être clarifié grâce à :

Proposition 3.26 ([FGN·Ana1, Exercice 4.44])**.**
Soit f une fonction convexe sur un intervalle I ouvert. Alors :
- ▷ *Ou f est monotone sur I ;*
- ▷ *Ou il existe $a \in I$ tel que f est décroissante sur $(+\infty, a] \cap I$ et croissante sur $[a, +\infty) \cap I$.*

Démonstration.
Supposons f non monotone. S'il existe $x < y < z$ dans I tels que $f(x) < f(y)$ et $f(z) < f(y)$, alors f n'est pas convexe. Donc il existe $x < y < z$ dans I tels que $f(y) < f(x)$ et $f(y) < f(z)$.

La fonction f étant continue sur I, elle atteint son minimum sur le segment $[x, z]$ en un point $a \in (x, y)$.

Soit b' un point de I tel que $f(b') < f(a)$. Alors $f(c) < f(a)$ pour tout c dans l'intervalle d'extrémités a et b', ce qui contredit la minimalité de $f(a)$ sur intervalle (x, y). donc a est un minimiseur global de f.

Soient $a \leq x' < y'$. Comme $f(a) \leq f(y')$, on a $f(x') \leq f(y')$ par convexité. Donc f est croissante sur $[a, +\infty) \cap I$. On montre de même que f est décroissante sur $(+\infty, a] \cap I$. □

Corollaire 3.27.
Soit f une fonction convexe sur un intervalle I. Alors f admet des limites au bord de cet intervalle. De plus, si le bas de l'intervalle a appartient à I,

$$f(a) \geq \lim_{x\to a^+} f(x),$$

et de même pour le haut de l'intervalle.

Démonstration.
L'existence des limites vient de la proposition précédente et de la monotonie. Si de plus f est définie au bord de l'intervalle, la convexité fournit une borne supérieure par une fonction affine qui implique l'inégalité voulue [12]. □

3.6.3 Comportement asymptotique

On peut aussi analyser le comportement asymptotique de fonctions convexes définies sur un intervalle infini.

12. Mais pas de borne inférieure – pour cela, il faudrait que f soit définie au-delà de a !

Proposition 3.28 ([FGN·Ana1, Exercice 4.45][Gou, Chapitre 2.3, Exercice 2][Rom, Exercice 8.4])**.**

Soit $a > 0$ et f une fonction convexe sur $(a, +\infty)$. Alors la limite $\lim_{x\to+\infty} \frac{f(x)}{x}$ existe et appartient à $\mathbb{R} \cup \{+\infty\}$.

Démonstration.
Par le Lemme 3.16, la fonction $x \mapsto \frac{f(x)-f(a+1)}{x-a-1}$ est croissante sur $(a+1, +\infty)$. Elle admet donc une limite ℓ appartenant à $\mathbb{R} \cup \{+\infty\}$.
Or $\frac{f(x)-f(a+1)}{x} = \frac{x-a-1}{x} \cdot \frac{f(x)-f(a+1)}{x} - a - 1$, donc $\lim_{x\to+\infty} \frac{f(x)-f(a+1)}{x} = \ell$ comme produit de limites. Enfin, $\lim_{x\to+\infty} \frac{f(x)}{x} = \ell$ comme somme de limites. □

Si la limite précédente est finie, alors on peut préciser le comportement asymptotique.

Proposition 3.29.
Soit $a > 0$ et f une fonction convexe sur $(a, +\infty)$. Posons $\ell := \lim_{x\to+\infty} \frac{f(x)}{x}$.
Si ℓ est finie, alors $\lim_{x\to+\infty}(f(x) - \ell x)$ existe et appartient à $\mathbb{R} \cup \{-\infty\}$.

Démonstration.
Posons $g(x) := f(x) - \ell x$. Alors g est convexe, $\lim_{x\to+\infty} \frac{g(x)}{x} = 0$ et on veut montrer que $\lim_{x\to+\infty} g(x)$ existe et appartient à $\mathbb{R} \cup \{-\infty\}$.
Soient $a < x < y$. La fonction $z \mapsto \frac{g(z)-g(x)}{z-x}$ est croissante par le Lemme 3.16 et a pour limite 0, donc est négative. En particulier, $\frac{g(y)-g(x)}{y-x} \leq 0$, donc $g(y) \leq g(x)$.
La fonction g est donc décroissante. Elle admet donc une limite en $+\infty$, dont la valeur appartient à $\mathbb{R} \cup \{-\infty\}$. □

Exercice 3.30.
Donnez des exemples de chaque situation autorisée par ces deux propositions :
- ▷ $\lim_{x\to+\infty} \frac{f(x)}{x} = +\infty$ *;*
- ▷ $\lim_{x\to+\infty} \frac{f(x)}{x} = \ell$ *est finie et* $\lim_{x\to+\infty} f(x) - \ell x = -\infty$ *;*
- ▷ $\lim_{x\to+\infty} \frac{f(x)}{x} = \ell$ *est finie et* $\lim_{x\to+\infty} f(x) - \ell x$ *est finie.*

3.7 Application 2 : Moyennes

Rappelons différents types de moyennes :

Définition 3.31. *Soient $(a_1, \dots, a_n)$ des réels strictements positifs. Leur* ***moyenne arithmétique*** *est définie par*

$$A = \frac{1}{n} \sum_{k=1}^{n} a_k.$$

Leur ***moyenne géométrique*** *est définie par*

$$G = \left(\prod_{k=1}^{n} a_k \right)^{\frac{1}{n}}.$$

Leur ***moyenne harmonique*** *est définie par*

$$\frac{1}{H} = \frac{1}{n} \sum_{k=1}^{n} \frac{1}{a_k}.$$

Proposition 3.32 (Inégalités entre moyennes)**.**

$$H \leq G \leq A.$$

Démonstration.
La fonction exp est convexe sur $\mathbb{R}$. On applique l'inégalité de convexité aux points $-\ln(a_k)$ pondérés par $1/n$:

$$\begin{aligned}
\exp\left(\frac{1}{n}\sum_{k=1}^{n}(-\ln(a_k))\right) &\leq \frac{1}{n}\sum_{k=1}^{n}\exp(-\ln(a_k)) \\
\frac{1}{\left(\prod_{k=1}^{n} a_k\right)^{\frac{1}{n}}} &\leq \frac{1}{\frac{1}{n}\sum_{k=1}^{n} a_k} \\
G &\geq H.
\end{aligned}$$

On applique l'inégalité de convexité aux points $\ln(a_k)$ pondérés par $1/n$:

$$\begin{aligned}
\exp\left(\frac{1}{n}\sum_{k=1}^{n}\ln(a_k)\right) &\leq \frac{1}{n}\sum_{k=1}^{n}\exp(\ln(a_k)) \\
\left(\prod_{k=1}^{n} a_k\right)^{\frac{1}{n}} &\leq \frac{1}{n}\sum_{k=1}^{n} a_k \\
G &\leq A.
\end{aligned}$$

□

Exercice 3.33.
Donner des exemples concrets dans lequels chacune de ces trois moyennes peut apparaître.

Remarque 3.34.
L'avantage de cette approche, comparée à des approches plus algébriques, est qu'elle se prête bien aux généralisations. Par exemple, il est très facile de donner des inégalités entre moyennes pondérées, en changeant simplement les poids des barycentres.

3.8 Application 3 : Young, Hölder, Minkowski

Présentons 3 inégalités : Young, Hölder et Minkowski. Dans ce qui suit, p et q sont dans $[1, \infty]$, et tels que $\frac{1}{p} + \frac{1}{q} = 1$ (on dit que p et q sont **conjugués**). Nous adoptons la convention $1/\infty = 0$.

Théorème 11 (Inégalité de Young [FGN·Ana3, Exercice 4.19] [Moi, Exercice 5 p. 135] [Ska, Exercice 7.8])**.**
Soient a, b deux réels positifs, et p, $q \in (1, +\infty)$ conjugués. Alors

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Démonstration.
On applique une inégalité de concavité au logarithme :

$$\ln\left(\frac{1}{p}a^p + \frac{1}{q}b^q\right) \geq \frac{1}{p}\ln(a^p) + \frac{1}{q}\ln(b^q) = \ln(a) + \ln(b).$$

En prenant l'exponentielle, on obtient l'inégalité voulue :

$$\frac{1}{p}a^p + \frac{1}{q}b^q \geq ab \qquad \square$$

L'inégalité de Young généralise l'inégalité plus classique $ab \leq \frac{a^2+b^2}{2}$, qui correspond au cas particulier important $p = q = 2$.

Théorème 12 (Inégalité de Hölder [FGN·Ana3, Exercice 4.19] [Moi, Exercice 5 p. 135] [Rom, Exercice 2.1] [Ska, Exercice 7.8])**.**
Soient f, g deux fonctions continues par morceaux sur un segment $[a,b]$. Soient p, $q \in (1,+\infty)$ conjugués. Alors

$$\int_a^b |f(x)g(x)| \, \mathrm{d}x \leq \left(\int_a^b |f(x)|^p \, \mathrm{d}x\right)^{\frac{1}{p}} \left(\int_a^b |g(x)|^q \, \mathrm{d}x\right)^{\frac{1}{q}}.$$

Démonstration.
Soit $\lambda > 0$. Remarquons qu'en appliquant l'inégalité de Young à λa et $\lambda^{-1}b$, on trouve

$$ab \leq \frac{\lambda^p}{p}a^p + \frac{\lambda^{-q}}{q}b^q.$$

Appliquons cette inégalité à $a = |f(x)|$ et $b = |g(x)|$, à x fixé, puis intégrons sur $[a,b]$.

$$\int_a^b |f(x)g(x)| \, \mathrm{d}x \leq \frac{\lambda^p}{p}\int_a^b |f(x)|^p \, \mathrm{d}x + \frac{\lambda^{-q}}{q}\int_a^b |g(x)|^q \, \mathrm{d}x.$$

On va choisir la meilleure valeur de λ possible. Cette fonction de λ est continue sur $\mathbb{R}_+^*$, tend vers $+\infty$ en 0 et $+\infty$, donc atteint son minimum. Elle est dérivable, donc son minimum est un point critique. Sa dérivée en λ valant

$$\lambda^{p-1}\int_a^b |f(x)|^p \, \mathrm{d}x - \lambda^{-(q+1)}\int_a^b |g(x)|^q \, \mathrm{d}x,$$

Le paramètre λ minimisant le membre de gauche est

$$\lambda_* = \left(\frac{\int_a^b |g(x)|^q \, \mathrm{d}x}{\int_a^b |f(x)|^p \, \mathrm{d}x}\right)^{\frac{1}{p+q}}.$$

En réinjectant ce choix de λ dans l'inégalité de Young, on trouve finalement

$$\begin{aligned}
\int_a^b |f(x)g(x)| \, \mathrm{d}x &\leq \frac{1}{p}\left(\frac{\int_a^b |g(x)|^q \, \mathrm{d}x}{\int_a^b |f(x)|^p \, \mathrm{d}x}\right)^{\frac{p}{p+q}} \int_a^b |f(x)|^p \, \mathrm{d}x + \frac{1}{q}\left(\frac{\int_a^b |g(x)|^q \, \mathrm{d}x}{\int_a^b |f(x)|^p \, \mathrm{d}x}\right)^{\frac{-q}{p+q}} \int_a^b |g(x)|^q \, \mathrm{d}x \\
&= \left(\frac{1}{p}+\frac{1}{q}\right)\left(\int_a^b |g(x)|^q \, \mathrm{d}x\right)^{\frac{p}{p+q}} \left(\int_a^b |f(x)|^p \, \mathrm{d}x\right)^{\frac{q}{p+q}} \\
&= \left(\int_a^b |g(x)|^q \, \mathrm{d}x\right)^{\frac{p}{p+q}} \left(\int_a^b |f(x)|^p \, \mathrm{d}x\right)^{\frac{q}{p+q}},
\end{aligned}$$

où l'on a utilisé l'égalité $p^{-1} + q^{-1}$ à la dernière ligne. Enfin, comme $p + q = pq$, on peut simplifier encore les exposants de la dernière ligne pour obtenir l'inégalité de Hölder. $\square$

Remarque 3.35.
Pour $p = q = 2$, on retrouve l'inégalité de Cauchy-Schwarz.

Remarque 3.36.
Cette inégalité est aussi valide pour des suites (finies ou infinies) en remplaçant les intégrales par des sommes.

Remarque 3.37.
Notons[13] *pour $p \in [1, +\infty)$,*

$$\|f\|_{\mathbb{L}^p([a,b])} := \left(\int_a^b |f(x)|^p \, \mathrm{d}x\right)^{\frac{1}{p}}.$$

Alors l'inégalité de Hölder s'écrit plus brièvement

$$\|fg\|_{\mathbb{L}^1([a,b])} \leq \|f\|_{\mathbb{L}^p([a,b])} \|g\|_{\mathbb{L}^q([a,b])}.$$

On déduit de l'inégalité de Hölder une troisième inégalité, dite de Minkowski.

Théorème 13 (Inégalité de Minkowski [FGN·Ana3, Exercice 4.19] [Moi, Exercice 5 p. 135] [Rom, Exercice 2.1])**.**
Soient f, g deux fonctions continues par morceaux sur un segment $[a, b]$. Soit $p \in [1, +\infty)$. Alors

$$\left(\int_a^b |f(x) + g(x)|^p \, \mathrm{d}x\right)^{\frac{1}{p}} \leq \left(\int_a^b |f(x)|^p \, \mathrm{d}x\right)^{\frac{1}{p}} + \left(\int_a^b |g(x)|^p \, \mathrm{d}x\right)^{\frac{1}{p}}.$$

Démonstration.
Pour $p = 1$, il s'agit de l'inégalité triangulaire pour les intégrales. Supposons $p > 1$. On sait que

$$\int_a^b |f(x) + g(x)|^p \, \mathrm{d}x \leq \int_a^b |f(x)| \cdot |f(x) + g(x)|^{p-1} \, \mathrm{d}x + \int_a^b |g(x)| \cdot |f(x) + g(x)|^{p-1} \, \mathrm{d}x.$$

On utilise l'inégalité de Hölder sur chacune des deux intégrales du membre de droite. L'exposant conjugué de p est $q = \frac{p}{p-1}$. Alors la première intégrale devient

$$\begin{aligned}
\int_a^b |f(x)| \cdot |f(x) + g(x)|^{p-1} \, \mathrm{d}x &\leq \left(\int_a^b |f(x)|^p \, \mathrm{d}x\right)^{\frac{1}{p}} \left(\int_a^b \left(|f(x) + g(x)|^{p-1}\right)^{\frac{p}{p-1}} \, \mathrm{d}x\right)^{\frac{p-1}{p}} \\
&= \left(\int_a^b |f(x)|^p \, \mathrm{d}x\right)^{\frac{1}{p}} \left(\int_a^b |f(x) + g(x)|^p \, \mathrm{d}x\right)^{\frac{p-1}{p}}.
\end{aligned}$$

La seconde intégrale se manipule similairement. On obtient

$$\int_a^b |f(x) + g(x)|^p \, \mathrm{d}x \leq \left(\int_a^b |f(x) + g(x)|^p \, \mathrm{d}x\right)^{\frac{p-1}{p}} \left[\left(\int_a^b |f(x)|^p \, \mathrm{d}x\right)^{\frac{1}{p}} + \left(\int_a^b |g(x)|^p \, \mathrm{d}x\right)^{\frac{1}{p}}\right]$$

$$\left(\int_a^b |f(x) + g(x)|^p \, \mathrm{d}x\right)^{\frac{1}{p}} \leq \left(\int_a^b |f(x)|^p \, \mathrm{d}x\right)^{\frac{1}{p}} + \left(\int_a^b |g(x)|^p \, \mathrm{d}x\right)^{\frac{1}{p}}. \qquad \square$$

13. Ce sera justifié par l'inégalité de Minkowski

Remarque 3.38.
En reprenant les notations précédentes, l'inégalité de Minkowski s'écrit

$$\|f+g\|_{\mathbb{L}^p([a,b])} \leq \|f\|_{\mathbb{L}^p([a,b])} + \|g\|_{\mathbb{L}^p([a,b])} \,.$$

Ainsi, $\|\cdot\|_{\mathbb{L}^p([a,b])}$ est une semi-norme sur l'espace des fonctions continues par morceaux (l'homogénéité est comparativement facile à vérifier).

Pour que $\|\cdot\|_{\mathbb{L}^p([a,b])}$ soit une norme, il reste à la définir sur un espace tel qu'elle soit définie positive. Cela peut se faire ou bien en restreignant l'ensemble de définition (par exemple aux polynômes ou aux fonctions continues), ou bien en identifiant les fonctions qui diffèrent sur un ensemble de mesure nulle (espaces de Lebesgue).

3.8.1 Un peu d'analyse dimensionnelle

L'inégalité de Hölder est un bon point pour introduire un peu d'analyse dimensionnelle. Celle-ci est classique en physique, mais aussi très utile en mathématiques ! Dans les grandes lignes, *une unité correspond à une action de groupe.*

Partons d'un exemple : on trace la distance parcourue par une hirondelle en fonction du temps. Si l'on se donne des unités de temps et de distance, on obtient une fonction f de $\mathbb{R}$ dans $\mathbb{R}$, par exemple la distance parcourue par l'hirondelle (en **km**) en fonction du temps (en **h**).

Du point de vue mathématique, les notions sensibles d'espace ou de temps n'ont pas de signification, de même que l'hirondelle. Cependant, il est possible de changer d'unité. Par exemple, si l'on prend comme unité de distance le mètre (divisant donc par 1000 cette unité), alors les valeurs numériques de distances sont partout multipliées par 1000 ; en particulier, f est remplacée par $1000f$.

De ce point de vue, *une quantité représentant une longueur est une quantité qui est multipliée par λ quand on divise l'unité de longueur par λ.* On peut même parler d'aires et de volumes : ce sont des quantités multipliées respectivement par λ^2 et λ^3 sous cette opération.

Dans certains arguments mathématiques, il peut être très intéressant de vérifier si des égalités ou inégalités restent vraies si l'on multiplie une des quantités par un facteur λ. L'utilisation d'unités fictives est un moyen très simple de garder une trace de la dépendance de chaque terme en λ. En particulier, pour qu'une égalité soit toujours vraie dans un espace vectoriel entier [14], il faut qu'elle soit dimensionnellement cohérente !

Exemple 3.39.
Soit $f \in \mathcal{C}^1(\mathbb{R},\mathbb{R})$. On peut mettre comme unité fictive la seconde à la source et le mètre à l'arrivée. Diviser l'unité de distance par un facteur λ revient à remplacer f par λf, donc f est en mètres (et sa valeur n'a pas de dépendance en secondes).

La fonction f' a toujours cette dépendance en l'unité de distance. De plus, diviser l'unité de temps par un facteur μ revient à remplacer f par $t \mapsto f(\mu^{-1}t)$, dont la dérivée est $t \mapsto \mu^{-1}f'(\mu^{-1}t)$. La fonction f' est donc elle aussi divisée par un facteur μ. Par conséquent, f' est en $m \cdot s^{-1}$, autrement dit en mètres par seconde.

Exemple 3.40.
Dans la démonstration de l'inégalité de Hölder, une application directe de l'inégalité de Young donnerait

$$\int_a^b |f(x)g(x)|\,\mathrm{d}x \leq \frac{1}{p}\int_a^b |f(x)|^p\,\mathrm{d}x + \frac{1}{q}\int_a^b |g(x)|^q\,\mathrm{d}x.$$

14. Ou dans un cône d'un espace vectoriel, par exemple l'ensemble des fonctions positives. Ce que l'on n'a pas le droit de faire, c'est de travailler avec des quantités normées, ou à valeurs dans un intervalle fixé.

Si f est (par exemple) en m et g en m^{-1}, cette inégalité n'est pas dimensionnellement cohérente : le membre de gauche est en m^0, celui de droite est somme d'un terme en m^p et d'un terme en m^{-q}. Cela ne signifie pas que l'inégalité est fausse, mais simplement qu'en remplaçant f par λf et g par $\lambda^{-1} g$, cette inégalité peut être plus ou moins forte. On est donc poussé à l'optimiser en fonction de lambda, ce qui conduit à l'inégalité de Hölder.

D'autres choix d'unités auraient été possibles (par exemple, f en m et g sans unité, ou f et g toutes deux en m). Ces choix conduisent à l'introduction de paramètres λ différents, mais la logique de la preuve et le résultat final restent les mêmes.

Il est intéressant de remarquer que l'inégalité de Hölder est bien dimensionnellement cohérente : si f est en m et g en s, alors les membres de gauche et de droite de l'inégalité $\|fg\|_{\mathbb{L}^1([a,b])} \leq \|f\|_{\mathbb{L}^p([a,b])} \|g\|_{\mathbb{L}^q([a,b])}$ sont bien en $m \cdot s$.

3.9 Application 4 : Transformée de Fenchel-Legendre

La transformée de Fenchel-Legendre a des applications très importantes en physique mathématiques [15] ainsi qu'en probabilités [16]. Nous en tirons quelques propositions élémentaires.

Définition 3.41 (Transformée de Fenchel-Legendre)**.**
*Soit $f : I \to \mathbb{R}$ une fonction convexe. Sa **transformée de Fenchel-Legendre** est la fonction*

$$f^*(\ell) := \sup_{x \in I} \{\ell x - f(x)\}.$$

Son domaine de définition est $I^* := \{\ell \in \mathbb{R} : \ f^*(\ell) < +\infty\}$.

Exemple 3.42.
Si $f(x) = Cx^2$ pour $x \in \mathbb{R}$ (avec $C > 0$), alors $f^*(\ell) = \frac{\ell^2}{4C}$ pour $\ell \in \mathbb{R}$.
Si $f(x) = e^{Cx}$ pour $x \in \mathbb{R}$ (avec $C > 0$), alors $f^*(\ell) = \frac{\ell}{C}[\ln(\ell/C) - 1]$.
Si $f(x) = C|x|$ pour $x \in \mathbb{R}$ (avec $C > 0$), alors f^* est la fonction valant 0 sur $I^* = [-C, C]$. On remarque au passage que l'information de f se retrouve encodée non pas dans la valeur de f^*, mais dans son domaine de définition.

Exercice 3.43.
Construisez graphiquement $f^*(\ell)$ pour une valeur donnée de ℓ.

Proposition 3.44.
La transformée de Fenchel-Legendre d'une fonction convexe est convexe, et son domaine de définition est un intervalle.

Démonstration.
Soient ℓ, $\ell' \in I^*$ et $t \in [0,1]$. Alors, pour tout $x \in I$,

$$f^*(\ell) \geq \ell x - f(x) \quad \text{ et } \quad f^*(\ell') \geq \ell' x - f(x).$$

Par combinaison linéaire positive de ces deux inégalités,

$$(1-t) f^*(\ell) + t f^*(\ell') \geq [(1-t)\ell + t\ell'] x - f(x).$$

Ceci étant valable pour tout $x \in I$,

$$(1-t) f^*(\ell) + t f^*(\ell') \geq \sup_{x \in I} \{[(1-t)\ell + t\ell'] x - f(x)\} = f^*((1-t)\ell + t\ell').$$

15. Par exemple pour passer d'une formulation lagrangienne à une formulation hamiltonienne de la mécanique classique.

16. Principes de grandes déviations, qui ne sont pas sans lien avec la physique mathématique.

De plus, cette dernière quantité est finie, donc $(1-t)\ell + t\ell' \in I^*$. Ceci étant valable pour tout $t \in [0,1]$, le domaine I^* est un intervalle et f^* est convexe. $\square$

Proposition 3.45.
Soit f une fonction convexe sur un intervalle I. Alors $I \subset (I^*)^*$, et $(f^*)^*(x) = f(x)$ pour tout x dans l'intérieur de I.

Démonstration.
Soit $x \in I$. Alors, pour tout $\ell \in I^*$,

$$\begin{aligned} f^*(\ell) &\geq \ell x - f(x) \\ f(x) &\geq x\ell - f^*(\ell). \end{aligned}$$

En prenant le supremum sur $\ell \in I^*$, on trouve $f(x) \geq (f^*)^*(x)$. En particulier, $x \in (I^*)^*$. Ceci étant valable pour tout $x \in I$, on a de plus $I \subset (I^*)^*$.

Supposons de plus que x est un point intérieur de I. Rappelons que la fonction $y \mapsto \frac{f(y)-f(x)}{y-x}$ est croissante sur $I \setminus \{x\}$; soit ℓ sa limite à droite en x, qui est nécessairement finie. Alors $f(y) - f(x) \geq \ell(y-x)$ pour tout $y \in I$ (on fera attention au signe de $y-x$), ou autrement dit, $\ell y - f(y) \leq \ell x - f(x)$. Par conséquent, $f^*(\ell) = \sup_{y \in I}\{\ell y - f(y)\} = \ell x - f(x)$. En réordonnant cette inégalité, on trouve $f(x) = \ell x - f^*(\ell)$, donc $f(x) \leq (f^*)^*(x)$ par passage au supremum sur ℓ. Comme on dispose des inégalités dans les deux sens, $(f^*)^*(x) = f(x)$. $\square$

Remarque 3.46.
On a utilisé dans cette démonstration le résultat intermédiaire suivant : si x est un point intérieur au domaine I d'une fonction convexe, alors il existe une sous-tangente passant par x, c'est-à-dire une droite passant par I et en-dessous (au sens large) du graphe de f. Si l'on suppose ce résultat connu, la démonstration est significativement plus simple.

Remarque 3.47.
On peut montrer que, si x est un point du bord de I, alors $(f^*)^*(x) = f(x)$ si et seulement si f est continue en x.

3.10 Application 5 : Théorème de Gauss-Lucas

Le théorème de Gauss-Lucas permet de localiser simplement les racines complexes de la dérivée P' d'un polynôme en fonction de celles de P.

Théorème 14 (Gauss-Lucas).
Soit $P \in \mathbb{C}[X]$ un polynôme non constant. Les racines de P' appartiennent à l'enveloppe convexe des racines de P.

Démonstration.
Soit $Q = \frac{P'}{P}$ la dérivée logarithmique de P, définie sur $\{z \in \mathbb{C} : \; P(z) \neq 0\}$. Alors, si l'on factorise $P(X) = a \prod_{i=1}^k (X - z_i)^{m_i}$,

$$Q(z) = \sum_{i=1}^k \frac{m_i}{z - z_i} = \sum_{i=1}^k \frac{m_i}{|z - z_i|^2} \overline{z - z_i}.$$

Soit z une racine de P'. Si $P(z) = 0$, alors z appartient à l'enveloppe convexe des racines de P. Sinon,

$$\overline{Q(z)} = \sum_{i=1}^k \frac{m_i}{|z - z_i|^2}(z - z_i) = 0,$$

donc

$$z = \frac{1}{\sum_{i=1}^{k} \frac{m_i}{|z-z_i|^2}} \sum_{i=1}^{k} \frac{m_i}{|z-z_i|^2} z_i$$

est barycentre de points racines de P. □

Remarque 3.48.
Ce théorème peut aussi se montrer de façons plus élémentaire pour des polynomes réels scindés à racines simples. Soit P un tel polynôme ; notons d son degré. Alors P s'annule exactement d fois sur $\mathbb{R}$; notons $x_1 < \ldots < x_d$ ces racines. Par le théorème de Rolle, P' s'annule au moins une fois sur chaque intervalle (x_i, x_{i+1}), donc par cardinalité, exactement une fois sur chacun de ces intervalles. En particulier, toutes les racines de P' appartiennent à $[x_1, x_d]$.

3.11 Application 6 : Théorème de Carathéodory

Théorème 15 (Carathéodory)**.**
Soit A une partie de $\mathbb{R}^n$ et $x \in Conv(A)$. Alors x est barycentre d'au plus $n+1$ points de A.

Démonstration.
Montrons dans un premier temps que, dans $\mathbb{R}^n$, un barycentre de $n+2$ points $(a_k)_{0 \leq k \leq n+1}$ est barycentre d'au plus $n+1$ de ces points. Soit $x = \sum_{k=0}^{n+1} t_k a_k$ un tel barycentre. Si l'un des t_k est nul, ou si deux des points a_k sont confondus, il n'y a rien à montrer (on élimine le point a_k correspondant). Sinon, la famille $(a_k)_{0 \leq k \leq n+1}$ est affinement liée ; il existe donc une combinaison linéaire non triviale

$$\sum_{k=1}^{n+1} \lambda_k \overrightarrow{a_0 a_k} = \overrightarrow{0}.$$

Posons $\lambda_0 := -\sum_{k=1}^{n} \lambda_k$. Alors, pour tout $s \in \mathbb{R}$,

$$x = \sum_{k=0}^{n+1} (t_k + s\lambda_k) a_k.$$

Or les fonctions $s \mapsto t_k + s\lambda_k$ sont toutes positives pour $n = 0$, et au moins l'un des λ_k est strictement négatif (car la combinaison linéaire est non triviale), donc au moins l'une de ces fonctions tend vers $-\infty$ quand s tend vers $+\infty$. Soit donc $s_* := \min\{s \geq 0 : \ \exists 0 \leq k \leq n, \ t_k + s\lambda_k = 0\}$. Alors tous les coefficients $(t_k + s_* \lambda_k)$ sont positifs, leur somme vaut 1 (car la somme des λ_k vaut 0 par définition de λ_0), et au moins l'un d'entre eux est nul. On a donc écrit x comme combinaison linéaire d'au plus $n+1$ points de la famille $(a_k)_{0 \leq k \leq n+1}$.

Il reste à terminer la démonstration. Montrons par récurrence sur $k \geq n+1$ que tout barycentre de k points de A est barycentre d'au plus $n+1$ points de A (il n'y a rien à montrer si $k \leq n$). Le résultat est vrai pour $k = n+1$. Supposons-le vrai au rang $k \geq n+1$, et soit x un barycentre de $k+1 \geq n+2$ points. Alors, parmis les $n+2$ premiers points, on peut en trouver un qui est barycentre des autres. On peut donc l'éliminer, et écrire x comme barycentre des k points restants. Par hypothèse de récurrence, x est alors barycentre d'au plus $n+1$ points. □

Corollaire 3.49.
Soient $n \geq 0$ et K un compact de $\mathbb{R}^n$. Alors $Conv(K)$ est compacte.

Démonstration.
L'enveloppe convexe de K est l'ensemble de ses barycentres. Par le théorème de Carathéodory, dans $\mathbb{R}^n$, tout barycentre est barycentre d'au plus $n+1$ points, donc l'enveloppe convexe de K est l'ensemble des barycentres d'au plus $n+1$ points de K. C'est même l'ensemble des barycentres d'exactement $n+1$ points de K si l'on autorisa les poids nuls :

$$\mathrm{Conv}(K) = \{\sum_{k=0}^{n} t_k a_k,\ (t_k)_{0\leq k\leq n} \in S_n,\ (a_k)_{0\leq k\leq n} \in K\},$$

où S_n est le simplexe $\{(t_k)_{0\leq k\leq n} \in [0,1]^{n+1} :\ \sum_{k=0}^{n} t_k = 1\}$. Or $S_n \times K$ est compact et l'application de $S_n \times K$ dans $\mathbb{R}^n$ qui a $(t_k)_{0\leq k\leq n},\ a_k)_{0\leq k\leq n}$ associe $\sum_{k=0}^{n} t_k a_k$ est continue, donc son image est compacte. Or cette image est $\mathrm{Conv}(K)$, donc $\mathrm{Conv}(K)$ est compacte. □

3.12 Application 7 : Projection sur un convexe

La dernière application proposée consiste à explorer la notion de projection sur un compact convexe. On reprend la notion de distance à un compact explorée en Sous-section 1.7. Cette partie s'appuie notamment sur [FGN·Ana3, Exercce 2.21].

Proposition 3.50.
Soit K un convexe compact non vide d'un espace vectoriel euclidien E. Alors, pour tout $x \in E$, il existe un unique $y \in K$ tel que $d(\{x\}, K) = d(x, y)$.

Démonstration.
Soit $x \in E$. Par compacité, la fonction distance atteint son minimum sur $\{x\} \times K$, ce qui démontre l'existence d'un tel point x.

Soient y, $y' \in K$ deux points tels que $d(\{x\}, K) = d(x, y) = d(x, y')$. Notons $\theta \in [0, \pi]$ l'angle tel que $\langle y - x, y' - x\rangle = \|y - x\|\,\|y' - x\| \cos(\theta)$, ou autrement dit l'angle entre les demi-droite $[x, y)$ et $[x, y')$. Soit y'' le milieu de $[y, y']$. Alors un peu de trigonométrie dans le plan affine engendré par x, y et y' donne

$$\left\|y'' - x\right\| = d(x, y)\cos(\theta/2).$$

Comme K est convexe, $y'' \in K$. Par minimalité de la distance, $\cos(\theta/2) = 1$, donc $\theta = 0$, donc y et y' sont portés par la même demi-droite issue de x et sont à la même distance de x, donc $y = y'$. □

Définition 3.51 (Projeté d'un point sur un convexe)**.**
*Soient K un convexe compact non vide d'un espace euclidien E et $x \in E$. Le **projeté de** x **sur** K, noté $p_K(x)$, est le point $y \in K$ tel que $d(x, y) = d(\{x\}, K)$.*

Remarquons que $p_K(x) = x$ pour tout $x \in K$.

Lemme 3.52.
Soit $x \in E \setminus K$. Soit H l'hyperplan affine passant par $p_K(x)$ et orthogonal à $p_K(x) - x$. Alors $\{x\}$ et K sont de part et d'autre (au sens large) de H.

Démonstration.
Cet hyperplan a pour équation $\langle p_K(x) - y, p_K(x) - x\rangle = 0$, la variable libre étant y. On sait que $\langle p_K(x) - x, p_K(x) - x\rangle > 0$; il faut donc montrer que $\langle p_K(x) - y, p_K(x) - x\rangle \leq 0$ pour tout $y \in K$.

Soit $y \in K$. L'application $t \mapsto \|x - (1-t)p_K(x) - ty\|^2$ définie sur $[0, 1]$ est minimale en 0, donc sa dérivée en 0 est positive. Or cette dérivée vaut $2\langle p_K(x) - y, x - p_K(x)\rangle$, donc $\langle p_K(x) - y, p_K(x) - x\rangle \leq 0$. □

Remarque 3.53.
Bien qu'on n'en dispose pas d'une interprétation géométrique utile, l'inégalité $\langle p_K(x) - y, p_K(x) - x\rangle \leq 0$ *pour tout* $y \in K$ *reste trivialement vraie si* $x \in K$.

Cette inégalité caractérise $p_K(x)$ *: si* y *est un point de* K *tel que* $\langle y - z, y - x\rangle \leq 0$ *pour tout* $z \in K$*, alors* $y = p_K(x)$*. En effet, on a alors* $\|x - z\|^2 \geq \|x - y\|^2 + \|y - z\|^2$ *pour tout* $z \in K$ *en développant les produit scalaires, donc* y *est bien le point de* K *minimisant la distance à* x*.*

Remarque 3.54.
Si le bord de K *est une hypersurface lisse (par exemple si* K *est un ellipsoïde), alors* $p_K(x) \in \partial K$*. Il y a alors un unique hyperplan* H *passant par* $p_K(x)$ *tel que* K *soit d'un seul côté de* H *: l'hyperplan tangent à* ∂K *en* $p_K(x)$*. Par le lemme précédent, cet hyperplan est orthogonal à* $p_K(x) - x$*. Par conséquent, la droite* $(xp_K(x))$ *est perpendiculaire à* ∂K*.*

Corollaire 3.55.
Soient K *un convexe compact non vide d'un espace euclidien. Alors* p_K *est* 1*-lipschitzienne.*

Démonstration.
Soient x, x' deux points de cet espace euclidien, et y, y' leurs projetés respectifs sur K. Alors $\langle y - y', y - x\rangle \leq 0$ et $\langle y' - y, y' - x'\rangle \leq 0$. En additionnant ces deux inégalités, on trouve $\langle y - y', y - y' - x + x'\rangle \leq 0$, donc $0 \leq \langle y - y', y - y'\rangle \leq \langle y - y', x - x'\rangle$. Par l'inégalité de Cauchy-Schwarz, $\|y - y'\|^2 \leq \|y - y'\| \, \|x - x'\|$, d'où le résultat souhaité après simplification. □

Proposition 3.56.
Soient K *un convexe compact non vide et* $y \in K$*. Alors* $p_K^{-1}(\{y\})$ *est un cône affine de sommet* y*, convexe et fermé.*

Démonstration.
Soit $A := p_K^{-1}(\{y\})$. Montrons que A est un cône affine de sommet y. Soit $x \in A \setminus \{y\}$ et $\lambda \geq 0$; nous voulons montrer que $x' := y + \lambda(x - y) \in A$, donc que $p_K(x') = y$.

Soit $z \in K$. Alors $\langle y - z, y - x\rangle \leq 0$, donc $\langle y - z, y - x'\rangle = \langle y - z, \lambda(y - x)\rangle \leq 0$. Par la Remarque 1.21, $p_K(x') = y$, donc $x' \in A$, ce que l'on voulait démontrer.

L'application p_K est continue car 1-lipschitzienne, donc A est fermé comme préimage d'un fermé par une application continue.

Il reste à montrer que A est convexe. Soient x, $x' \in A$ et $t \in [0, 1]$. Posons $x'' := (1 - t)x + tx'$. Soit $z \in K$. Alors

$$\langle y - z, y - x''\rangle = (1 - t)\langle y - z, y - x\rangle + t\langle y - z, y - x'\rangle \leq 0.$$

Ceci est vrai pour tout $z \in K$, donc encore une fois par la Remarque 1.21, le point x'' appartient à A, ce que l'on voulait démontrer. □

Exercice 3.57.
Dessiner les ensembles $p_K^{-1}(\{y\})$ *quand* K *est un disque, un triangle quelconque, un cube.*

4 Références

[FGN·Ana1] : *Oraux X-ENS. Analyse 1.* S. Francinou, H. Gianella, S. Nicolas.
[FGN·Ana3] : *Oraux X-ENS. Analyse 3.* S. Francinou, H. Gianella, S. Nicolas.
[Gou] : *Les maths en tête, Analyse.* X. Gourdon.
[Moi] : *Mathématiques supérieures, Analyse.* J. Moisan, F. Chanet, F. Delmas, N. Tosel.
[Rom] : *Éléments d'analyse réelle.* J.-E. Rombaldi.
[Ska] : *Analyse.* G. Skandalis.
[RW·L2] : *Mathématiques tout-en-un pour la Licence 2.* J.P. Ramis et A. Warusfel.