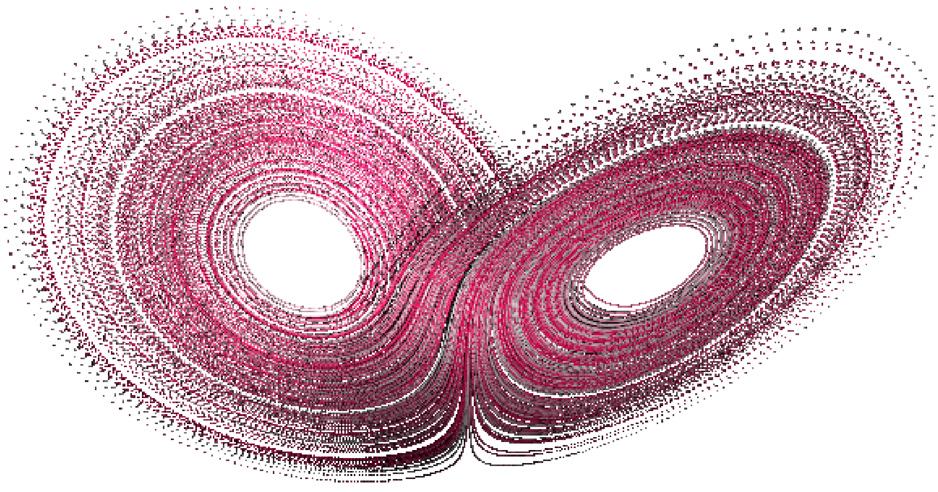
UMONS
Université de Mons

Analyse Mathématique I



Toutes suggestions et corrections peuvent être envoyées à
Christophe.Troestler@umons.ac.be
Je remercie Stéphanie BRIDOUX, Quentin BROUETTE, Matthieu DEMEY, Damien DETRAIN, Julie DE PRIL, Damien GALANT, Marie JULIEN et François STEPHANY pour leur relecture attentive.

Table des matières

Chapitre I

Convergence et nombres réels

Nous supposons ici connus les résultats classiques sur les limites de suites vus dans le cours de Calculus. Tous les critères suffisants de convergence qui y ont été vus imposaient de connaître à priori la limite de la suite dont on voulait prouver la convergence. En effet, soit cette limite est connue à partir d'opérations algébriques sur d'autres limites, soit on cherche à majorer $|x_n - a|$ où a est la limite pressentie de la suite (x_n). Cette section va développer des outils qui permettent de montrer la convergence d'une suite sans avoir aucune idée de la valeur de sa limite. Les retombées de cette exploration dévoileront la nature véritable des nombres réels.

I.1 Suites de Cauchy et complétion

Nous avons déjà vu des critères *nécessaires* de convergence qui ne faisaient pas intervenir la valeur de la limite. Nous avons vu que toute suite convergeante vers un réel doit être bornée. Cependant, être borné est loin d'être une condition suffisante de convergence : la suite $\big((-1)^n\big)_{n\in\mathbb{N}}$ est bornée par $R = 1$ mais elle ne converge pas. Essayons de trouver une condition nécessaire plus fine — plus proche de la notion de convergence. Informellement, $x_n \to a$ dit que les éléments x_n se rapprochent de a. Mais, dans ce cas, ils doivent aussi être proches les uns des autres ! Cette dernière propriété est très importante et mène à la notion suivante.

Définition I.1. Une suite $(x_n)_{n\in I} \subseteq \mathbb{R}$ est dite *de Cauchy* si

$$\forall\varepsilon > 0,\ \exists n_0 \in \mathbb{N},\ \forall m,n \in I,\ \big(m \geqslant n_0 \wedge n \geqslant n_0 \Rightarrow |x_n - x_m| \leqslant \varepsilon\big).$$

La propriété des suites convergentes expliquée ci-dessus s'énonce alors comme suit.

Proposition I.2. *Soit $(x_n)_{n\in I}$ une suite de nombres réels. Si $(x_n)_{n\in I}$ converge vers un nombre réel, alors $(x_n)_{n\in I}$ est de Cauchy.*

Démonstration. Appelons $a \in \mathbb{R}$ la limite de (x_n). La définition de $x_n \to a$ s'écrit

$$\forall \varepsilon_1 > 0,\ \exists n_1 \in \mathbb{N},\ \forall n \geqslant n_1,\ |x_n - a| \leqslant \varepsilon_1. \tag{I.1}$$

Il faut montrer que (x_n) est de Cauchy. Soit $\varepsilon > 0$. En prenant $\varepsilon_1 = \varepsilon/2 > 0$ dans (I.1), on a

$$\exists n_1 \in \mathbb{N},\ \forall n \geqslant n_1,\ |x_n - a| \leqslant \varepsilon/2. \tag{I.2}$$

Prenons $n_0 := n_1$. Soit m et n deux naturels $\geqslant n_0$. De (I.2) on tire que $|x_m - a| \leqslant \varepsilon/2$ et $|x_n - a| \leqslant \varepsilon/2$. Dès lors, en utilisant l'inégalité triangulaire, on a

$$\begin{aligned} |x_m - x_n| = |x_m - a + a - x_n| &\leqslant |x_m - a| + |x_n - a| \\ &\leqslant \varepsilon/2 + \varepsilon/2 = \varepsilon. \end{aligned}$$

□

Intuitivement, on a envie de dire que l'inverse est vrai : si les éléments d'une suite se rapprochent les uns des autres, ils doivent forcément aussi se rapprocher d'un certain nombre a. Il faut cependant se méfier et essayer de valider son intuition. Pour vous persuader que ce n'est peut-être pas aussi évident qu'il n'y parait, montrons que ce n'est pas vrai si on travaille dans $\mathbb{Q}$. On pourrait en effet se demander pourquoi on a choisi de travailler sur $\mathbb{R}$. La définition de $x_n \to a$ et toutes les propriétés vues jusquà présent restent valables si on se restreint aux suites $(x_n) \subseteq \mathbb{Q}$ et dont la limite $a \in \mathbb{Q}$. Ce qui n'est pas vrai est que si une suite $(x_n)_{n\in I} \subseteq \mathbb{Q}$ est de Cauchy, alors elle converge vers un élément $a \in \mathbb{Q}$. Pour donner un exemple de telle suite, considérons la méthode de Newton pour calculer $\sqrt{2}$. Particularisée à ce cas, elle se réduit à définir la suite $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{Q}$ par

$$x_0 = 2 \qquad \text{et} \qquad x_{n+1} = \frac{x_n}{2} + \frac{1}{x_n}, \quad n \in \mathbb{N}.$$

La suite (x_n) est strictement décroissante [1] et minorée par 1 :

$$\forall n \in \mathbb{N}, \quad x_{n+1} < x_n \tag{I.3}$$

$$\forall n \in \mathbb{N}, \quad x_n \geqslant 1. \tag{I.4}$$

1. Veuillez aussi regarder les graphiques dans l'introduction du syllabus du cours de Calculus I.

Prouvons maintenant rigoureusement ces affirmations. En remplaçant x_{n+1} par sa définition en fonction de x_n, on voit que (I.3) est équivalent à

$$\forall n \in \mathbb{N}, \quad 2 < x_n^2. \tag{I.5}$$

D'autre part, il est facile de prouver par récurrence (faites le !) que $x_n > 0$ pour tout n. Dès lors, (I.5) implique que $x_n^2 > 2 > 1$ et donc que $x_n > 1$. Il reste donc à établir (I.5). Faisons le par récurrence. Pour $n = 0$, l'inégalité devient $2 < 4$ ce qui est vrai. Supposons que $2 < x_n^2$ et montrons que $2 < x_{n+1}^2$. En remplaçant x_{n+1} par $x_n/2 + 1/x_n$, on trouve que $x_{n+1}^2 > 2$ est équivalent à (faites les calculs !) $(x_n^2 - 2)^2 > 0$ ce qui est vrai puisque, par hypothèse de récurrence, $x_n^2 \neq 2$.

Nous verrons à la section suivante (théorème I.7) que les propriétés (I.3) et (I.4) impliquent que la suite (x_n) soit de Cauchy. Supposons que celle-ci converge vers un élément a. Bien sûr on a alors aussi que $x_{n+1} \to a$ (pouvez-vous le montrer ?). De plus, par les règles de calcul (limites de sommes et quotients), on a

$$a = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} \left(\frac{x_n}{2} + \frac{1}{x_n} \right) = \frac{a}{2} + \frac{1}{a}. \tag{I.6}$$

On pouvait appliquer la règle concernant le quotient $1/x_n$ car de (I.4) on déduit que $a \geqslant 1$ et donc $a \neq 0$. On peut réécrire (I.6) sous la forme :

$$a^2 = 2.$$

Or il n'y a aucune solution $a \in \mathbb{Q}$ à cette équation.[2] Cela montre bien que la suite $(x_n) \subseteq \mathbb{Q}$ ne peut converger vers un élément de $\mathbb{Q}$.

L'avantage de l'exemple précédent est qu'il n'utilise que $\mathbb{Q}$. Si on accepte qu'on connait $\mathbb{R}$, on peut donner un exemple qui est peut-être plus facile à comprendre. Considérons $a = \sqrt{2}$ dont on vient de voir qu'il n'appartient pas à $\mathbb{Q}$. On peut regarder son écriture décimale :

$$a = \sqrt{2} = 1{,}414213562\ldots = 1{,}a_{-1}a_{-2}a_{-3}\ldots$$

où $a_{-i} \in \{0,\ldots,9\}$ est la i^{e} décimale de $\sqrt{2}$. La suite $(x_n)_{n\in\mathbb{N}}$ définie par

$$x_n = 1{,}a_{-1}\ldots a_{-n} = \frac{1a_{-1}\ldots a_{-n}}{10^n} \in \mathbb{Q}$$

2. Soit $p/q \in \mathbb{Q}$ une solution, c'est-à-dire que $p^2 = 2q^2$ avec $p,q \in \mathbb{Z}$. En simplifiant par 2 autant de fois que nécessaire la fraction p/q, on peut supposer que p ou q est impair. Mais $p^2 = 2q^2$ implique que p^2 et donc p est pair (faites les détails). Donc, $p = 2r$ pour un $r \in \mathbb{Z}$ et $q^2 = 2r^2$. Mais alors q est aussi pair ce qui contredit le fait qu'un des deux devait être impair.

($1a_{-1}\dots a_{-n}$ représente le naturel dont les digits sont $1, a_{-1}, \dots, a_{-n}$) converge bien vers a puisque
$$|x_n - a| \leqslant \frac{1}{10^n} \xrightarrow[n\to\infty]{} 0$$
(pouvez-vous justifier l'inégalité ?). Ainsi, la suite de rationnels (x_n) — qui est bien de Cauchy en vertu de la proposition I.2 — tend vers $a = \sqrt{2} \notin \mathbb{Q}$.

Ces deux exemples montrent que les suites de Cauchy de $\mathbb{Q}$ ne convergent pas nécessairement dans $\mathbb{Q}$ — en fait elles convergent dans $\mathbb{R}$. Les espaces dont les suites de Cauchy possèdent une limite dans ce même espace sont fondamentaux en Analyse. Ils sont dit « complets ».

Définition I.3. Un espace X est dit *complet* si toute suite de Cauchy dans X converge vers un élément de X.

D'après ce qui vient d'être dit, $\mathbb{Q}$ n'est pas complet. Par contre l'espace $\mathbb{R}$ l'est et c'est sa caractéristique essentielle par rapport à $\mathbb{Q}$. Plus précisément :

Axiome I.4. $\mathbb{R}$ est le plus petit espace complet qui contient $\mathbb{Q}$. On dit que $\mathbb{R}$ est le *complété* de $\mathbb{Q}$.

À ce stade, il n'est pas clair qu'un tel « complété de $\mathbb{Q}$ » existe ni qu'il puisse être muni d'une structure de corps. Pour ceux qui sont intéressés, une construction de $\mathbb{R}$ et la preuve de diverses propriétés sera donnée à la section I.5. Pour les autres, vous pouvez penser que $\mathbb{R}$ est essentiellement $\mathbb{Q}$ auquel on a rajouté des éléments pour « boucher les trous » afin que toutes les suites de Cauchy convergent. Par ailleurs, l'usage du « le » implique l'unicité, ce qu'on n'a pas ici en tant qu'ensemble. Néanmoins, il est utilisé parce que si $\mathbb{R}_1$ et $\mathbb{R}_2$ sont deux ensembles qui vérifient la propriété de l'axiome I.4, alors $\mathbb{R}_1$ et $\mathbb{R}_2$ sont *isomorphes en tant que corps*, ce qui veut intuitivement dire que *la seule différence* entre les deux ensembles peut-être au niveau *des « noms » des éléments* mais que, sinon, ils ont exactement la même structure. Nous ne nous attarderons pas sur ce point ici.

I.2 Supremum, infimum et suites monotones

Nous allons maintenant tirer diverses conséquences du fait que $\mathbb{R}$ est complet. Commençons par définir clairement quelques notions.

Définition I.5. Soit $(x_n)_{n\in I}$ une suite de nombres réels. On dit que $(x_n)_{n\in I}$ est

- *croissante* si $\forall n \in I,\ x_{n+1} \geqslant x_n$;
- *décroissante* si $\forall n \in I,\ x_{n+1} \leqslant x_n$;
- *monotone* si elle est croissante ou décroisssante ;
- *strictement croissante* si $\forall n \in I,\ x_{n+1} > x_n$;
- *strictement décroissante* si $\forall n \in I,\ x_{n+1} < x_n$;
- *strictement monotone* si elle est strictement croissante ou strictement décroissante.

Revoyez aussi les définitions de suites majorée et minorée. Ces notions s'étendent de manière naturelle aux ensembles. Faisons le explicitement pour éviter toute confusion.

Définition I.6. Un ensemble $A \subseteq \mathbb{R}$ est dit *majoré* (resp. *minoré*) s'il existe un $R \in \mathbb{R}$ tel que, pour tout $a \in A$, $a \leqslant R$ (resp. $a \geqslant R$). Un tel R est appelé un *majorant* (resp. un *minorant*) de A.

Voici quelques exemples. La suite $(1/n)_{n\geqslant 1}$ (figure I.1) est strictement décroissante (donc aussi décroissante) puisque $1/(n+1) < 1/n$ pour tout $n \geqslant 1$. Les suites constantes sont les seules suites à la fois croissantes et décroissantes. Elles ne sont pas strictement monotones. La suite $(x_n)_{n\in\mathbb{N}}$ définie par $x_n = \lceil n/3 \rceil$ (voir figure I.2) est croissante mais pas strictement croissante. Elle n'est pas non plus majorée car, quel que soit $R \in \mathbb{R}$, $x_n > R$ pour $n = \max\{3\lceil R\rceil + 3, 0\} \in \mathbb{N}$ (pouvez-vous faire les détails ?). La suite $\big((-1)^n\big)_{n\in\mathbb{N}}$ n'est pas monotone (à fortiori pas strictement monotone) mais elle est majorée et minorée (donc bornée). La suite $(x_n)_{n\in\mathbb{N}}$ définie par $x_n = n/(25+n^2)$ (voir figure I.3) n'est pas monotone mais sa sous-suite $(x_n)_{n\geqslant 6}$ est strictement décroissante. Elle est minorée par zéro et majorée par $1/10$.

Voici une première conséquence de la complétude de $\mathbb{R}$ sur les suites monotones.

Théorème I.7. *Soit $(x_n)_{n\in I}$ une suite de nombres réels. Si $(x_n)_{n\in I}$ est croissante et majorée (resp. décroissante et minorée), elle est de Cauchy. En particulier, puisque $\mathbb{R}$ est complet, il existe un $a \in \mathbb{R}$ tel que $x_n \to a$.*

Démonstration. Nous allons seulement faire la preuve dans le cas où (x_n) est croissante et majorée — l'autre cas étant similaire. Supposons par l'absurde que

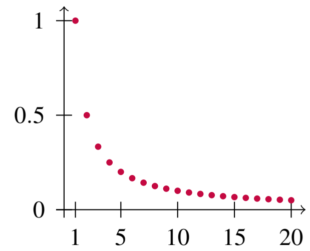


FIGURE I.1 – $(1/n)$

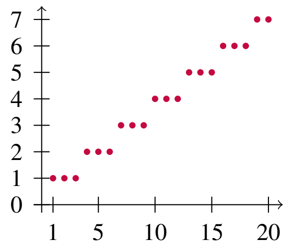


FIGURE I.2 – $(\lceil n/3 \rceil)$

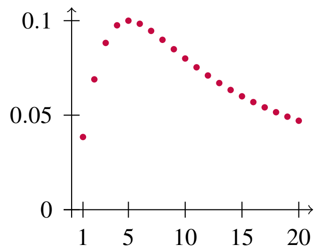


FIGURE I.3 – $\left(\dfrac{n}{25+n^2}\right)$

(x_n) ne soit pas de Cauchy, c'est-à-dire que

$$\exists \varepsilon > 0,\ \forall n_0 \in \mathbb{N},\ \exists m,n \geqslant n_0, \quad |x_n - x_m| > \varepsilon.$$

Quitte à échanger m et n, on peut supposer que $m \geqslant n$ (pouvez vous justifier cette affirmation en détail *en tenant compte* des quantificateurs qui précèdent ?). Puisque la suite est croissante, cela implique que $x_m \geqslant x_n$. Le fait que (x_n) ne soit pas de Cauchy se réécrit donc comme

$$\exists \varepsilon > 0,\ \forall n_0 \in \mathbb{N},\ \exists m \geqslant n \geqslant n_0, \quad x_m \geqslant x_n + \varepsilon. \tag{I.7}$$

On sait que $I = \{n \in \mathbb{N} : n \geqslant n_I\}$ pour un certain $n_I \in \mathbb{N}$. En prenant $n_0 := n_I$ dans (I.7), on obtient l'existence de

$$\mu_0, \nu_0 \in I \text{ tels que} \qquad \mu_0 \geqslant \nu_0 \quad \text{et} \quad x_{\mu_0} \geqslant x_{\nu_0} + \varepsilon.$$

En réutilisant (I.7) avec $n_0 := \mu_0 + 1$, on déduit qu'il existe

$$\mu_1, \nu_1 \in I \text{ tels que} \qquad \mu_1 \geqslant \nu_1 > \mu_0 \quad \text{et} \quad x_{\mu_1} \geqslant x_{\nu_1} + \varepsilon.$$

On continue de la même manière : à partir de

$$\mu_k, \nu_k \in I \text{ tels que} \qquad \mu_k \geqslant \nu_k > \mu_{k-1} \quad \text{et} \quad x_{\mu_k} \geqslant x_{\nu_k} + \varepsilon \tag{I.8}$$

on déduit l'existence de μ_{k+1}, ν_{k+1} vérifiant les mêmes propriétés simplement en prenant $n_0 := \mu_k + 1$ dans (I.7).

Considérons la suite $(x_{\nu_k})_{k\in\mathbb{N}}$ (voit figure I.4). La propriété (I.8) utilisée pour k et $k+1$ implique que $\nu_{k+1} > \mu_k \geqslant \nu_k$ et $x_{\mu_k} \geqslant x_{\nu_k} + \varepsilon$. En se rappelant que (x_n) est croissante, on déduit $x_{\nu_{k+1}} \geqslant x_{\mu_k}$. En rassemblant les éléments précédents qui sont valables quel que soit k, on obtient que

$$\forall k \in \mathbb{N}, \quad x_{\nu_{k+1}} \geqslant x_{\nu_k} + \varepsilon.$$

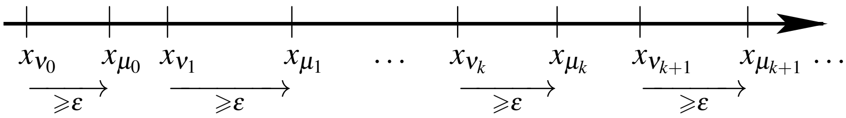


FIGURE I.4 – Sous-suite d'une suite (x_n) non de Cauchy

Une simple preuve par récurrence (faites la !) permet alors de conclure que

$$\forall k \in \mathbb{N}, \quad x_{\nu_k} \geqslant x_{\nu_0} + k\varepsilon. \tag{I.9}$$

Rappelons maintenant que $(x_n)_{n \in I}$ est majorée, c'est-à-dire qu'il existe un $R \in \mathbb{R}$ tel que $x_n \leqslant R$ pour tout n. En particulier,

$$\forall k \in \mathbb{N}, \quad x_{\nu_k} \leqslant R. \tag{I.10}$$

De (I.9) et (I.10), il vient que $\forall k \in \mathbb{N},\ R \geqslant x_{\nu_0} + k\varepsilon$, c'est-à-dire que

$$\forall k \in \mathbb{N}, \quad k \leqslant \frac{R - x_{\nu_0}}{\varepsilon}.$$

ce qui est faux car il suffit de prendre $k = \lceil (R - x_{\nu_0})/\varepsilon \rceil + 1$. □

Grâce à ce théorème nous allons pouvoir généraliser l'idée de maximum et de minimum. Rappelons d'abord ces notions.

Définition I.8. Soit $A \subseteq \mathbb{R}$ et $a \in \mathbb{R}$. On dit que a est le *maximum* (resp. *minimum*) de A si $a \in A$ et si $\forall a' \in A,\ a' \leqslant a$ (resp. $\forall a' \in A,\ a' \geqslant a$).

Notez que cette définition dit « *le* maximum ». En vérité, pour pouvoir employer le « le », il faudrait montrer l'unicité du maximum. Ceci est laissé au lecteur. On note le maximum (resp. minimum) de A par $\max A$ (resp. $\min A$) quand celui-ci existe.

Insistons sur le fait que le maximum de A doit être un majorant de A ($\forall a' \in A,\ a' \leqslant a$) mais doit aussi appartenir à A. Par exemple, le maximum de $]0,1]$ est 1 car $1 \in]0,1]$ et $\forall x \in]0,1],\ x \leqslant 1$. Par contre, $]0,1[$ n'a pas de maximum. En effet, si $a \in]0,1[$ en était un, en prenant $a' = (1+a)/2 \in]0,1[$ on aurait $a' > a$ et on contredirait ainsi la maximalité de a (voir figure I.5). On a peut-être envie de dire que 1 est le maximum de $]0,1[$ mais ce n'est pas le cas car $1 \notin]0,1[$. 1 est simplement un majorant de $]0,1[$. Il y en a cependant beaucoup d'autres : $2, 3, 4, 5, 1.5, \pi, \dots$ sont aussi des majorants. Ce qui distingue 1 des autres c'est

FIGURE I.5 – Non existence du maximum

que c'est le « meilleur » au sens où c'est le plus petit de tous les majorants ou que c'est le seul qui « colle » à l'ensemble.

Ce que nous venons de décrire sur cet exemple est l'idée de supremum d'un ensemble. Formalisons cette notion en toute généralité.

Définition I.9. Soit $A \subseteq \mathbb{R}$ et $a \in \mathbb{R}$. On dit que a est le *supremum* (resp. l'*infimum*) de A si a est le plus petit des majorants (resp. le plus grand des minorants).

Écrivons à l'aide de formules quantifiées le fait que a soit le supremum de A :

- a majore A : $\forall a' \in A,\ a' \leqslant a$;
- a est le plus petit des majorants : $\forall b \in \mathbb{R},\ (\forall a' \in A,\ a' \leqslant b) \Rightarrow b \geqslant a$.

On laisse au lecteur le soin de faire de même avec la notion d'infimum.

De nouveau, nous avons utilisé l'article « le » qui exprime l'unicité. Donnons une preuve de ce fait. Soient a_1 et a_2 deux supremums de A. Puisque a_2 majore A et que a_1 est le plus petit des majorants, on a $a_1 \leqslant a_2$. De manière analogue, $a_2 \leqslant a_1$. Donc $a_1 = a_2$ et l'unicité du supremum est prouvée. La preuve est similaire pour l'infimum. On note le supremum (resp. l'infimum) d'un ensemble A par

$$\sup A \qquad (\text{resp. } \inf A).$$

Il est facile de voir que $\sup A = -\inf(-A)$ où $-A$ désigne l'ensemble $\{-a : a \in A\}$. C'est pourquoi dans la suite nous ferons les démonstrations uniquement pour le supremum, les faits concernant l'infimum s'en déduisant grâce à cette relation. (Ce peut être néanmoins un bon exercice que d'adapter les preuves données au cas de l'infimum.)

L'avantage du supremum (resp. de l'infimum) vis à vis du maximum (resp. du minimum) est que celui-ci existe toujours. C'est l'objet du théorème I.10 et des définitions I.13 et I.15 ci-dessous.

Théorème I.10. *Soit $A \subseteq \mathbb{R}$. Si $A \neq \varnothing$ est majoré (resp. minoré), alors le supremum (resp. l'infinum) de A existe.*

Nous ne ferons la démonstration que pour le supremum — celle pour l'infimum étant similaire.

Afin de faciliter la preuve de ce théorème, introduisons la notion de « maximum approximatif ». Si $a \in \mathbb{R}$ est le maximum de $A \subseteq \mathbb{R}$, c'est que $a \in A$ et $A \subseteq \,]-\infty, a]$. L'idée de « maximum approximatif » est qu'on va garder la propriété $a \in A$ mais on va seulement demander que A soit « approximativement recouvert » par $]-\infty, a]$. Plus précidément, on a

Définition I.11. Soit $A \subseteq \mathbb{R}$, $a \in \mathbb{R}$ et $\varepsilon > 0$. On dit que a est un ε-*maximum* de A si $a \in A$ et $A \subseteq \,]-\infty, a+\varepsilon[$.

Autrement dit, $a \in A$ et $a+\varepsilon$ est un majorant de A. Cela est illustré à la figure I.6. Contrairement aux maximums, les ε-maximums ne sont pas uniques.

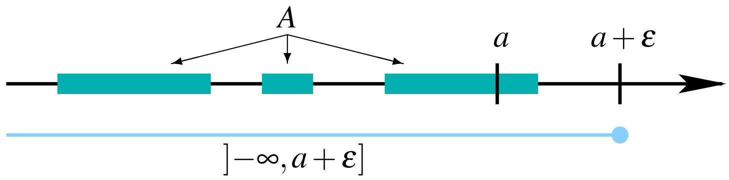


FIGURE I.6 – ε-maximum

Par exemple, pour $]0,1[$ et $\varepsilon < 1$, $1-\varepsilon/2$, $1-\varepsilon/3$, $1-\varepsilon/4$,... sont tous des ε-maximums. En fait, si a est un ε-maximum et $a' \in A$ est plus grand que a, alors a' est aussi un ε-maximum. L'avantage des ε-maximums par rapport aux maximums est qu'il leur faut très peu d'hypothèses pour exister.

Proposition I.12. *Soit $A \neq \varnothing$ un sous-ensemble de $\mathbb{R}$. Si A est majoré alors, quel que soit $\varepsilon > 0$, A possède (au moins) un ε-maximum.*

Démonstration. Soit $\varepsilon > 0$. Procédons par l'absurde et supposons que A ne possède aucun ε-maximum. En niant la définition « $\exists a \in A,\ a+\varepsilon$ soit un majorant de A », on obtient :

$$\forall a \in A,\ \exists a' \in A,\ a' > a+\varepsilon. \tag{I.11}$$

Puisque A est non vide, on peut prendre (au hasard) un élément $a_0 \in A$. En employant (I.11) avec $a = a_0$, on déduit l'existence d'un $a' \in A$, que nous allons noter a_1, tel que $a_1 > a_0 + \varepsilon$. On peut de nouveau appliquer (I.11) avec $a = a_1$ pour avoir l'existence d'un $a_2 \in A$ tel que $a_2 > a_1 + \varepsilon$. En continuant de la sorte, on construit une suite

$$(a_n)_{n\in\mathbb{N}} \subseteq A \quad \text{telle que} \quad \forall n \in \mathbb{N},\ a_{n+1} > a_n + \varepsilon.$$

(Pouvez-vous expliciter la construction par récurrence qui se cache derrière ce qu'on vient de dire ?) De cette propriété, on tire aisément par récurrence que

$$\forall n \in \mathbb{N},\ a_n \geqslant a_0 + n\varepsilon. \tag{I.12}$$

Rappelons que A est majoré, c'est-à-dire qu'il existe un $R \in \mathbb{R}$ tel que $\forall a \in A,\ a \leqslant R$. De (I.12), on déduit alors que

$$\forall n \in \mathbb{N},\ n \leqslant \frac{R - a_0}{\varepsilon}$$

ce qui est impossible — il suffit de prendre $n = \lceil (R - a_0)/\varepsilon \rceil + 1 \in \mathbb{N}$ pour le voir. □

Revenons maintenant à la démonstration de la preuve de l'existence du supremum.

Démonstration du théorème I.10. Nous allons construire une suite croissante d'ε-maximums qui vont converger vers le supremum. Commençons avec $\varepsilon = 1$. En vertu de la proposition I.12, il existe un

$$a_1 \in A \quad \text{tel que} \quad \forall a \in A,\ a \leqslant a_1 + 1.$$

Si on utilise de nouveau la proposition I.12 mais cette fois avec $\varepsilon = 1/2$, on obtient un $a_2' \in A$ tel que a_2' est un $1/2$-maximum de A. Posons $a_2 := \max\{a_1, a_2'\}$. Clairement $a_2 \in A$ — puisque a_1 et a_2' appartiennent à A. De plus, comme $a_2 \geqslant a_2'$, a_2 est aussi un $1/2$-maximum de A (vérifiez-le !). Par construction $a_2 \geqslant a_1$. En répétant l'argument, on construit un $a_3 \in A$ tel que $a_3 \geqslant a_2$ et a_3 est un $1/3$-maximum de A. En continuant de la sorte, on obtient une suite

$$(a_n)_{n \geqslant 1} \subseteq A \text{ telle que } \forall n \geqslant 1,\ a_{n+1} \geqslant a_n \text{ et}$$
$$a_n \text{ est un } 1/n\text{-maximum de } A$$

(voir aussi la figure I.7). Comme la suite (a_n) est croissante et qu'elle est majorée par une borne supérieure de A, le théorème I.7 dit qu'elle est de Cauchy et donc qu'il existe un $a^* \in \mathbb{R}$ tel que $a_n \to a^*$. Nous allons montrer que A vérifie la définition du supremum de A.

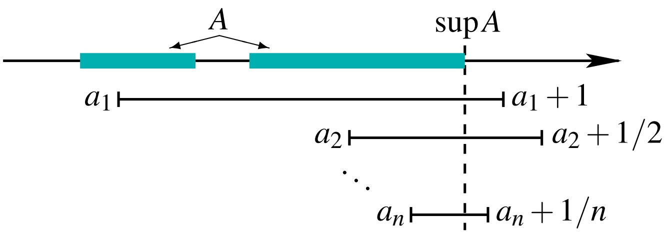


FIGURE I.7 – Suite croissante (a_n) convergeant vers le supremum

- *Pour tout $a \in A$, $a \leqslant a^*$.* Soit $a \in A$. Puisque a_n est un $1/n$-maximum, on a $a \leqslant a_n + 1/n$ pour tout n. Vu que $a_n \to a^*$ et $1/n \to 0$, les règles de calculs sur les limites et le résultat sur la limite d'inégalités impliquent que

$$a \leqslant \lim_{n\to\infty} \left(a_n + \frac{1}{n}\right) = a^*.$$

- *Si b est un majorant de A, alors $a^* \leqslant b$.* Puisque $a_n \in A$ et que b est un majorant, il vient $a_n \leqslant b$. En passant à la limite, on en déduit $a^* = \lim_{n\to\infty} a_n \leqslant b$. □

L'hypothèse que A soit majoré est assez naturelle. Si ce n'est pas le cas, il n'y a aucune chance pour que le supremum tel que défini page 8 existe. En effet, supposons un instant que $\sup A$ existe et montrons que, quel que soit $x \in \mathbb{R}$, $\sup A \geqslant x$. Soit $x \in \mathbb{R}$. Puisque A n'est pas majoré, il ne peut certainement pas être borné par x, ce qui signifie qu'il existe un $a \in A$ tel que $x < a$. Mais puisque, par définition, le supremum est un majorant, on a $x < a \leqslant \sup A$. En résumé, le supremum d'un ensemble non majoré doit être plus grand que tous les nombres réels (et par conséquent ne peut pas être un nombre réel). Cela motive la définition suivante.

Définition I.13. Si $A \subseteq \mathbb{R}$ est un ensemble qui n'est pas majoré (resp. pas minoré), on pose $\sup A = +\infty$ (resp. $\inf A = -\infty$).

La seconde hypothèse sur A dans le théorème I.10 est que A soit non-vide. À l'instar de ce qu'on a fait ci-dessus pour A non borné, est-il possible de donner une valeur raisonnable à $\sup \varnothing$? Pour le découvrir, intéressons nous au comportement du supremum vis à vis des opérations ensemblistes.

Proposition I.14. *Soit A et B deux sous-ensembles de $\mathbb{R}$. Les relations suivantes sont vraies :*

$$\begin{array}{cc} A \subseteq B \Rightarrow \sup A \leqslant \sup B & A \subseteq B \Rightarrow \inf A \geqslant \inf B \\ \sup(A \cup B) = \max\{\sup A, \sup B\} & \inf(A \cup B) = \min\{\inf A, \inf B\} \\ \sup(A \cap B) \leqslant \min\{\sup A, \sup B\} & \inf(A \cap B) \geqslant \max\{\inf A, \inf B\} \end{array}$$

Démonstration. Comme d'habitude, nous ne donnerons des preuves que des relations concernant le supremum — les autres étant similaires.

- Commençons par montrer $A \subseteq B \Rightarrow \sup A \leqslant \sup B$. Si $\sup B = +\infty$, c'est évident. On peut donc supposer $\sup B < +\infty$. Puisque $A \subseteq B$, tout $a \in A$ appartient aussi à B et, vu que $\sup B$ est un majorant de B, $a \leqslant \sup B$. Donc $\sup B$ est un majorant de A. Comme le supremum est le plus petit des majorants, on a $\sup A \leqslant \sup B$.

- $\sup(A \cup B) = \max\{\sup A, \sup B\}$. Vu que A et B sont inclus à $A \cup B$, $\sup A \leqslant \sup(A \cup B)$ et $\sup B \leqslant \sup(A \cup B)$ d'où

$$\max\{\sup A, \sup B\} \leqslant \sup(A \cup B).$$

D'autre part, si $x \in A \cup B$, on a que x appartient à A, auquel cas $x \leqslant \sup A \leqslant \max\{\sup A, \sup B\}$, ou x appartient à B auquel cas $x \leqslant \sup B \leqslant \max\{\sup A, \sup B\}$. Autrement dit, $\max\{\sup A, \sup B\}$ est un majorant de $A \cup B$ et la définition du supremum implique que $\sup(A \cup B) \leqslant \max\{\sup A, \sup B\}$.

- $\sup(A \cap B) \leqslant \min\{\sup A, \sup B\}$. Cela découle du premier point car $A \cap B$ est inclus à A et à B. □

Remarquons que l'inégalité de la propriété d'intersection n'est en général pas une égalité. Par exemple, si $A = \{1, 2\}$ et $B = \{1, 3\}$, on a $\sup(A \cap B) = \sup\{1\} = 1 < \min\{\sup A, \sup B\} = \min\{2, 3\} = 2$.

Revenons maintenant à la question de la valeur à attribuer à $\sup \varnothing$. Nous voudrions que la proposition précédente reste valable — c'est pourquoi nous n'avons pas mis des restrictions du type $A \neq \varnothing$ dans son énoncé. Soit $r \in \mathbb{R}$. Comme $\varnothing \subseteq \{r\}$, nous voudrions que $\sup \varnothing \leqslant \sup\{r\} = r$. Étant donné que r est quelconque, cela signifie que $\sup \varnothing$ doit être plus petit que n'importe quel réel — et ne peut donc être un réel. Au vu de ceci, il est naturel d'adopter la définition suivante.

Définition I.15. On pose $\sup \varnothing = -\infty$ et $\inf \varnothing = +\infty$.

Le lecteur est invité à vérifier que, grâce à cette définition, les autres propriétés de la propostion I.14 sont vraies même si A ou B est vide.

À ce moment, il est bon de répéter que le travail qu'on vient de faire nous permet d'attribuer une valeur dans $[-\infty, +\infty]$ à $\sup A$ et $\inf A$ pour un ensemble arbitraire $A \subseteq \mathbb{R}$. Nous avons vu que la complétude de $\mathbb{R}$ était une condition suffisante pour l'existence de $\sup A$ et $\inf A$. En fait, elle est essentielle. Par exemple, dans $\mathbb{Q}$, le supremum de $\{x \in \mathbb{Q} : x^2 \leqslant 2\}$ n'existe pas.

Dans le discours qui précédait la définition de supremum, nous avions noté que le supremum était le meilleur des majorants car il était le plus petit ou le seul qui « collait » à l'ensemble. Explicitons cette seconde caractérisation.

Proposition I.16. *Soit $A \subseteq \mathbb{R}$ et $a \in \mathbb{R}$. Le réel a est le supremum (resp. l'infimum) de A si et seulement si a est un majorant (resp. minorant) de A et l'une des trois propriétés (équivalentes) suivantes est vérifiée :*

(i) *il existe une suite $(x_n)_{n \in I} \subseteq A$ telle que $x_n \to a$;*

(ii) *il existe une suite croissante (resp. décroissante) $(x_n)_{n \in I} \subseteq A$ telle que $x_n \to a$;*

(iii) $\forall \varepsilon > 0,\ \exists a' \in A,\ a - \varepsilon \leqslant a'$ *(resp. $a' \leqslant a + \varepsilon$).*

Démonstration. Il est clair que (ii) $\Rightarrow$ (i). On a aussi (i) $\Rightarrow$ (iii). En effet, étant donné un $\varepsilon > 0$, la définition de $x_n \to a$ implique qu'il existe un $x_n \in A$ tel que $|x_n - a| \leqslant \varepsilon$ et donc tel que $a - \varepsilon \leqslant x_n$.

Condition nécessaire. Il suffit de prouver que si $a = \sup A$, alors a satisfait (ii). La preuve du théorème I.10 (page 10) construit en effet une suite croissante de A convergeant vers le supremum.

Condition suffisante. Il suffit de montrer que si a est un majorant satisfaisant (iii), alors $a = \sup A$. Comme a est un majorant, il reste à prouver que c'est le plus petit d'entre eux. Soit b un majorant de A. Soit $n \in \mathbb{N} \setminus \{0\}$. En appliquant (iii) avec $\varepsilon = 1/n$, on a l'existence d'un $a' \in A$ tel que $a - 1/n \leqslant a'$. Puisque b est un majorant, on en déduit que $a - 1/n \leqslant b$. Vu que n est arbitraire, on peut passer à la limite $n \to \infty$ ce qui donne $a = \lim(a - 1/n) \leqslant b$ comme désiré. $\square$

Remarque I.17. Il faut faire attention au fait que l'équivalence de (i), (ii) et (iii) a lieu sous l'hypothèse que a est un majorant de A.

Il est facile de démontrer directement (i) $\Rightarrow$ (ii). En effet, étant donné $(x_n)_{n\in I}$, il suffit de considérer la suite $\big(\max\{x_n : n \in I,\ n \leqslant k\}\big)_{k\in I}$.

Strictement parlant, on n'a pas démontré que (iii) $\Rightarrow$ (ii). Pour le faire, il suffit de reprendre les idées qui permettent la construction de la suite (a_n) dans la démonstration du théorème I.10 (page I) mais d'employer (iii) au lieu de l'existence des ε-maximums. Les détails sont laissés au lecteur (c'est un bon exercice !).

Comme nous avons précisé $a \in \mathbb{R}$ dans l'énoncé de la proposition précédente, celle-ci ne s'applique pas au cas où le supremum prend une valeur infinie. Évidemment, si $A = \varnothing$, il n'y a aucune chance de trouver des suites dans A ! Lorsque A est non majoré cependant, on peut trouver un analogue à la proposition I.16. Le voici.

Proposition I.18. *Soit $A \subseteq \mathbb{R}$. Le supremum (resp. infimum) de A vaut $+\infty$ (resp. $-\infty$) si et seulement si une des propriétés (équivalentes) suivantes est satisfaite :*

(i) il existe une suite $(x_n)_{n\in I} \subseteq A$ telle que $x_n \to +\infty$ (resp. $x_n \to -\infty$) ;

(ii) il existe une suite croissante (resp. décroissante) $(x_n)_{n\in I} \subseteq A$ telle que $x_n \to +\infty$ (resp. $x_n \to -\infty$) ;

(iii) $\forall \rho \in \mathbb{R},\ \exists a' \in A,\ a' \geqslant \rho$ (resp. $a' \leqslant \rho$).

Démonstration. Comme d'habitude nous ne ferons la démonstration que pour le supremum. Puisque (iii) ne veut rien dire d'autre que A est non majoré, on a par définition que $\sup A = +\infty \Leftrightarrow$ (iii). De plus il est clair que (ii) $\Rightarrow$ (i). Il reste donc à prouver que (i) $\Rightarrow$ (iii) $\Rightarrow$ (ii).

(i) $\Rightarrow$ (iii). Soit $\rho > 0$. Comme $x_n \to +\infty$, il existe au moins un n tel que $x_n \geqslant \rho$ (pouvez-vous faire les détails ?). De plus $x_n \in A$ puisque la suite est incluse à A. Il suffit donc de prendre $a' := x_n$.

(iii) $\Rightarrow$ (ii). Construisons d'abord une suite $(x'_n)_{n\in\mathbb{N}} \subseteq A$. En utilisant (iii) avec $\rho = n$, on obtient l'existence d'un $x'_n \in A$ tel que $x'_n \geqslant n$. Posons $x_n := \max\{x'_m : m \leqslant n\}$. Clairement, $(x_n)_{n\in\mathbb{N}}$ est une suite croissante. De plus $x_n \to +\infty$. Il suffit en effet d'utiliser la convergence dominée et de remarquer que, pour tout $n \in \mathbb{N}$, $x_n \geqslant x'_n \geqslant n \xrightarrow[n\to\infty]{} +\infty$. □

Les deux propositions précédentes montrent que le supremum d'un ensemble non-vide est la limite d'une suite croissante d'éléments de cet ensemble. Inversément, étant donné une suite croissante, le supremum peut caractériser sa limite.

Proposition I.19. *Soit $(x_n)_{n\in I} \subseteq \mathbb{R}$. Si $(x_n)_{n\in I}$ est croissante (resp. décroissante), alors, au sens large, $x_n \xrightarrow[n\to\infty]{} \sup\{x_n : n \in I\}$ (resp. $x_n \xrightarrow[n\to\infty]{} \inf\{x_n : n \in I\}$).*

Démonstration. Nous ne ferons la preuve que pour les suites croissantes, le cas des suites décroissantes est laissé au lecteur. Distinguons deux cas.

- Si $\sup\{x_n; n \in I\} = +\infty$, c'est que l'ensemble $\{x_n : n \in I\}$ n'est pas majoré. Montrons que $x_n \to +\infty$, c'est-à-dire que $\forall \rho \in \mathbb{R}$, $\exists n_0 \in \mathbb{N}$, $\forall n \geqslant n_0$, $x_n \geqslant \rho$. Soit $\rho \in \mathbb{R}$. Puisque ρ ne peut être un majorant de $\{x_n : n \in I\}$, il existe un $n_0 \in I$ tel que $x_{n_0} > \rho$. Pour tout $n \geqslant n_0$, la croissance de la suite implique que $x_n \geqslant x_{n_0}$ et donc que $x_n \geqslant \rho$ comme désiré.

- L'autre possibilité est que $a := \sup\{x_n; n \in I\} \in \mathbb{R}$ (en effet le supremum ne peut valoir $-\infty$ car l'ensemble n'est pas vide). Il faut prouver que $x_n \to a$, c'est-à-dire que $\forall \varepsilon > 0$, $\exists n_0 \in \mathbb{N}$, $\forall n \geqslant n_0$, $|x_n - a| \leqslant \varepsilon$. Soit $\varepsilon > 0$. En vertu du point (iii) de la proposition I.16, il existe un $x_{n_0} \in \{x_n : n \in I\}$ tel que $a - \varepsilon \leqslant x_{n_0}$. Soit $n \geqslant n_0$. Vu que la suite est croissante, on a $x_n \geqslant x_{n_0}$. De plus, comme a est le supremum de $\{x_n : n \in I\}$, on a en particulier que $x_n \leqslant a$. Ainsi

$$a - \varepsilon \leqslant x_{n_0} \leqslant x_n \leqslant a < a + \varepsilon$$

et donc $|x_n - a| \leqslant \varepsilon$ comme voulu. □

Intéressons nous maintenant à la préservation ou non des inégalités par passage au supremum. Au vu des propositions I.16 et I.18 qui disent que les supremums peuvent s'obtenir par un processus de limite, on s'attend à ce que la situation soit semblable à celle de la règle sur la limite d'inégalités : les inégalités larges sont préservées tandis que les inégalités strictes peuvent devenir larges. C'est effectivement ce qui se passe.

Proposition I.20. *Soient A et B deux sous ensembles de $\mathbb{R}$. Si $\forall a \in A$, $\exists b \in B$, $a \leqslant b$ (resp. $b \leqslant a$), alors $\sup A \leqslant \sup B$ (resp. $\inf A \geqslant \inf B$).*

Cette proposition est une généralisation de $A \subseteq B \Rightarrow \sup A \leqslant \sup B$.

Démonstration. Soit $a \in A$. Par hypothèse, on sait qu'il existe un $b \in B$ tel que $a \leqslant b$. Vu que $\sup B$ est un majorant de B, on a $b \leqslant \sup B$ et donc $a \leqslant \sup B$. Puisque a est arbitraire, cela signifie que $\sup B$ est un majorant de A. Par définition du supremum de A, on a $\sup A \leqslant \sup B$. □

Ce n'est pas parce qu'on aurait $\forall a \in A,\ \exists b \in B,\ a < b$ qu'on pourrait en déduire que $\sup A < \sup B$. Il suffit pour s'en convaincre de prendre $A = B = \left]-1,0\right[$. Quel que soit $a \in \left]-1,0\right[$, on peut prendre $b := a/2 \in \left]-1,0\right[$ qui est $> a$. Pourtant $\sup A = \sup B$.

Dans cette section, nous avons présenté le supremum comme une « généralisation » du maximum qui a l'avantage de toujours exister. Terminons en expliquant quand le supremum d'un ensemble est en fait un maximum.

Proposition I.21. *Soit $A \subseteq \mathbb{R}$. L'ensemble A possède un maximum (resp. minimum) si et seulement si $\sup A \in A$ (resp. $\inf A \in A$), auquel cas $\sup A = \max A$ (resp. $\inf A = \min A$).*

Démonstration. Comme d'habitude, nous ne ferons la démonstration que pour le supremum.

Condition nécessaire. Soit $a \in A$ le maximum de A. Par définition, a est un majorant de A. De plus, si b est un autre majorant de A, $a \leqslant b$ puisque $a \in A$. Dès lors a satisfait la définition du supremum et on a $\sup A = a = \max A \in A$.

Condition suffisante. Posons $a := \sup A$. Par hypothèse $a \in A$. Par définition du supremum, a est un majorant de A. Donc a satisfait la définition du maximum — qui dès lors existe. □

I.3 Limite supérieure et inférieure

Dans cette section, nous allons utiliser les notions de supremum et d'infimum pour créer des limites qui existent toujours au sens large. Nous verrons les liens avec le concept de limite vu précédemment, ce qui nous donnera un outil supplémentaire pour prouver l'existence de limites.

L'idée de base est d'essayer de définir la limite « par au-dessus » et « par en-dessous » d'une suite. Si (x_n) est une suite et $n_0 \in \mathbb{N}$, toutes les valeurs de x_n pour $n \geqslant n_0$ se trouvent dans l'intervalle $\big[\inf\{x_n : n \geqslant n_0\}, \sup\{x_n : n \geqslant n_0\}\big]$. Ainsi, on peut voir cet intervalle comme l'espace dans lequel x_n peut se mouvoir pour $n \geqslant n_0$. Pour que cet intervalle soit représentatif de ce qui se passe « à la fin » de la suite, il faut prendre n_0 de plus en plus grand, c'est-à-dire passer à la limite $n_0 \to +\infty$. Cela conduit à la définition suivante.

Définition I.22. La *limite supérieure* (resp. *limite inférieure*) d'une suite $(x_n)_{n\in I} \subseteq \mathbb{R}$, notée $\overline{\lim}_{n\to\infty} x_n$ (resp. $\underline{\lim}_{n\to\infty} x_n$), est définie comme

$$\overline{\lim_{n\to\infty}}\, x_n := \lim_{n_0\to\infty} \sup_{n\geqslant n_0} x_n \qquad (\text{resp. } \underline{\lim_{n\to\infty}}\, x_n := \lim_{n_0\to\infty} \inf_{n\geqslant n_0} x_n).$$

Comme c'est l'usage, nous avons noté $\sup_{n\geqslant n_0} x_n$ (resp. $\inf_{n\geqslant n_0} x_n$) au lieu de $\sup\{x_n : n \geqslant n_0\}$ (resp. $\inf\{x_n : n \geqslant n_0\}$). Certains auteurs utilisent les notations $\limsup_{n\to\infty} x_n$ (resp. $\liminf_{n\to\infty} x_n$) à la place de $\overline{\lim}_{n\to\infty} x_n$ (resp. $\underline{\lim}_{n\to\infty} x_n$). Nous avons dit que ces limites existent toujours. En effet, puisque $\{x_n : n \geqslant n_0\} \supseteq \{x_n : n \geqslant n_0 + 1\}$, la proposition I.14 implique que la suite $(\sup\{x_n : n \geqslant n_0\})_{n_0\in\mathbb{N}}$ est décroissante et donc que sa limite existe et vaut l'infimum de ses valeurs (voir la propostition I.19). On peut faire le même raisonnement pour la limite inférieure. En résumé, on peut écrire

$$\overline{\lim_{n\to\infty}}\, x_n = \inf_{n_0\in\mathbb{N}} \sup_{n\geqslant n_0} x_n \qquad \text{et} \qquad \underline{\lim_{n\to\infty}}\, x_n = \sup_{n_0\in\mathbb{N}} \inf_{n\geqslant n_0} x_n.$$

Il est aussi facile de voir que $\underline{\lim}\, x_n \leqslant \overline{\lim}\, x_n$. Puisque les limites supérieure et inférieure sont des estimations du comportement de la suite à l'infini par le dessus et par le dessous respectivement, il est naturel de penser que la suite aura une limite si et seulement si les limites supérieure et inférieure coïncident.

Proposition I.23. *Soit $(x_n)_{n\in I}$ une suite de nombres réels. La suite $(x_n)_{n\in I}$ converge au sens large si et seulement si $\underline{\lim_{n\to\infty}}\, x_n = \overline{\lim_{n\to\infty}}\, x_n$, auquel cas $\lim_{n\to\infty} x_n = \underline{\lim_{n\to\infty}}\, x_n = \overline{\lim_{n\to\infty}}\, x_n$.*

Démonstration. Condition suffisante. Posons $a := \underline{\lim}\, x_n = \overline{\lim}\, x_n \in [-\infty, +\infty]$. Puisque

$$\forall n \in I, \quad \inf_{m\geqslant n} x_m \leqslant x_n \leqslant \sup_{m\geqslant n} x_m$$

et que les suites $(\inf_{m\geqslant n} x_m)_{n\in I}$ et $(\sup_{m\geqslant n} x_m)_{n\in I}$ convergent toutes deux vers a, la convergence dominée (pour la convergence vers un réel ou vers $\pm\infty$) implique que $x_n \to a$.

Condition nécessaire. Ceci découle de la proposition I.24 et du fait que les sous-suites d'une suite convergeante ont la même limite. En effet, les limites supérieures et inférieures étant des limites de sous-suites et les sous-suites ayant même limite que la suite de départ (qui existe par hypothèse), on a $\underline{\lim}\, x_n = \lim x_n = \overline{\lim}\, x_n$. □

Cette proposition donne une manière supplémentaire de prouver que la suite $\big((-1)^n\big)_{n\in\mathbb{N}}$ ne converge pas. En effet, $-1 = \underline{\lim}(-1)^n < \overline{\lim}(-1)^n = 1$. En fait, on peut relier ceci à l'existence de deux sous-suites convergeant vers des limites différentes car les limites supérieure et inférieure sont les limites de sous-suites bien choisies.

Proposition I.24. *Soit $(x_n)_{n\in I} \subseteq \mathbb{R}$. Il existe des sous-suites $(x'_m)_{m\in J}$ et $(x''_p)_{p\in K}$ telles que, au sens large, $x'_m \xrightarrow[m\to\infty]{} \overline{\lim}_{n\to\infty} x_n$ et $x''_p \xrightarrow[p\to\infty]{} \underline{\lim}_{n\to\infty} x_n$.*

Démonstration. Nous ne ferons la démonstration que pour la limite supérieure — celle pour la limite inférieure étant similaire. Appelons a la limite supérieure de (x_n). Nous ne considérerons que le cas $a \in \mathbb{R}$ — si $a = +\infty$, il suffit de remplacer les appels à la définition de convergence vers un réel et à la proposition I.16 par des utilisations de la définition de convergence vers $+\infty$ et de la proposition I.18 ; si $a = -\infty$, $\underline{\lim} x_n = \overline{\lim} x_n = -\infty$ et donc $x_n \to -\infty$. Soit $\varepsilon > 0$ et $n_0 \in \mathbb{N}$. De $\sup_{m\geqslant n} x_m \xrightarrow[n\to\infty]{} a$, on déduit qu'il existe un n_1 tel que

$$\forall n \geqslant n_1, \quad \big|\sup_{m\geqslant n} x_m - a\big| \leqslant \varepsilon/2. \tag{I.13}$$

Posons $n := \max\{n_0, n_1\}$. Par la définition équivalente du supremum (proposition I.16), on sait qu'il existe un $\mu \geqslant n$ tel que

$$-\varepsilon/2 + \sup_{m\geqslant n} x_m \leqslant x_\mu \leqslant \sup_{m\geqslant n} x_m. \tag{I.14}$$

En mettant (I.13) et (I.14) ensemble, on trouve que x_μ vérifie $a - \varepsilon \leqslant x_\mu \leqslant a + \varepsilon$. Pour résumer, on vient de montrer que

$$\forall \varepsilon > 0,\ \forall n_0 \in \mathbb{N},\ \exists \mu \geqslant n_0,\ |x_\mu - a| \leqslant \varepsilon. \tag{I.15}$$

En prenant $\varepsilon = 1$ et $n_0 = 0$ dans (I.15), on obtient qu'il existe un μ_1 tel que $|x_{\mu_1} - a| \leqslant 1$. En recommençant avec $\varepsilon = 1/2$ et $n_0 = \mu_1 + 1$, on trouve qu'il existe un $\mu_2 > \mu_1$ tel que $|x_{\mu_2} - a| \leqslant 1/2$. En continuant de la sorte, on crée une sous-suite $(x_{\mu_k})_{k\geqslant 1}$ de (x_n) telle que $\forall k \geqslant 1,\ |x_{\mu_k} - a| \leqslant 1/k$. En vertu des de la convergence dominée, la sous-suite $(x_{\mu_k})_{k\geqslant 1}$ converge vers a. Ceci termine la preuve. □

I.4 Propriété des intervalles emboités

La complétude de $\mathbb{R}$ peut s'exprimer de manière plus géométrique. Si on a une suite d'intervalles qui se « rétrécissent », on s'attend à trouver au moins un point « à la limite ». Plus précisément, si $([a_n, b_n])_{n\in\mathbb{N}}$ est une suite d'intervalles (avec $a_n, b_n \in \mathbb{R}$) emboités, i.e., $[a_{n+1}, b_{n+1}] \subseteq [a_n, b_n]$ pour tout n, alors $\bigcap_{n\in\mathbb{N}}[a_n, b_n] \neq \varnothing$ (voir figure I.8). Cette propriété, qui s'appelle la *propriété des intervalles emboités*, est équivalente à la complétude de $\mathbb{R}$.

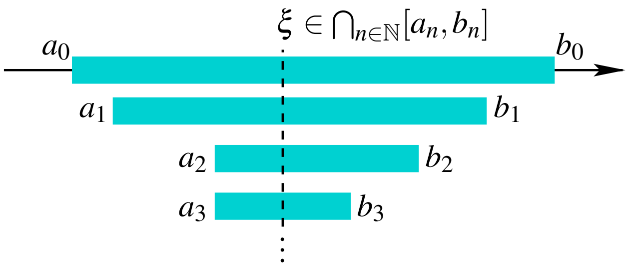


FIGURE I.8 – Propriété des intervalles emboités

Proposition I.25. *De la complétude de $\mathbb{R}$ on peut déduire la propriété des intervalles emboités et vice-versa.*

Démonstration. ($\Rightarrow$) Supposons qu'on sache que $\mathbb{R}$ est complet et déduisons-en la propriété des intervalles emboités. Soit $([a_n, b_n])_{n\in\mathbb{N}}$ une suite d'intervalles emboités. Sans perte de généralité, on peut supposer que $a_n \leqslant b_n$ (sinon les échanger). L'inclusion des intervalles s'exprime alors par

$$\forall n \in \mathbb{N}, \quad a_n \leqslant a_{n+1} \leqslant b_{n+1} \leqslant b_n.$$

Autrement dit $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ sont des suites croissante et décroissante respectivement. Vu que $(a_n)_{n\in\mathbb{N}}$ est majorée par b_0 et $(b_n)_{n\in\mathbb{N}}$ est minorée par a_0, ces deux suites convergent respectivement vers $a^* := \sup_{n\in\mathbb{N}} a_n \in \mathbb{R}$ et $b^* := \inf_{n\in\mathbb{N}} b_n \in \mathbb{R}$ (proposition I.19). Comme $a_n \leqslant b_n$ pour tout n, on a en passant à la limite que $a^* \leqslant b^*$. On va montrer que

$$\bigcap_{n\in\mathbb{N}} [a_n, b_n] = [a^*, b^*]$$

ce qui établira la thèse puisque $[a^*, b^*] \neq \varnothing$ — il contient au moins a^*. Si $x \in \bigcap_{n\in\mathbb{N}}[a_n, b_n]$, alors $a_n \leqslant x \leqslant b_n$ quel que soit n et donc, en passant à la limite,

$a^* \leqslant x \leqslant b^*$. Ceci établit l'inclusion « $\subseteq$ ». Inversément, si $x \in [a^*, b^*]$ et $n \in \mathbb{N}$, on a

$$a_n \leqslant \sup_{n\in\mathbb{N}} a_n = a^* \leqslant x \leqslant b^* = \inf_{n\in\mathbb{N}} b_n \leqslant b_n$$

et par conséquent $x \in [a_n, b_n]$.

($\Leftarrow$) Supposons maintenant que la propriété des intervalles emboités soit vraie. Soit $(x_n)_{n\in\mathbb{N}}$ une suite de Cauchy. Montrons qu'il existe un $x^* \in \mathbb{R}$ tel que $x_n \to x^*$. Pour trouver x^*, construisons une suite d'intervalles emboités telle que x^* soit dans l'intersection. Soit $k \in \mathbb{N} \setminus \{0\}$. En prenant $\varepsilon = 1/k$ dans la définition de « (x_n) est de Cauchy », on trouve qu'il existe un $n_k \in \mathbb{N}$ tel que

$$\forall n \geqslant n_k, \quad |x_n - x_{n_k}| \leqslant 1/k. \tag{I.16}$$

Quitte à remplacer n_k par $\max\{n_\ell : \ell \leqslant k\}$, on peut supposer que la suite $(n_k)_{k\geqslant 1}$ est croissante. Posons $J_k := [x_{n_k} - 1/k, x_{n_k} + 1/k]$ et $I_k := \bigcap_{\ell\leqslant k} J_\ell$. L'ensemble I_k étant une intersection (finie) d'intervalles, c'en est un lui-même s'il est non vide. C'est le cas car (I.16) implique que $|x_{n_k} - x_{n_\ell}| \leqslant 1/\ell$ si $k \geqslant \ell$ (vu qu'alors $n_k \geqslant n_\ell$) et donc que $x_{n_k} \in J_\ell$ pour tout $\ell \leqslant k$. Enfin, $I_{k+1} = I_k \cap J_{k+1} \subseteq I_k$ ce qui signifie que la suite d'intervalles $(I_k)_{k\geqslant 1}$ est emboitée. Par hypothèse, il existe donc un $x^* \in \bigcap_{k\geqslant 1} I_k$.

Nous allons montrer que $x_n \to x^*$ en vérifiant la définition de convergence. Soit $\varepsilon > 0$. Il existe un $k \geqslant 1$ tel que $1/k \leqslant \varepsilon/2$ — par exemple $k = \lceil 2/\varepsilon \rceil$. Posons $n_0 = n_k$. Soit $n \geqslant n_0$. Par (I.16), $|x_n - x_{n_k}| \leqslant 1/k$. D'autre part, comme $x^* \in I_k \subseteq J_k$, on a $|x^* - x_{n_k}| \leqslant 1/k$. Donc

$$|x^* - x_n| \leqslant |x^* - x_{n_k}| + |x_{n_k} - x_n| \leqslant 1/k + 1/k = 2/k \leqslant \varepsilon. \qquad \square$$

I.5 Annexe : construction des réels

Comme promis, nous expliquons ici une manière de construire $\mathbb{R}$ à partir des suites de Cauchy de $\mathbb{Q}$. Il y en a d'autres. On peut par exemple construire $\mathbb{R}$ à partir de coupures de $\mathbb{Q}$. Le lecteur intéressé se reportera à [1] pour plus de détails.

Comment peut-on ajouter des éléments à $\mathbb{Q}$ de manière à assurer une limite à toutes les suites de Cauchy de $\mathbb{Q}$? Après tout, nous n'avons aucune idée en général de la valeur de cette limite ! Si on réfléchit un peu, on se rend compte que les valeurs à ajouter existent par le fait qu'elles sont pointées par les suites de

Cauchy. Cependant, il faut bien se rendre compte qu'une même valeur peut être pointée par diverses suites : par exemple, les deux suites $(1/n)$ et $(1/n^2)$ tendent toutes deux vers zéro. Comment exprimer que deux suites de Cauchy pointent vers le même élément ? Il ne faut pas oublier en effet qu'il faut le faire sans parler de la limite elle-même. Intuitivement, deux suites vont avoir la même limite si et seulement si leurs éléments sont proches les uns des autres. On peut montrer ceci dans le cas où on suppose qu'on connait $\mathbb{R}$.

Proposition I.26. *Soit $(x_n)_{n\in I}$ et $(y_n)_{n\in J}$ deux suites convergentes de nombres réels. Alors les deux propriétés suivantes sont équivalentes :*

$$\lim_{n\to\infty} x_n = \lim_{n\to\infty} y_n\,; \tag{I.17}$$

$$\forall \varepsilon > 0,\ \exists n_0 \in \mathbb{N},\ \forall n \in I \cap J, \quad n \geqslant n_0 \Rightarrow |x_n - y_n| \leqslant \varepsilon. \tag{I.18}$$

Démonstration. Laissée au lecteur. □

Revenons à la construction de $\mathbb{R}$. On va prendre les suites de Cauchy dans $\mathbb{Q}$. Ces suites pointent vers des nombres réels — qui sont de cette manière implicitement définis. Deux suites de Cauchy pointent vers le même nombre réel si elles satisfont la propriété (I.18). à ce stade, nous ne connaissons rien de plus sur les nombres réels. Nous allons donc identifier un réel à *l'ensemble* des suites qui pointent sur lui.

Formalisons maintenant ces idées. Notons $\mathscr{C}$ l'ensemble des suites $(x_n)_{n\in\mathbb{N}}$ de nombres rationnels qui sont de $\mathbb{Q}$-Cauchy au sens où

$$\forall \varepsilon \in \mathbb{Q} : \varepsilon > 0,\ \exists n_0 \in \mathbb{N},\ \forall m,n \in \mathbb{N}, \quad (m \geqslant n_0 \wedge n \geqslant n_0) \Rightarrow |x_m - x_n| \leqslant \varepsilon.$$

(Nous sommes forcés à cette définition vu que nous ne pouvons pas parler des nombres réels.) Définissons la relation d'équivalence « $\sim$ » sur $\mathscr{C}$ par

$$(x_n)_{n\in\mathbb{N}} \sim (y_n)_{n\in\mathbb{N}}$$

ssi [3]

$$\forall \varepsilon \in \mathbb{Q} : \varepsilon > 0,\ \exists n_0 \in \mathbb{N},\ \forall n \in \mathbb{N}, \quad n \geqslant n_0 \Rightarrow |x_n - y_n| \leqslant \varepsilon.$$

Vérifions que c'est bien une relation d'équivalence.

3. Ceci peut être compris comme $x_n - y_n \to 0$ mais avec une définition qui n'utilise que des $\varepsilon \in \mathbb{Q}$ vu qu'on est ici en train de construire $\mathbb{R}$. Lorsqu'on disposera de l'ensemble des réels, on pourra montrer que cette définition est équivalente à $x_n - y_n \to 0$ avec la définition classique dans laquelle $\varepsilon \in \mathbb{R}$.

▪ $\sim$ est réflexive : $\forall (x_n) \in \mathscr{C},\ (x_n) \sim (x_n)$. C'est évident. En effet, pour un $\varepsilon \in \mathbb{Q}$, $\varepsilon > 0$, il suffit de prendre $n_0 := 0$ puisque, quel que soit $n \in \mathbb{N}$, $|x_n - x_n| = 0 < \varepsilon$.

▪ $\sim$ est symétrique : $\forall (x_n), (y_n) \in \mathscr{C},\ (x_n) \sim (y_n) \Leftrightarrow (y_n) \sim (x_n)$. C'est évident vu que $|x_n - y_n| = |y_n - x_n|$.

▪ $\sim$ est transitive : $\forall (x_n), (y_n), (z_n) \in \mathscr{C},\ (x_n) \sim (y_n) \wedge (y_n) \sim (z_n) \Rightarrow (x_n) \sim (z_n)$. Il faut prouver $(x_n) \sim (z_n)$, c'est-à-dire

$$\forall \varepsilon \in \mathbb{Q} : \varepsilon > 0,\ \exists n_0,\ \forall n \geqslant n_0,\ |x_n - z_n| \leqslant \varepsilon.$$

Soit $\varepsilon \in \mathbb{Q}$, $\varepsilon > 0$. Les définitions de $(x_n) \sim (y_n)$ et $(y_n) \sim (z_n)$ impliquent respectivement que

$$\exists n_1,\ \forall n \geqslant n_1,\ |x_n - y_n| \leqslant \varepsilon/2 \qquad \text{et} \qquad \exists n_2,\ \forall n \geqslant n_2,\ |y_n - z_n| \leqslant \varepsilon/2.$$

Posons $n_0 := \max\{n_1, n_2\}$. Si $n \geqslant n_0$, on a $|x_n - z_n| \leqslant |x_n - y_n| + |y_n - z_n| \leqslant \varepsilon/2 + \varepsilon/2 = \varepsilon$.

Nous sommes maintenant prêts à définir $\mathbb{R}$.

Définition I.27. $\mathbb{R} := (\mathscr{C}/\sim) := \{[(x_n)] : (x_n) \in \mathscr{C}\}$ où $[(x_n)] := \{(y_n) \in \mathscr{C} : (y_n) \sim (x_n)\}$.

L'ensemble $[(x_n)]$ est appelée la *classe d'équivalence* de (x_n). Elle est constituée des suites qui pointent vers le même réel que (x_n). C'est bien ce à quoi on avait décidé d'identifier les nombres réels. Le fait que $\sim$ est une relation d'équivalence implique que

$$[(x_n)] = [(y_n)] \Leftrightarrow (x_n) \sim (y_n) \tag{I.19}$$

(démontrez-le !). On identifie les rationnels aux classes d'équivalence des suites constantes. Plus précisément, on définit l'injection

$$i : \mathbb{Q} \to \mathbb{R} : q \mapsto \big[(q)_{n\in\mathbb{N}}\big].$$

C'est bien une injection car $[(p)_{n\in\mathbb{N}}] = [(q)_{n\in\mathbb{N}}]$ implique que $p = q$ (adaptez la preuve de l'unicité de la limite).

Le lemme suivant renforce l'idée qu'on considère les nombres qui sont « pointés par » les suites de Cauchy au sens où, si la suite converge dans $\mathbb{Q}$, alors le nombre vers lequel elle pointe est précisément cette limite.

Lemme I.28. *Soit* $(x_n)_{n\in\mathbb{N}} \in \mathscr{C}$ *et* $q \in \mathbb{Q}$. *Si* $(x_n)_{n\in\mathbb{N}}$ $\mathbb{Q}$*-converge vers* q *au sens où*

$$\forall \varepsilon \in \mathbb{Q} : \varepsilon > 0,\ \exists n_0 \in \mathbb{N},\ \forall n \geqslant n_0,\ |x_n - q| \leqslant \varepsilon \tag{I.20}$$

alors $[(x_n)_{n\in\mathbb{N}}] = [(q)_{n\in\mathbb{N}}]$.

Démonstration. C'est évident car (I.20) implique immédiatement que $(x_n)_{n\in\mathbb{N}} \sim (q)_{n\in\mathbb{N}}$. □

Nous avons vu que toute sous-suite d'une suite convergente avait même limite que celle-ci. Il est donc naturel qu'une sous-suite d'une suite de $\mathbb{Q}$-Cauchy pointe vers le même nombre réel. Ce résultat sera utile à diverses reprises.

Lemme I.29. *Soit* $(x_n)_{n\in\mathbb{N}} \in \mathscr{C}$. *Si* $(x'_m)_{m\in\mathbb{N}}$ *est une sous-suite de* $(x_n)_{n\in\mathbb{N}}$ *alors* $(x'_m)_{m\in\mathbb{N}} \in \mathscr{C}$ *et* $(x'_m)_{m\in\mathbb{N}} \sim (x_n)_{n\in\mathbb{N}}$, *c'est-à-dire* $[(x'_m)_{m\in\mathbb{N}}] = [(x_n)_{n\in\mathbb{N}}]$.

Démonstration. Comme $(x'_m)_{m\in\mathbb{N}} \subseteq (x_n)_{n\in\mathbb{N}}$, on a par définition qu'il existe une fonction strictement croissante $\varphi : \mathbb{N} \to \mathbb{N}$ telle que

$$\forall m \in \mathbb{N}, \quad x'_m = x_{\varphi(m)}.$$

Comme dans la démonstration de la convergence des sous-suites d'une suite convergeante, il est facile de prouver par récurrence que

$$\forall m \in \mathbb{N}, \quad \varphi(m) \geqslant m. \tag{I.21}$$

D'autre part, comme $(x_n) \in \mathscr{C}$, on a

$$\forall \varepsilon_1 \in \mathbb{Q} : \varepsilon_1 > 0,\ \exists n_0 \in \mathbb{N},\ \forall n_1, n_2 \geqslant n_0,\ |x_{n_1} - x_{n_2}| \leqslant \varepsilon_1. \tag{I.22}$$

- Commençons par prouver que $(x'_m) \in \mathscr{C}$, c'est-à-dire que

$$\forall \varepsilon \in \mathbb{Q} : \varepsilon > 0,\ \exists m_0 \in \mathbb{N},\ \forall m_1, m_2 \geqslant m_0,\ |x'_{m_1} - x'_{m_2}| \leqslant \varepsilon.$$

Soit $\varepsilon \in \mathbb{Q}$, $\varepsilon > 0$. Par (I.22) avec $\varepsilon_1 = \varepsilon$, on trouve qu'il existe un $n_0 \in \mathbb{N}$ tel que $\forall n_1, n_2 \geqslant n_0$, $|x_{n_1} - x_{n_2}| \leqslant \varepsilon$. Prenons $m_0 := n_0$. Si $m_1, m_2 \geqslant m_0$, par (I.21) $\varphi(m_1)$ et $\varphi(m_2)$ sont plus grand ou égaux à $m_0 = n_0$ et donc $|x'_{m_1} - x'_{m_2}| = |x_{\varphi(m_1)} - x_{\varphi(m_2)}| \leqslant \varepsilon$ comme désiré.

- Montrons maintenant que $(x'_m) \sim (x_n)$, c'est-à-dire que

$$\forall \varepsilon \in \mathbb{Q} : \varepsilon > 0,\ \exists k_0 \in \mathbb{N},\ \forall k \geqslant k_0,\ |x'_k - x_k| \leqslant \varepsilon.$$

Soit $\varepsilon \in \mathbb{Q}$, $\varepsilon > 0$. Comme précédemment, on utilise (I.22) avec $\varepsilon_1 = \varepsilon$ pour trouver un n_0 et on pose $k_0 := n_0$. Si $k \geqslant k_0$, $\varphi(k) \geqslant k \geqslant k_0$ et donc (I.22) implique que $|x'_k - x_k| = |x_{\varphi(k)} - x_k| \leqslant \varepsilon$. □

La suite de l'exposé s'agence comme suit. Nous allons d'abord donner un sens aux définitions de convergence et d'être de Cauchy en munissant $\mathbb{R}$ d'opérations algébriques et d'une relation d'ordre qui étendent celles de $\mathbb{Q}$. Ensuite nous montrerons que toute suite $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{Q}$ de Cauchy pour ces définitions est en fait de $\mathbb{Q}$-Cauchy et donc converge dans $\mathbb{R}$. Enfin, un argument diagonal sera utilisé pour prouver que toute suite de Cauchy $(x_n)_{n\in I} \subseteq \mathbb{R}$ converge — c'est-à-dire que $\mathbb{R}$ est complet.

Commençons par définir les opérations d'addition et de multiplication sur $\mathbb{R}$ par

$$[(x_n)] + [(y_n)] := [(x_n + y_n)] \quad \text{et} \quad [(x_n)] \cdot [(y_n)] := [(x_n y_n)] \tag{I.23}$$

Ces définitions posent néanmoins a priori un problème. En effet, nous avons défini l'addition de deux classes d'équivalence $[(x_n)]$ et $[(y_n)]$ en choisissant des représentants (x_n) et (y_n) de celles-ci et en constituant la classe $[(x_n + y_n)]$. Mais si on avait pris d'autres représentants (x'_n) et (y'_n) de ces mêmes classes, c'est-à-dire si $[(x'_n)] = [(x_n)]$ et $[(y'_n)] = [(y_n)]$, aurait-on eu le même résultat : $[(x'_n + y'_n)] = [(x_n + y_n)]$? Au vu de (I.19), répondre positivement à cette question revient à montrer

$$\big((x'_n) \sim (x_n) \text{ et } (y'_n) \sim (y_n)\big) \Rightarrow (x'_n + y'_n) \sim (x_n + y_n). \tag{I.24}$$

De même, pour que la multiplication soit bien définie, il faut vérifier que :

$$\big((x'_n) \sim (x_n) \text{ et } (y'_n) \sim (y_n)\big) \Rightarrow (x'_n y'_n) \sim (x_n y_n). \tag{I.25}$$

Proposition I.30. *(I.24) et (I.25) sont vraies quelles que soient les suites (x_n), $(x'_n), (y'_n), (y_n) \in \mathscr{C}$.*

Nous aurons besoin du lemme suivant.

Lemme I.31. *Toute suite $(x_n)_{n\in\mathbb{N}}$ de $\mathscr{C}$ est bornée au sens où $\exists R \in \mathbb{Q}$, $\forall n \in \mathbb{N}$, $|x_n| \leqslant R$.*

Démonstration. En prenant $\varepsilon = 1 \in \mathbb{Q}$ dans la définition du fait que (x_n) est de $\mathbb{Q}$-Cauchy, on obtient

$$\exists n_0 \in \mathbb{N},\ \forall m, n \geqslant n_0,\ |x_n - x_m| \leqslant 1. \tag{I.26}$$

Posons $R := \max\{|x_0|, \dots, |x_{n_0-1}|, 1 + |x_{n_0}|\} \in \mathbb{Q}$ et montrons que c'est une borne. Soit $n \in \mathbb{N}$. Si $n < n_0$, il est évident que $|x_n| \leqslant R$. Si $n \geqslant n_0$, en utilisant (I.26) avec $m = n_0$, on déduit que $|x_n| \leqslant |x_n - x_{n_0}| + |x_{n_0}| \leqslant 1 + |x_{n_0}| \leqslant R$. □

Démonstration de la proposition I.30. Cette démonstration est fort similaire à celle des règles de calculs pour les limites de suites.

- Commençons par (I.24). Il faut prouver que $(x'_n + y'_n) \sim (x_n + y_n)$ c'est-à-dire que

$$\forall \varepsilon \in \mathbb{Q} : \varepsilon > 0,\ \exists n_0 \in \mathbb{N},\ \forall n \geqslant n_0,\ |x'_n + y'_n - x_n - y_n| \leqslant \varepsilon.$$

Soit $\varepsilon \in \mathbb{Q}$, $\varepsilon > 0$. Par définition de $(x'_n) \sim (x_n)$ et $(y'_n) \sim (y_n)$, il existe des naturels n_1 et n_2 tels que

$$\forall n \geqslant n_1,\ |x'_n - x_n| \leqslant \varepsilon/2 \qquad \text{et} \qquad \forall n \geqslant n_2,\ |y'_n - y_n| \leqslant \varepsilon/2.$$

Il suffit de prendre $n_0 := \max\{n_1, n_2\}$ car, pour tout $n \geqslant n_0$, $|x'_n + y'_n - x_n - y_n| \leqslant |x'_n + x_n| + |y'_n - y_n| \leqslant \varepsilon/2 + \varepsilon/2 = \varepsilon$.

- Pour (I.25), il faut établir que $(x_n y_n) \sim (x_n y_n)$, c'est-à-dire que

$$\forall \varepsilon \in \mathbb{Q} : \varepsilon > 0,\ \exists n_0 \in \mathbb{N},\ \forall n \geqslant n_0,\ |x'_n y'_n - x_n y_n| \leqslant \varepsilon.$$

Soit $\varepsilon \in \mathbb{Q}$, $\varepsilon > 0$. On sait qu'il existe des bornes $R, R' \in \mathbb{Q}$ telles que $|x_n| \leqslant R$ et $|y'_n| \leqslant R'$ pour tout n. Sans perte de généralité, on peut supposer que $R > 0$ est $R' > 0$ — sinon les redéfinir comme $R := \max\{R, 1\}$ et $R' := \max\{R', 1\}$. Des hypothèses $(x'_n) \sim (x_n)$ et $(y'_n) \sim (y_n)$ on déduit qu'il existe des naturels n_1 et n_2 tels que

$$\forall n \geqslant n_1,\ |x'_n - x_n| \leqslant \varepsilon/(2R') \qquad \text{et} \qquad \forall n \geqslant n_2,\ |y'_n - y_n| \leqslant \varepsilon/(2R).$$

Posons $n_0 := \max\{n_1, n_2\}$ Dès lors, pour tout $n \geqslant n_0$, on a $|x'_n y'_n - x_n y_n| = |(x'_n - x_n) y'_n + x_n (y'_n - y_n)| \leqslant |x'_n - x_n|\,|y'_n| + |x_n||y'_n - y_n| \leqslant \varepsilon/(2R')R' + R(\varepsilon/2R) = \varepsilon$. □

Pour que les définitions (I.23) soient satisfaisantes, nous voudrions qu'elles étendent les opérations de $\mathbb{Q}$. Plus précisément, si $p, q \in \mathbb{Q}$, on peut faire $p+q$ dans $\mathbb{Q}$ ou voir p et q comme les réels $[(p)_{n\in\mathbb{N}}]$ et $[(q)_{n\in\mathbb{N}}]$ et former $[(p)_{n\in\mathbb{N}}] + [(q)_{n\in\mathbb{N}}]$. On voudrait que les deux opérations donnent le même résultat au sens où le réel correspondant à $p+q$ et $[(p)_{n\in\mathbb{N}}] + [(q)_{n\in\mathbb{N}}]$ sont les mêmes, c'est-à-dire $[(p+q)_{n\in\mathbb{N}}] = [(p)_{n\in\mathbb{N}}] + [(q)_{n\in\mathbb{N}}]$. C'est évidemment le cas au vu de (I.23). Un raisonnement similaire montre que la multiplication sur les réels étend celle sur $\mathbb{Q}$. On peut représenter graphiquement ces résultats en disant que les diagrammes des figures I.9 et I.10 sont commutatifs, c'est-à-dire que les deux fonctions allant de $\mathbb{Q}\times\mathbb{Q}$ vers $\mathbb{R}$ sont les mêmes.

$$\begin{tikzcd}
\mathbb{Q}\times\mathbb{Q} \arrow[r, "+"] \arrow[d, "i\times i"'] & \mathbb{Q} \arrow[d, "i"] \\
\mathbb{R}\times\mathbb{R} \arrow[r, "+"] & \mathbb{R}
\end{tikzcd}
\qquad
\begin{tikzcd}
(p,q) \arrow[r, mapsto] \arrow[d, mapsto] & p+q \arrow[d, mapsto] \\
\big([(p)],[(q)]\big) \arrow[r, mapsto] & {[(p)]+[(q)] = [(p+q)]}
\end{tikzcd}$$

FIGURE I.9 – Extension de l'addition de $\mathbb{Q}$

$$\begin{tikzcd}
\mathbb{Q}\times\mathbb{Q} \arrow[r, "\cdot"] \arrow[d, "i\times i"'] & \mathbb{Q} \arrow[d, "i"] \\
\mathbb{R}\times\mathbb{R} \arrow[r, "\cdot"] & \mathbb{R}
\end{tikzcd}
\qquad
\begin{tikzcd}
(p,q) \arrow[r, mapsto] \arrow[d, mapsto] & pq \arrow[d, mapsto] \\
\big([(p)],[(q)]\big) \arrow[r, mapsto] & {[(p)]\cdot[(q)] = [(pq)]}
\end{tikzcd}$$

FIGURE I.10 – Extension de la multiplication de $\mathbb{Q}$

Le résultat suivant montre que la structure algébrique qu'on vient de mettre sur $\mathbb{R}$ répond bien à nos attentes.

Proposition I.32. $(\mathbb{R}, +, \cdot)$ *est un corps commutatif.*

Démonstration. Vu que (I.23) définit l'addition et la multiplication sur $\mathbb{R}$ à partir de celle sur $\mathbb{Q}$, dont on sait qu'il est un corps commutatif, les choses suivantes sont évidentes :

- l'addition sur $\mathbb{R}$ est associative et commutative, a pour neutre $0_\mathbb{R} := [(0)_{n\in\mathbb{N}}]$ et l'opposé de $[(x_n)_{n\in\mathbb{N}}]$ est donné par $[(-x_n)_{n\in\mathbb{N}}]$;

- la multiplication sur $\mathbb{R}$ est associative, commutative et a pour neutre $1_{\mathbb{R}} := [(1)_{n\in\mathbb{N}}]$;
- la multiplication se distribue sur l'addition.

La seule chose qu'il reste à montrer est que, tout $x \in \mathbb{R} \setminus \{0_{\mathbb{R}}\}$ possède un inverse. Si $x = [(x_n)_{n\in\mathbb{N}}]$, on a envie de définir l'inverse par $[(1/x_n)_{n\in\mathbb{N}}]$ car alors $[(x_n)] \cdot [(1/x_n)] = [(x_n/x_n)] = 1_{\mathbb{R}}$. Mais pour faire ceci, il faut montrer qu'on peut trouver un représentant (x_n) de la classe d'équivalence x tel que $\forall n \in \mathbb{N},\ x_n \neq 0$. Supposons au contraire qu'il n'y ait pas de tel représentant et montrons que $x = 0_{\mathbb{R}}$. Par hypothèse on a donc

$$\forall (x_n) \in x,\ \exists n \in \mathbb{N},\ x_n = 0. \tag{I.27}$$

Prenons au hasard un représentant $(\xi_n)_{n\in\mathbb{N}} \in x$ (les classes d'équivalence sont toujours non-vides par définition). Par (I.27), il existe un $v_0 \in \mathbb{N}$ tel que $\xi_{v_0} = 0$. Considérons la suite $(\xi_{n+v_0+1})_{n\in\mathbb{N}}$ qui est une sous-suite de (ξ_n) et par conséquent $(\xi_{n+v_0+1}) \sim (x_n)$ (adaptez la preuve de la convergence des sous-suites d'une suite convergeante). Dès lors, $(\xi_{n+v_0+1}) \in x$ et (I.27) implique qu'il existe un $n_1 \in \mathbb{N}$, tel que $\xi_{n_1+v_0+1} = 0$. Posons $v_1 := n_1 + v_0 + 1$. On a

$$v_1 > v_0 \quad \text{et} \quad \xi_{v_1} = 0.$$

En répétant le même argument avec v_1 au lieu de v_0, on va trouver un $v_2 \in \mathbb{N}$ tel que

$$v_2 > v_1 \quad \text{et} \quad \xi_{v_2} = 0.$$

En continuant de la sorte, on obtiendra une suite $(v_k)_{k\in\mathbb{N}} \subseteq \mathbb{N}$ telle que

$$(v_k)_{k\in\mathbb{N}} \text{ est strictement croissante} \quad \text{et} \quad \forall k \in \mathbb{N},\ \xi_{v_k} = 0. \tag{I.28}$$

Ainsi (ξ_{v_k}) est une sous-suite de (x_n) et donc $(\xi_{v_k}) \sim (x_n)$ ce qui implique que $x = [(\xi_{v_k})_{k\in\mathbb{N}}] = [(0)_{n\in\mathbb{N}}] = 0_{\mathbb{R}}$. □

Définissons maintenant une relation d'ordre sur $\mathbb{R}$ qui étend celle de $\mathbb{Q}$.

Définition I.33. Soit $x \in \mathbb{R}$. On dit que $x \geqslant 0_{\mathbb{R}}$ si et seulement si $x = 0_{\mathbb{R}}$ ou $x > 0_{\mathbb{R}}$ où $x > 0_{\mathbb{R}}$ est défini comme

$$\forall (x_n) \in x,\ \exists \varepsilon \in \mathbb{Q} : \varepsilon > 0,\ \exists n_0 \in \mathbb{N},\ \forall n \geqslant n_0,\ x_n \geqslant \varepsilon. \tag{I.29}$$

Si $x, y \in \mathbb{R}$, on note $x \geqslant y$ pour $x - y \geqslant 0_{\mathbb{R}}$.

L'intuition qui se cache derrière cette définition est que, si $x > 0_{\mathbb{R}}$, alors, quelle que soit la suite (x_n) qui converge vers x, ses éléments x_n seront $\geqslant \varepsilon := x/2 > 0$ pour n assez grand. Il suffit en fait de le vérifier pour une seule suite (x_n) car toutes les autres s'en rapprochent.

Lemme I.34. *Soit $x \in \mathbb{R}$. Le fait que $x > 0_{\mathbb{R}}$ est équivalent à*

$$\exists (x_n) \in x,\ \exists \varepsilon \in \mathbb{Q} : \varepsilon > 0,\ \exists n_0 \in \mathbb{N},\ \forall n \geqslant n_0,\ x_n > \varepsilon. \tag{I.30}$$

Démonstration. Il est clair que (I.29) $\Rightarrow$ (I.30). Il reste à montrer l'implication inverse. Prenons $(x'_n)_{n\in\mathbb{N}} \in x$ et vérifions que

$$\exists \varepsilon' \in \mathbb{Q} : \varepsilon' > 0,\ \exists n'_0 \in \mathbb{N},\ \forall n \geqslant n'_0,\ x'_n > \varepsilon'.$$

La définition de $(x'_n) \sim (x_n)$ implique que

$$\exists n_1 \in \mathbb{N},\ \forall n \geqslant n_1,\ |x'_n - x_n| \leqslant \varepsilon/2$$

où le ε est celui de (I.30). Prenons $\varepsilon' := \varepsilon/2$ et $n'_0 := \max\{n_0, n_1\}$. Si $n \geqslant n'_0$, on a $x'_n = x'_n - x_n + x_n \geqslant -|x'_n - x_n| + x_n \geqslant -\varepsilon/2 + \varepsilon = \varepsilon'$. □

Il est clair que cet ordre étend celui de $\mathbb{Q}$ car, pour tout $p, q \in \mathbb{Q}$, $p > q$ si et seulement si $\exists \varepsilon \in \mathbb{Q} : \varepsilon > 0,\ p - q > \varepsilon$ si et seulement si $[(p)_{n\in\mathbb{N}}] > [(q)_{n\in\mathbb{N}}]$.

Nous voudrions maintenant vérifier que :

Proposition I.35. *$\geqslant$ est une relation d'ordre total sur $\mathbb{R}$ qui est compatible avec l'addition et la multiplication.*

Rappelons ces différents termes. Le fait que « $\geqslant$ » soit une relation d'ordre veut dire que[4]

- $\geqslant$ est réflexive : $\forall x,\ x \geqslant x$;
- $\geqslant$ est antisymétrique : $\forall x, y\ (x \geqslant y \wedge y \geqslant x) \Rightarrow x = y$;
- $\geqslant$ est transitive : $\forall x, y, z,\ (x \geqslant y \wedge y \geqslant z) \Rightarrow x \geqslant z$.

La relation d'ordre

4. Techniquement, dans les formules quantifiées qui suivent, nous aurions dû préciser que $x, y, z \in \mathbb{R}$. Nous ne l'avons pas fait, d'abord pour ne pas alourdir les notations, mais surtout parce que nous voulons attirer l'attention sur le caractère général de ces propriétés : un corps ordonné est précisément un corps K muni d'une relation $\geqslant$ qui vérifie ces six propriétés — et donc toutes leurs conséquences.

- $\geqslant$ est totale : $\forall x, y,\ x \geqslant y \vee y \geqslant x$.

Enfin « $\geqslant$ » est compatible avec l'addition et la multiplication signifie que

- compatibilité avec l'addition : $\forall x, y, z,\ x \geqslant y \Rightarrow x + z \geqslant y + z$;
- compatibilité avec la multiplication : $\forall x, y\ (x \geqslant 0 \wedge y \geqslant 0) \Rightarrow xy \geqslant 0$.

Ces propriétés permettent de déduire tous les faits usuels que vous connaissez à partir de l'ordre sur $\mathbb{R}$. Établissons pour commencer que $1 \geqslant 0$. Puisque l'ordre est total, $1 \geqslant 0$ ou $0 \geqslant 1$. Si $0 \geqslant 1$, on ajoutant -1 aux deux membres, on obtient $-1 \geqslant 0$. La compatibilité avec la multiplication implique alors que $1 = (-1)(-1) \geqslant 0$. Or $0 \geqslant 1$ et $1 \geqslant 0$ impliquent la contradiction [5] $1_{\mathbb{R}} = 0_{\mathbb{R}}$, ce qui montre que le cas $0 \geqslant 1$ ne peut avoir lieu. Une autre conséquence de ces propriétés est que

$$\forall x \in \mathbb{R}, \quad x \geqslant 0 \text{ ou } -x \geqslant 0.$$

En effet, soit $x \in \mathbb{R}$. Il suffit de montrer que si $x \not\geqslant 0$ alors $-x \geqslant 0$. Puisque l'ordre est total, on a $x \geqslant 0 \vee 0 \geqslant x$. Comme $x \not\geqslant 0$, on a nécessairement que $0 \geqslant x$. Il suffit d'ajouter $-x$ aux deux membres de cette inégalité. Comme troisième exemple, montrons que

$$\forall x \in \mathbb{R}, \quad x \geqslant 0 \Leftrightarrow 0 \geqslant -x. \tag{I.31}$$

C'est évident car il suffit d'ajouter $-x$ (pour $\Rightarrow$) ou x (pour $\Leftarrow$) aux deux membres de l'inégalité. De la même manière, on peut déduire (essayez !) les règles habituelles :

- $\forall x, y, z,\ (x \geqslant y \wedge z \geqslant 0) \Rightarrow xz \geqslant yz$;
- $\forall x, y, z,\ (x \geqslant y \wedge z \leqslant 0) \Rightarrow xz \leqslant yz$;
- $\forall x,\ x^2 \geqslant 0$.

On peut bien entendu définir la valeur absolue d'un nombre réel par

$$|x| := \begin{cases} x & \text{si } x \geqslant 0 \\ -x & \text{si } x \leqslant 0 \end{cases}$$

et montrer, grâce à (I.31), que $\forall x,\ |x| \geqslant 0$.

Revenons à la preuve de la proposition I.35.

Démonstration de la proposition I.35. Puisque $x \geqslant y$ est défini comme $x - y \in K$ où $K := \{x \in \mathbb{R} : x \geqslant 0\}$, il suffit de montrer que K vérifie

5. Rappelons que $\mathbb{Q} \to \mathbb{R} : q \mapsto [(q)_{q \in \mathbb{N}}] =: q_{\mathbb{R}}$ est injective.

(i) $\forall x, y \in K,\ x+y \in K$;

(ii) $\forall x, y \in K,\ xy \in K$;

(iii) $K \cap (-K) = \{0\}$;

(iv) $K \cup (-K) = \mathbb{R}$

où $-K := \{-x \in \mathbb{R} : x \in K\}$ En effet, supposons que K vérifie (i)–(iv) et montrons que $\geqslant$ est une relation d'ordre total compatible avec l'addition et la multiplication :

- réflexivité : $x \geqslant x$ car $0 \in K$ (vu que $0 \in \{0\} = K \cap (-K) \subseteq K$);
- antisymétrie : si $x - y \in K$ et $y - x \in K$, on a $x - y \in K \cap (-K) = \{0\}$ et donc $x - y = 0$;
- transitivité : si $x - y \in K$ et $y - z \in K$, de (i) on déduit que $x - z = (x-y) + (y-z) \in K$;
- ordre total : en effet $x - y \in \mathbb{R} = K \cup (-K)$ et donc $x - y \in K$ ou $x - y \in -K$, c'est-à-dire $x - y \in K$ ou $-(x-y) \in K$;
- compatibilité avec l'addition : si $x - y \in K$ alors $(x+z) - (y+z) = x - y \in K$;
- compatibilité avec la multiplication : si $x, y \in K$, (ii) implique que $xy \in K$.

Il reste donc à prouver (i)–(iv).

(i) Soient $x, y \in K$. Il faut prouver que $x + y \in K$, c'est-à-dire que

$$\forall (z_n) \in x+y,\ \exists \varepsilon \in \mathbb{Q} : \varepsilon > 0,\ \exists n_0 \in \mathbb{N},\ \forall n \geqslant n_0,\ z_n \geqslant \varepsilon.$$

Soit $(z_n) \in x+y$. Prenons $(x_n) \in x$ et $(y_n) \in y$. Comme $x+y = [(x_n+y_n)]$, il vient en vertu de (I.19) que $(z_n) \sim (x_n+y_n)$. Par hypothèse $x \in K$ et $y \in K$ et donc il existe $\varepsilon_1, \varepsilon_2 \in \mathbb{Q}$, $\varepsilon_1 > 0$, $\varepsilon_2 > 0$ et $n_1, n_2 \in \mathbb{N}$ tels que

$$\forall n \geqslant n_1, x_n \geqslant \varepsilon_1 \quad \text{et} \quad \forall n \geqslant n_2, y_n \geqslant \varepsilon_2. \tag{I.32}$$

Posons $\varepsilon_3 := \min\{\varepsilon_1, \varepsilon_2\}$. Puisque $\varepsilon_3 \in \mathbb{Q}$ et $\varepsilon_3 > 0$, la définition de $(z_n) \sim (x_n+y_n)$ implique qu'il existe un $n_3 \in \mathbb{N}$ tel que

$$\forall n \geqslant n_3, \quad |z_n - (x_n+y_n)| \leqslant \varepsilon_3. \tag{I.33}$$

Prenons $\varepsilon := \varepsilon_3$ et $n_0 := \max\{n_1, n_2, n_3\}$. Soit $n \geqslant n_0$. Il faut prouver que $z_n \geqslant \varepsilon$. (I.33) implique que $z_n \geqslant x_n + y_n - \varepsilon_3$ et donc, en utilisant (I.32), on a $z_n \geqslant \varepsilon_1 + \varepsilon_2 - \varepsilon_3 \geqslant \varepsilon_3 + \varepsilon_3 - \varepsilon_3 = \varepsilon_3 = \varepsilon$.

(ii) Ce point se démontre de manière similaire à (i) et est laissé au lecteur.

(iii) Par définition de l'ordre, il est clair que $0 \in K$ et donc que $0 \in K \cap (-K)$ d'où $\{0\} \subseteq K \cap (-K)$. Il reste à prouver l'inclusion inverse, c'est-à-dire $x \in K \cap (-K) \Rightarrow x = 0$. Comme on a déjà montré que $0 \in K \cap (-K)$, il suffit de voir qu'il n'y a pas de $x \neq 0$ qui appartienne à $K \cap (-K)$. Supposons au contraire qu'il existe un $x \neq 0$ tel que $x \in K$ et $-x \in K$, c'est-à-dire un $x \in \mathbb{R}$ tel que $x > 0$ et $-x > 0$. En d'autres termes, au vu de la définition I.33, nous supposons que, quel que soit $(x_n) \in x$,

$$\begin{aligned} &\exists \varepsilon_1 \in \mathbb{Q} : \varepsilon_1 > 0,\ \exists n_1 \in \mathbb{N},\ \forall n \geqslant n_1,\ x_n \geqslant \varepsilon_1 \\ \text{et}\quad &\exists \varepsilon_2 \in \mathbb{Q} : \varepsilon_2 > 0,\ \exists n_2 \in \mathbb{N},\ \forall n \geqslant n_2,\ -x_n \geqslant \varepsilon_2 \end{aligned}$$

Posons $n := \max\{n_1, n_2\}$. Comme ce n est à la fois plus grand que n_1 et n_2, nous en déduisons que $x_n \leqslant -\varepsilon_2 < 0 < \varepsilon_1 \leqslant x_n$. Cette contradiction termine l'argument.

(iv) Comme il est évident que $K \cup (-K) \subseteq \mathbb{R}$, il suffit de démontrer l'inclusion opposée. Soit $x \in \mathbb{R}$. Nous allons prouver que si ni $x > 0$, ni $-x > 0$, alors $x = 0$. Le fait que $x \not> 0$ implique qu'il existe une suite $(x_n) \in x$ telle que

$$\forall \varepsilon \in \mathbb{Q} : \varepsilon > 0,\ \forall n_0 \in \mathbb{N},\ \exists n \geqslant n_0,\ x_n < \varepsilon. \tag{I.34}$$

En prenant $\varepsilon = 1$ et $n_0 = 1$, (I.34) nous dit qu'il existe un n, que nous appellerons μ_0, tel que $x_{\mu_0} < 1$. Ensuite, en réutilisant (I.34) avec $\varepsilon = 1/2$ et $n_0 = \mu_0 + 1$, nous obtenons l'existence d'un $\mu_1 > \mu_0$ tel que $x_{\mu_1} < 1/2$. En continuant de la sorte, on crée une suite $(\mu_k)_{k \in \mathbb{N}}$ telle que

$$\forall k \in \mathbb{N}, \qquad \mu_k < \mu_{k+1} \quad \text{et} \quad x_{\mu_k} < \frac{1}{k+1}.$$

Comme $(x_{\mu_k})_{k \in \mathbb{N}}$ est une sous-suite de (x_n), on a que $x = [(x_{\mu_k})_{k \in \mathbb{N}}]$. En utilisant le lemme I.34 et $-x \not> 0$, on peut recommencer la même procédure à partir de la suite $(x_{\mu_k})_{k \in \mathbb{N}} \in x$ pour obtenir une sous-suite $(x_{\mu_{\kappa(\ell)}})_{\ell \in \mathbb{N}}$ de (x_{μ_k}) telle que

$$\forall \ell \in \mathbb{N}, \qquad -x_{\mu_{\kappa(\ell)}} > \frac{1}{\ell+1} \quad \text{et} \quad x_{\mu_{\kappa(\ell)}} < \frac{1}{\kappa(\ell)+1}.$$

On prouve par récurrence à partir de la croissance stricte de $(\kappa(\ell))_{\ell \in \mathbb{N}}$ que $\kappa(\ell) \geqslant \ell$. Dès lors il vient aisément (copier la démonstration de la convergence dominée) que $x_{\mu_{\kappa(\ell)}} \xrightarrow[\ell\to\infty]{} 0$. Comme $(x_{\mu_{\kappa(\ell)}})$ est une sous-suite de (x_n), on a $x = [(x_{\mu_{\kappa(\ell)}})_{\ell\in\mathbb{N}}]$ et le lemme I.28 entraîne que $x = 0$. $\square$

Une dernière propriété de l'ordre sur $\mathbb{R}$ est que tout réel x est majoré par un entier. cette propriété est appelée « axiome d'Archimède ». C'est celui-ci qui permet de définir $\lceil x\rceil$, le plus petit entier $\geqslant x$, que nous avons utilisé à diverses reprises dans les sections précédentes.

Proposition I.36 (Axiome d'Archimède)**.** *Pour tout $x \in \mathbb{R}$, il existe un $n \in \mathbb{N}$ tel que $x \leqslant n$.*

N'oubliez pas que nous avons identifié $\mathbb{Q}$ — à fortiori $\mathbb{N}$ — à un sous-ensemble de $\mathbb{R}$. L'inégalité $x \leqslant n$ a donc un sens — et signifie explicitement que $x \leqslant [(n)_{k\in\mathbb{N}}]$.

Démonstration. C'est vrai si $x \in \mathbb{Q}$ puisque, si $x = p/q$, $p \in \mathbb{Z}$, $q \in \mathbb{N}$, alors $x \leqslant |p|$. Soit $x \in \mathbb{R}$. Vu ce qui vient d'être dit, il suffit de trouver un $q \in \mathbb{Q}$ tel que $x < q$, c'est-à-dire, vu le lemme I.34, tel que

$$\exists (x_n) \in x,\ \exists \varepsilon \in \mathbb{Q} : \varepsilon > 0,\ \exists n_0,\ \forall n \geqslant n_0,\ q - x_n \geqslant \varepsilon.$$

Soit $(x_n)_{n\in\mathbb{N}} \in x$. Puisque (x_n) est de $\mathbb{Q}$-Cauchy, il vient

$$\exists n_1,\ \forall m,n \geqslant n_1,\ |x_n - x_m| \leqslant 1.$$

Posons $q := 2 + |x_{n_1}| \in \mathbb{Q}$, $\varepsilon := 1$ et $n_0 := n_1$. Si $n \geqslant n_0$, en utilisant le fait que $\big||x_{n_1}| - |x_n|\big| \leqslant |x_{n_1} - x_n|$, on déduit que $q - x_n = 2 + |x_{n_1}| - x_n \geqslant 2 - |x_{n_1} - x_n| \geqslant 2 - 1 = \varepsilon$. $\square$

Cette proposition implique que $\mathbb{R}$ ne peut contenir d'éléments infinitésimaux. En effet, si $\varepsilon > 0$ est infinitésimal, il doit nécessairement être plus petit que $1/n$ quel que soit $n \in \mathbb{N} \setminus \{0\}$. Donc $1/\varepsilon$ doit être plus grand que tout $n \in \mathbb{N}$ ce qui contredit l'axiome d'Archimède.

Maintenant que nous avons muni $\mathbb{R}$ d'une structure de corps commutatif, d'un ordre total et que nous avons montré que l'usage de $\lceil\cdot\rceil$ est licite, les définitions de convergence de suites et de suites de Cauchy prennent leur sens.

Lemme I.37. *Soit $(x_n)_{n\in\mathbb{N}} \in \mathscr{C}$ et $r = [(x_n)_{n\in\mathbb{N}}]$. Alors $x_n \xrightarrow[n\to\infty]{} r$ (au sens de la définition de la convergence des suites vers un nombre réel).*

Démonstration. Il faut prouver que $x_n \to r$, c'est-à-dire que

$$\forall \varepsilon \in \mathbb{R} : \varepsilon > 0,\ \exists n_0,\ \forall n \geqslant n_0,\ -\varepsilon \leqslant x_n - r \leqslant \varepsilon.$$

Soit $\varepsilon \in \mathbb{R}$, $\varepsilon > 0$. Par l'axiome d'Archimède, il existe un $p \in \mathbb{N}$ tel que $1/\varepsilon \leqslant p$ i.e., $1/p \leqslant \varepsilon$. Il suffit de montrer que

$$\exists n_0,\ \forall n \geqslant n_0,\ -1/p < x_n - r < 1/p.$$

Dans cette phrase, il faut bien se rendre compte que $x_n \in \mathbb{Q}$ et donc représente le réel $[(x_n)_{k\in\mathbb{N}}]$ où $(x_n)_{k\in\mathbb{N}}$ est la suite constante de valeur x_n. L'inégalité $x_n - r < 1/p$ signifie, grâce au lemme I.34, que

$$\exists \eta \in \mathbb{Q} : \eta > 0,\ \exists k_0 \in \mathbb{N},\ \forall k \geqslant k_0,\ 1/p - (x_n - x_k) \geqslant \eta. \tag{I.35}$$

On peut faire le même raisonnement pour l'inégalité $-1/p < x_n - r$. Vu que (x_n) est de $\mathbb{Q}$-Cauchy, on sait qu'il existe un $n_1 \in \mathbb{N}$ tel que

$$\forall m, n \geqslant n_1, \quad |x_n - x_m| \leqslant 1/(2p). \tag{I.36}$$

Prenons $n_0 := n_1$. Soit $n \geqslant n_0$. Nous allons prouver que $x_n - r < 1/p$, c'est-à-dire que (I.35) est vérifié — l'inégalité $-1/p < x_n - r$ est laissée au lecteur. Prenons $\eta := 1/(2p)$ et $k_0 := n_1$. Si $k \geqslant k_0$, (I.36) implique que

$$|x_n - x_k| \leqslant 1/(2p).$$

Dès lors on a $x_n - x_k \leqslant 1/(2p)$ ou encore $1/p - (x_n - x_k) \geqslant 1/(2p) = \eta$ ce qui est bien l'inégalité recherchée. □

Proposition I.38. *Soit $(x_n)_{n\in I} \subseteq \mathbb{Q}$ une suite de Cauchy. Alors, il existe un $r \in \mathbb{R}$ tel que $x_n \to r$.*

Démonstration. On sait que $I = \{n \in \mathbb{N} : n \geqslant n_I\}$ pour un certain $n_I \in \mathbb{N}$. Puisque $\mathbb{Q} \subseteq \mathbb{R}$, le fait que (x_n) soit de Cauchy entraîne que (x_n) soit de $\mathbb{Q}$-Cauchy (on considère moins d'ε) et donc que $(x_{n+n_I})_{n\in\mathbb{N}} \in \mathscr{C}$. Posons $r := [(x_{n+n_I})_{n\in\mathbb{N}}]$. D'après le lemme I.37, $x_{n+n_I} \to r$, c'est-à-dire $x_n \to r$. □

On peut aussi voir que $\mathbb{R}$ n'est pas « trop gros » : il consiste juste en les points qu'il faut ajouter à $\mathbb{Q}$ pour que les suites de Cauchy de ce dernier convergent.

Proposition I.39 (Densité de $\mathbb{Q}$ dans $\mathbb{R}$)**.** *Tout réel $r \in \mathbb{R}$ est limite d'une suite de rationnels $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{Q}$.*

Démonstration. C'est évident au vu du lemme I.37 car il suffit de prendre une suite $(x_n) \in r$. $\square$

Pour finir, voici le résultat qui a motivé toute la construction ci-dessus.

Théorème I.40. *$\mathbb{R}$ est complet.*

Démonstration. Soit $(x_n)_{n\in I} \subseteq \mathbb{R}$ une suite de Cauchy. Sans perte de généralité, on peut supposer $I = \mathbb{N}$ — sinon considérer $(x_{n+n_I})_{n\in\mathbb{N}}$ si $I = \{n : n \geqslant n_I\}$. Grâce à la propostion I.39, pour tout $n \in \mathbb{N}$, il existe un $x'_n \in \mathbb{Q}$ tel que

$$|x'_n - x_n| \leqslant \frac{1}{n+1}. \tag{I.37}$$

▪ Commençons par prouver que $(x'_n)_{n\in\mathbb{N}}$ est de $\mathbb{Q}$-Cauchy. Soit $\varepsilon \in \mathbb{Q}$, $\varepsilon > 0$. Il faut trouver un $n_0 \in \mathbb{N}$ tel que $\forall m, n \geqslant n_0$, $|x'_n - x_n| \leqslant \varepsilon$. Puisque (x_n) est de Cauchy, on sait qu'il existe un $n_1 \in \mathbb{N}$ tel que

$$\forall m, n \geqslant n_1, \quad |x_n - x_m| \leqslant \varepsilon/3. \tag{I.38}$$

Prenons $n_0 := \max\{n_1, \lceil 3/\varepsilon \rceil\}$. Soient $m, n \geqslant n_0$. En utilisant (I.38) et (I.37), on déduit

$$\begin{aligned} |x'_n - x'_m| &\leqslant |x'_n - x_n| + |x_n - x_m| + |x_m - x'_m| \\ &\leqslant \frac{1}{n+1} + \frac{\varepsilon}{3} + \frac{1}{m+1} \leqslant \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

▪ Posons $r := [(x'_n)_{n\in\mathbb{N}}]$ et montrons que $x_n \to r$. Nous savons par le lemme I.37 que $x'_n \to r$. Soit $\varepsilon \in \mathbb{R}$, $\varepsilon > 0$. Il faut trouver un $n_0 \in \mathbb{N}$ tel que $\forall n \geqslant n_0$, $|x_n - r| \leqslant \varepsilon$. Par définition de $x'_n \to r$, nous savons qu'il existe un $n_1 \in \mathbb{N}$ tel que

$$\forall n \geqslant n_1, \quad |x'_n - r| \leqslant \varepsilon/2. \tag{I.39}$$

Posons $n_0 := \max\{n_1, \lceil \varepsilon/2 \rceil\}$. Si $n \geqslant n_0$, on déduit de (I.39) et (I.37) que

$$\begin{aligned} |x_n - r| &\leqslant |x_n - x'_n| + |x'_n - r| \\ &\leqslant \frac{1}{n+1} + \frac{\varepsilon}{2} \leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$ $\square$

I.6 Exercices

Les exercices ci-dessous portent également sur l'utilisation des définitions en ε-n et ε-δ de la notion de limite puisque la théorie correspondante est dans le syllabus de Calculus mais que ce cours n'exerce pas ces définitions.

Exercice I.1. Montrez, à partir de la définition de convergence d'une suite, que :

(i) $\dfrac{1}{n+1} \to 0$

(ii) $\dfrac{1}{n^2} \to 0$

(iii) $\dfrac{1}{n^p} \to 0$, où $p \in \mathbb{R}^{>0}$

(iv) $\dfrac{-n}{n+5} \to -1$

(v) $\dfrac{3n+2}{5n-4} \to \dfrac{3}{5}$

(vi) $x_n \to 1$ avec $x_n = 0,\underbrace{9\ldots9}_{n \text{ fois}}$

(vii) $n^2 \xrightarrow[n\to\infty]{} +\infty$

(viii) $\ln\dfrac{1}{n} \to -\infty$

Exercice I.2. Montrez que les trois propriétés suivantes sont équivalentes :

$$\forall\varepsilon > 0,\ \exists n_0 \in \mathbb{N},\ \forall n \geqslant n_0, \quad |x_n - a| < \varepsilon \tag{I.40}$$

$$\forall\varepsilon > 0,\ \exists n_0 \in \mathbb{N},\ \forall n \geqslant n_0, \quad |x_n - a| \leqslant \varepsilon \tag{I.41}$$

$$\forall\varepsilon > 0,\ \exists n_0 \in \mathbb{N},\ \forall n \geqslant n_0, \quad |x_n - a| \leqslant \varepsilon/2 \tag{I.42}$$

Exercice I.3. Prouvez que $a^n \xrightarrow[n\to\infty]{} \infty$ lorque $a > 1$.

Exercice I.4. Soit $(x_n) \subseteq \mathbb{R}$ définie par

$$\begin{cases} x_0 = \dfrac{\sqrt{2}}{2} \\ x_{n+1} = \sqrt{\dfrac{x_n + 1}{2}} \end{cases}$$

(i) Montrez que, pour tout $n \in \mathbb{N}$, x_n peut être définie par $x_n = \cos\theta_n$ pour un unique $\theta_n \in [0, \pi/2]$.

(ii) Calculez $\lim_{n\to\infty} x_n$.

Exercice I.5. Prouvez les affirmations suivantes.

- $\sqrt[n]{a} \xrightarrow[n\to\infty]{} 1$ où $a \in\,]0, +\infty[$;
- $\dfrac{a^n}{n!} \xrightarrow[n\to\infty]{} 0$ quel que soit $a \in \mathbb{R}$;
- $n^k a^n \xrightarrow[n\to\infty]{} 0$ quel que soit $k \in \mathbb{N}$ et $|a| < 1$;

- $\sqrt[n]{n} \xrightarrow[n\to\infty]{} 1.$

Exercice I.6. Les suites suivantes sont-elles bornées ? Justifiez. (Lorsque vous répondez par l'affirmative, veuillez donner explicitement une borne.)

(i) $x_n = (-1)^n + \dfrac{2}{n+1}$

(ii) $x_n = \dfrac{4n+2}{4+3n}$

(iii) $x_n = \cos n + \sin\sqrt{n}$

(iv) $x_n = \dfrac{5^n}{n!}$

(v) $x_n = \dfrac{n!}{(2n)!}$

(vi) $x_n = \dfrac{n!}{3^n+4^n}$

(vii) $x_n = \dfrac{n^2+1}{n}$

(viii) $x_n = (\sqrt{5})^{2n}$

(ix) $x_n = \dfrac{(-2)^n}{n\,(\sqrt{7})^{2n}}$

(x) $x_0 = \dfrac{\sqrt{2}}{2}$, $x_{n+1} = \sqrt{\dfrac{x_n+1}{2}}$

(xi) $x_n = \displaystyle\sum_{k=1}^{n} \frac{1}{k^2+k}$

(xii) $x_n = \cos(n!) + \sin\bigl(\dfrac{n\pi}{2}\bigr)$

(xiii) $x_n = \dfrac{(\sqrt{n})^{2n}}{3^n}$

(xiv) $x_n = \dfrac{a^n}{n!}$ où $a \in \mathbb{R}$.

Exercice I.7. Soit $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$. On suppose que les sous-suites définies par

$$y_n = x_{2n}, \quad z_n = x_{2n+1}, \quad u_n = x_{3n}$$

convergent. Montrez que ces trois sous-suites convergent vers la même limite et que (x_n) converge aussi vers cette limite.

Exercice I.8. On dit que deux suites (x_n) et (y_n) sont *équivalentes*[6] si et seulement si

$$\lim_{n\to\infty} \frac{x_n}{y_n} = 1.$$

Ceci se note $(x_n) \sim (y_n)$. Montrez que si $y_n \to a$ et $(x_n) \sim (y_n)$ alors $x_n \to a$.

Exercice I.9 (août 2007). Soit $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$. Dites, pour chacune des affirmations suivantes, si elle est vraie ou fausse et justifiez votre réponse par un court argument ou un contre exemple.

(a) Vrai : ☐ Faux : ☐ Si (x_n) converge vers π, alors (x_n) est minorée par 0.

6. Cette définition suppose que les y_n ne s'annulent pas pour n suffisament grand.

(b) Vrai : ☐ Faux : ☐ Si (x_n) est décroissante et minorée par 0, alors (x_n) converge vers 0.

(c) Vrai : ☐ Faux : ☐ Si (x_n) converge vers 0, alors $\exists n_0 \in \mathbb{N},\ \forall n \geqslant n_0,\ x_n < \pi$.

(d) Vrai : ☐ Faux : ☐ Si (x_n) converge vers π, alors $\exists n \in \mathbb{N},\ x_n > 0$.

Exercice I.10. Soit $(x_n)_{n\in\mathbb{N}}$ définie par

$$\begin{cases} x_0 = 0, \\ x_1 = 2, \\ x_{n+1} = \frac{1}{2}(x_n + x_{n-1}) \end{cases}$$

Montrez que $(x_n)_{n\in\mathbb{N}}$ est de Cauchy en prouvant que la définition I.1 est vérifiée.

Exercice I.11. On note $\overline{\mathbb{R}}$ ou encore $[-\infty,+\infty]$ l'ensemble $\mathbb{R} \cup \{-\infty,+\infty\}$. On étend la relation d'ordre $\leqslant$ sur $\mathbb{R}$ à $[-\infty,+\infty]$ en posant

$$\forall x \in \mathbb{R}, \quad -\infty \leqslant x \leqslant +\infty \tag{I.43}$$

La définition I.8 de maximum (resp. de minimum) a d'un ensemble A s'étend de manière naturelle aux $a \in [-\infty,+\infty]$ et aux ensembles $A \subseteq [-\infty,+\infty]$. Pour $A \subseteq [-\infty,+\infty]$, on défini

$$\overline{\mathrm{Maj}}(A) := \big\{x \in [-\infty,+\infty] : \forall a \in A,\ a \leqslant x\big\}$$

- Montrez que $\overline{\mathrm{Maj}}(A) \neq \varnothing$ quel que soit $A \subseteq [-\infty,+\infty]$.
- Pour $A \subseteq \mathbb{R}$, montrez que $\sup A \in [-\infty,+\infty]$ vérifie

$$\sup A = \min\big(\overline{\mathrm{Maj}}(A)\big)$$

Exercice I.12. Soit $(x_n)_{n\in\mathbb{N}}$ une suite de nombres réels pour laquelle il existe un $c \in \,]0,1[$ tel que, pour tout $n \in \mathbb{N}$, $|x_{n+1} - x_n| \leqslant c|x_n - x_{n-1}|$. Montrez que (x_n) est de Cauchy.

Exercice I.13. Soit $(x_n)_{n\in\mathbb{N}}$ la suite définie par $x_n = 1 + \frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2^n}$. Montrez que $(x_n)_{n\in\mathbb{N}}$ est de Cauchy.

Exercice I.14. On définit la suite $(x_n)_{n\in\mathbb{N}}$ par $x_n = \sum_{k=0}^{n} \frac{1}{k!}$.

(i) Montrez que la suite $(x_n)_{n\in\mathbb{N}}$ est convergeante. On note $\mathrm{e} := \lim\limits_{n\to\infty} x_n$.

(ii) Prouvez que $\mathrm{e} = \lim\limits_{n\to\infty}\left(1+\dfrac{1}{n}\right)^n$.

Exercice I.15 (janvier 2002). (i) Montrez, à partir de la définition de limite, que

$$\frac{-n+1}{2n+5} \xrightarrow[n\to\infty]{} -\frac{1}{2}.$$

(ii) En utilisant *uniquement*[7] (i) et la définition de convergence de suites, prouvez que

$$\text{si } a_n \to a, \text{ alors} \quad a_n + \frac{-n+1}{2n+5} \xrightarrow[n\to\infty]{} a - \frac{1}{2}.$$

Exercice I.16 (janvier 2002). Soient les ensembles

$$A := \left\{\sin\left(\frac{3\pi}{2} - \frac{1}{n^2}\right) : n \in \mathbb{N}_0\right\} \qquad \text{et} \qquad B := \left\{2 + \frac{(-5)^n}{n!} : n \in \mathbb{N}_0\right\}.$$

Calculez, s'ils existent,

- $\inf A$
- $\sup A$
- $\inf B$
- $\sup B$
- $\min A$
- $\max B$

Les réponses ne suffisent pas. Justifiez-les !

Exercice I.17 (janvier 2002). Soit la suite $(v_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ définie par

$$v_0 > 1, \qquad v_{n+1} = \sqrt{3v_n - 2}.$$

Étudiez, en fonction de v_0, la convergence de (v_n) et calculez sa limite lorsqu'elle existe.

Exercice I.18 (janvier 2002). Soit $\alpha \in \mathbb{R}$ et $(x_n)_{n\in\mathbb{N}_0}$ la suite définie par

$$\forall n \in \mathbb{N}_0, \quad x_n = n^\alpha.$$

Pour quelles valeurs de α la suite (x_n) converge-t-elle et quelle est alors sa limite ? Justifiez en détail votre réponse.

7. Cela signifie que vous pouvez utiliser (i) ainsi que la définition de convergence mais que tout le reste doit être démontré.

Exercice I.19 (janvier 2003). Étudiez la convergence de la suite (x_n) définie par :

$$x_n = \sqrt{1 + \frac{(-1)^n}{n}}, \qquad n \in \mathbb{N} \setminus \{0\}.$$

Justifiez en détail. Toute affirmation non vue au cours doit être démontrée.

Exercice I.20 (janvier 2003). Soit $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$. Supposons que $(x_n)_{n\in\mathbb{N}}$ converge vers 2. Montrez, en utilisant la définition en termes de ε, que la suite $(y_n)_{n\in\mathbb{N}}$ définie par

$$y_n := -2x_n + 3$$

converge vers -1.

Exercice I.21 (janvier 2003). Étudiez la convergence de la suite $(v_n)_{n\geqslant 1}$ définie par

$$v_1 = 2, \qquad v_{n+1} = 3 - \frac{1}{v_n} \quad \text{si } n \geqslant 1.$$

Calculez sa limite, si elle existe.

Exercice I.22 (janvier 2003). Soit l'ensemble

$$E := \left\{ 2 - \frac{3^n}{(n+1)!} : n \in \mathbb{N} \right\}.$$

Calculez, s'ils existent, $\inf E$, $\min E$, $\sup E$, $\max E$. Toutes vos réponses doivent être justifiées.

Exercice I.23 (janvier 2003). Pour chacune des affirmations suivantes, cochez la case adéquate selon que vous pensez qu'elle est vraie ou fausse. Justifiez par un bref argument ou un contre-exemple. *Les résultats du cours utilisés doivent être clairement identifiés.*

(a) Vrai : ☐ Faux : ☐ Le maximum d'un ensemble fini existe toujours.

(b) Vrai : ☐ Faux : ☐ Toute suite de rationnels converge dans $\mathbb{R}$.

(c) Vrai : ☐ Faux : ☐ Si une suite est croissante, alors elle converge au sens large.

(d) Vrai : ☐ Faux : ☐ Un ensemble $A \subseteq \mathbb{R}$ est borné si et seulement si il est borné minoré et majoré.

(e) Vrai : ☐ Faux : ☐ Si $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ converge, alors elle est telle que $|x_n - x_{n+1}| \to 0$.

(f) Vrai : ☐ Faux : ☐ Il est possible que le suprémum d'un ensemble A appartienne à A.

(g) Vrai : ☐ Faux : ☐ Toute suite convergente est bornée.

(h) Vrai : ☐ Faux : ☐ Toute suite bornée est convergente.

Exercice I.24 (janvier 2002). Soit $\lambda \in \mathbb{R}$ et $(x_n)_{n\in\mathbb{N}}$ la suite définie par

$$x_n := \left(\frac{1}{\lambda - 1}\right)^n.$$

Pour quelle(s) valeur(s) de λ la suite $(x_n)_{n\in\mathbb{N}}$ converge-t-elle et quelle est alors sa limite ? Justifiez en détail.

Exercice I.25 (juin 2003). Soit $a \in \mathbb{R}$ et $(x_n)_{n\in\mathbb{N}}$ la suite définie par

$$x_n = \left(\frac{a}{5}\right)^n.$$

Pour quelle(s) valeur(s) de a la suite converge-t-elle et quelle est alors sa limite ? Justifiez en détail votre réponse.

Exercice I.26 (juin 2003). Soient $(x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}}$ deux suites convergeant respectivement vers les réels a et b. Montrez, en utilisant la définition en termes de ε que $2x_n - y_n \xrightarrow[n\to+\infty]{} 2a - b$.

Exercice I.27. Montrez que les fonctions suivantes sont continues, premièrement en utilisant la définition de limite en termes de suites et ensuite en utilisant la définition en ε-δ.

(i) $f : \mathbb{R} \to \mathbb{R} : x \mapsto \alpha x + \beta$ où $\alpha, \beta \in \mathbb{R}$;

(ii) $f : \mathbb{R} \to \mathbb{R} : x \mapsto x^2$;

(iii) $f : \mathbb{R} \to \mathbb{R} : x \mapsto x^3$;

(iv) $f : \mathbb{R} \to \mathbb{R} : x \mapsto x^2 + x$.

Exercice I.28. Montrez que la fonction $f : \mathbb{R} \to \mathbb{R} : x \mapsto \cos(x)$ est continue en utilisant la définition en ε-δ.

Exercice I.29 (juin 2003). Considérons la fonction

$$g : \mathbb{R} \to \mathbb{R} : x \mapsto \begin{cases} 2x-3 & \text{si } x \geqslant 1 \\ x+5 & \text{si } x < 1 \end{cases}$$

En utilisant la définition et ε-δ de la continuité, montrez que la fonction g est continue en 2. La fonction g est-elle continue sur $\mathbb{R}$? Justifiez en détail votre réponse.

Exercice I.30 (août 2006). Indentifiez la ou les phrases quantifiées (en cochant la case qui précède) qui *tradui(sen)t* le fait « $x+x^5 \geqslant 1$ pour x suffisament proche de 1 » :

(a) ☐ $\exists \varepsilon > 0,\ \forall x,\ x > 1-\varepsilon \Rightarrow x+x^5 \geqslant 1$

(b) ☐ $\forall \varepsilon > 0,\ \exists x,\ x > 1-\varepsilon \Rightarrow x+x^5 \geqslant 1$

(c) ☐ $\exists \varepsilon > 0,\ \forall x,\ 1-\varepsilon < x < 1+\varepsilon \Rightarrow x+x^5 \geqslant 1$

(d) ☐ $\forall \varepsilon > 0,\ \exists x,\ 1-\varepsilon < x < 1+\varepsilon \Rightarrow x+x^5 \geqslant 1$

(e) ☐ $\exists (x_n)_{n\in\mathbb{N}},\ x_n \to 1 \Rightarrow \exists n_0,\ \forall n \geqslant n_0,\ x_n + x_n^5 \geqslant 1$

(f) ☐ $\forall (x_n)_{n\in\mathbb{N}},\ x_n \to 1 \Rightarrow \exists n_0,\ \forall n \geqslant n_0,\ x_n + x_n^5 \geqslant 1$

Parmi la ou les cases (a)–(f) cochées ci-dessus, choisissez-en une et prouvez qu'elle est vraie.

Exercice I.31 (janvier 2007). Pour chacune des suites ci-dessous, calculez sa limite au sens large si elle existe. Détaillez les différentes étapes de vos calculs et énoncez les résultats que vous utilisez.

- $\left(\dfrac{n^2+\sin n}{1-n^2}\right)_{n\geqslant 2}$
- $\left(n! - 2^n\right)_{n\geqslant 0}$
- $\left(\dfrac{(-2)^n}{n}\right)_{n\geqslant 1}$

Exercice I.32 (janvier 2007). Soit $x_0 \in [1,+\infty[$. On définit la suite $(x_n)_{n\in\mathbb{N}}$ commençant à x_0 par la récurrence

$$x_{n+1} = \frac{3x_n+1}{x_n+3}$$

Montrez que $(x_n)_{n\in\mathbb{N}}$ converge et donnez la valeur de sa limite. Justifiez les différentes étapes de votre raisonnement.

Exercice I.33 (janvier 2007). Soit la suite $(x_n)_{n\in\mathbb{N}}$ définie par

$$\forall n \in \mathbb{N}, \quad x_n = n^2(\lambda - 2)^n$$

où λ est un paramètre réel. Étudiez la convergence *au sens large* de la suite $(x_n)_{n\in\mathbb{N}}$ en fonction de λ. Lorsque $(x_n)_{n\in\mathbb{N}}$ converge, donnez la valeur de sa limite. Justifiez toutes vos réponses.

Exercice I.34 (janvier 2007). Soit la suite $(x_n)_{n\in\mathbb{N}}$ définie par

$$\forall n \in \mathbb{N}, \quad x_n = \frac{5^{n+3}}{(n+1)!}$$

Dites si la suite $(x_n)_{n\in\mathbb{N}}$ vérifie ou non chacune des affirmations suivantes. Donnez pour chacune d'entre elles une preuve de votre réponse.

(a) Vrai : $\square$ Faux : $\square$ La suite (x_n) est croissante.

(b) Vrai : $\square$ Faux : $\square$ $\exists R_1, R_2 \in \mathbb{R},\ \forall n \in \mathbb{N},\ R_1 \leqslant x_n \leqslant R_2$

(c) Vrai : $\square$ Faux : $\square$ $\forall n \in \mathbb{N},\ \exists R \in \mathbb{R},\ x_n \leqslant R$

Exercice I.35 (août 2007). Soient A et B deux sous-ensembles de $\mathbb{R}$. On définit

$$A - B := \{a - b : a \in A \text{ et } b \in B\}$$

Prouvez, en utilisant la définition de votre choix, que $\sup(A-B) = \sup A - \inf B$.

Exercice I.36 (mars 2008). Calculez $\inf\{x^2 - y \mid 0 < x \leqslant 1 \text{ et } 0 \leqslant y < 1\}$ au sens large (c'est-à-dire dans $[-\infty, +\infty]$).

Chapitre II

Limite de suites de vecteurs

II.1 Normes

La seule partie de la définition de convergence qui pose problème lorsqu'on cherche à la généraliser est celle qui fait intervenir la valeur absolue. Si on se rappelle l'intention de la définition, on se rend compte que $|x_n - a|$ sert à calculer la distance entre x_n et a. Il suffit donc de proposer une définition générale de distance. Celle-ci est assez simple : une distance sur un ensemble X est une fonction $d : X \times X \to [0, +\infty[$ qui satisfait

- $\forall x_1, x_2 \in X, \quad d(x_1, x_2) = 0 \Leftrightarrow x_1 = x_2$;
- $\forall x_1, x_2 \in X, \quad d(x_1, x_2) = d(x_2, x_1)$;
- $\forall x_1, x_2, x_3 \in X, \quad d(x_1, x_2) \leqslant d(x_1, x_3) + d(x_3, x_2)$.

La troisième propriété est appelée « inégalité triangulaire » car elle dit que la longueur d'un coté d'un triangle est inférieure à la somme des longueurs des deux autres cotés (voir figure II.1). Cependant, nous ne sommes pas intéressés ici à tra-

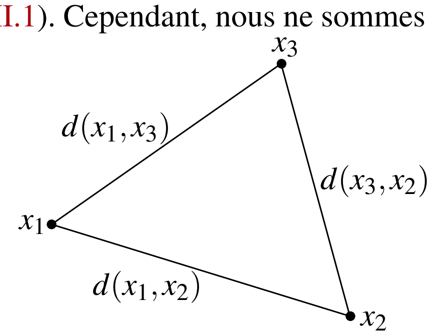


FIGURE II.1 – Inégalité triangulaire

vailler avec une distance sur un ensemble quelconque mais sur $\mathbb{R}^N$. Or la structure fondamentale de $\mathbb{R}^N$ est celle d'espace vectoriel et nous voudrions donc que la distance soit compatible avec l'addition et la multiplication scalaire. Plus précisément, nous voudrions que la distance $d : \mathbb{R}^N \times \mathbb{R}^N \to [0, +\infty[$ soit

- invariante par translation : $\forall x_1, x_2, x_3 \in \mathbb{R}^N, \quad d(x_1 + x_3, x_2 + x_3) = d(x_1, x_2)$;
- respecte les dilatations : $\forall x_1, x_2 \in X,\ \forall \lambda \in \mathbb{R}, \quad d(\lambda x_1, \lambda x_2) = |\lambda|\, d(x_1, x_2)$.

Ces deux propriétés sont assez naturelles. La première dit que la distance ne dépend pas de l'endroit où l'on se trouve dans l'espace tandis que la seconde dit qu'une dilatation ou une symétrie centrale de facteur $|\lambda|$ multitplie les distances par ce facteur $|\lambda|$ (ces faits sont illustrés à la figure II.2). Grâce à la propriété d'in-

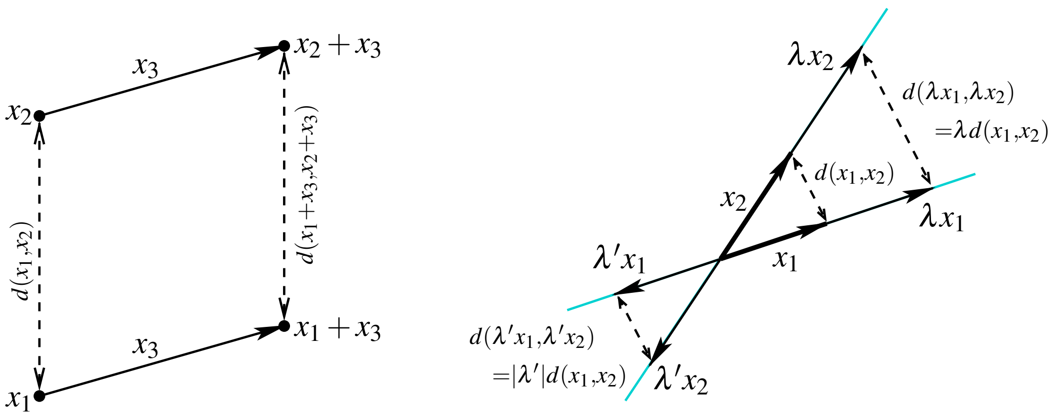


FIGURE II.2 – Comportement de la distance sous translations et dilatations

variance par translation, toutes les distances peuvent être ramenées à des distances à zéro : $d(x_1, x_2) = d(x_1 - x_2, 0)$. On en arrive ainsi à la définition de norme.

Définition II.1. Une *norme* sur $\mathbb{R}^N$ est une fonction $\|\cdot\| : \mathbb{R}^N \to [0, +\infty[: x \mapsto \|x\|$ qui possède les trois propriétés suivantes :

(i) $\forall x \in \mathbb{R}^N, \quad \|x\| = 0 \Rightarrow x = 0$;

(ii) $\forall x \in \mathbb{R}^N,\ \forall \lambda \in \mathbb{R}, \quad \|\lambda x\| = |\lambda|\, \|x\|$;

(iii) $\forall x, y \in \mathbb{R}^N, \quad \|x + y\| \leqslant \|x\| + \|y\|$.

La distance engendrée par une norme est définie par $d(x, y) := \|x - y\|$. L'implication inverse de la première propriété est vraie — c'est une conséquence de la seconde avec $\lambda = 0$. On appellera souvent la troisième propriété « inégalité triangulaire » puisqu'elle provient de celle pour les distances. La conséquence suivante de cette propriété est abondamment utilisée.

Lemme II.2. *Quel que soit* $x, y \in \mathbb{R}^N$, $\big|\|x\| - \|y\|\big| \leqslant \|x - y\|$.

Démonstration. De $\|x\| = \|x - y + y\| \leqslant \|x - y\| + \|y\|$, on déduit que $\|x\| - \|y\| \leqslant \|x - y\|$. En échangeant x et y, on a aussi que $-(\|x\| - \|y\|) \leqslant \|x - y\|$. Les deux dernières inégalités impliquent la thèse. □

Il y a trois normes que nous utiliserons tout particulièrement dans ce cours. Si $x = (x_1, \dots, x_N) \in \mathbb{R}^N$, on définit

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

À la norme $|\cdot|_1$ correspond ce qui est appelé la « taxi distance » car c'est celle qu'un taxi parcourerait dans une ville où toutes les rues sont soit horizontales soit verticales (voir figure II.3). À la norme $|\cdot|_2$ correspond la distance Euclidienne (c'est celle de la géométrie Euclidienne) qui mesure la distance entre deux points x et y par $\big(\sum_{i=1}^{N} (x_i - y_i)^2\big)^{1/2}$ (voir figure II.4). Avant d'aller plus loin, vérifions

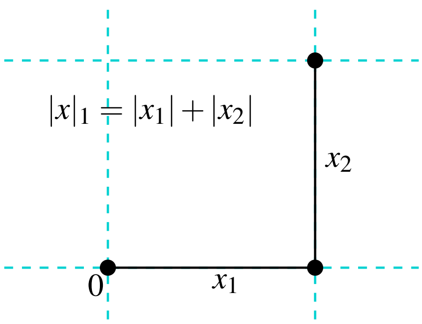


FIGURE II.3 – Taxi distance

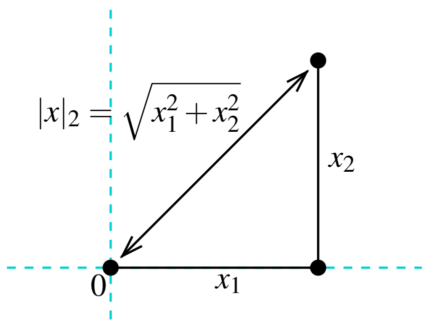


FIGURE II.4 – Distance Euclidienne

qu'on parle bien de normes.

Proposition II.3. $|\cdot|_1$ *est une norme sur* $\mathbb{R}^N$.

Démonstration. Quel que soit $x \in \mathbb{R}^N$, $|x|_1 \geqslant 0$ et donc $|\cdot|_1$ est bien une fonction de $\mathbb{R}^N$ dans $[0, +\infty[$.

- Soit $x \in \mathbb{R}^N$ tel que $|x|_1 = 0$. Pour tout $i = 1, \dots, N$, on a $|x_i| \leqslant |x|_1$. Cela implique que tous les x_i sont nuls et donc aussi x.
- Il est clair que si $x \in \mathbb{R}^N$ et $\lambda \in \mathbb{R}$, on ait $|\lambda x|_1 = |\lambda|\,|x|_1$.

- Enfin, si $x, y \in \mathbb{R}^N$, on déduit du fait que $|x_i + y_i| \leqslant |x_i| + |y_i|$ pour tout i, que $|x+y|_1 = \sum_{i=1}^N |x_i + y_i| \leqslant \sum_{i=1}^N (|x_i| + |y_i|) = |x|_1 + |y|_1$. □

Proposition II.4. $|\cdot|_\infty$ *est une norme sur* $\mathbb{R}^N$.

Démonstration. De nouveau, il est clair que $|x|_\infty \geqslant 0$ quel que soit $x \in \mathbb{R}^N$.

- Si $|x|_\infty = 0$, on déduit aisément de $|x_i| \leqslant |x|_\infty$ pour tout i, que $x = 0$.
- Soit $x \in \mathbb{R}^N$ et $\lambda \in \mathbb{R}$. Par définition, $|x|_\infty = |x_{i^*}|$ pour un certain $i^* \in \{1, \ldots, N\}$ qui vérifie $|x_{i^*}| \geqslant |x_i|$ pour tout i. On en déduit que $|\lambda x_{i^*}| = |\lambda|\,|x_{i^*}| \geqslant |\lambda|\,|x_i| = |\lambda x_i|$ pour tout i et donc que $|\lambda x|_\infty = |\lambda x_{i^*}| = |\lambda|\,|x_{i^*}| = |\lambda|\,|x|_\infty$.
- Soient $x, y \in \mathbb{R}^n$. Puisque $|x_i + y_i| \leqslant |x_i| + |y_i| \leqslant |x|_\infty + |y|_\infty$ pour tout i et $|x+y|_\infty = |x_{i^*} + y_{i^*}|$ pour un certain i^*, on a forcément que $|x+y|_\infty \leqslant |x|_\infty + |y|_\infty$. □

Avant de faire la preuve du fait que $|\cdot|_2$ est une norme, introduisons la notion de produit scalaire. Si $x, y \in \mathbb{R}^N$, on définit leur *produit scalaire* par

$$(x|y) := \sum_{i=1}^N x_i y_i. \tag{II.1}$$

Le lien avec $|\cdot|_2$ est évident :

$$\text{pour tout } x \in \mathbb{R}^N, \qquad |x|_2 = \sqrt{(x|x)}.$$

Nous supposerons que le lecteur est familier avec la notion de produit scalaire (au moins dans $\mathbb{R}^2$ et $\mathbb{R}^3$) et le fait qu'il définisse une relation d'orthogonalité : $x \perp y \Leftrightarrow (x|y) = 0$. Nous supposerons aussi connues les propriétés suivantes (qui sinon sont des exercices simples) :

(i) $\forall x \in \mathbb{R}^N \setminus \{0\}, \quad (x|x) > 0$;

(ii) $\forall x, y \in \mathbb{R}^N, \quad (x|y) = (y|x)$;

(iii) $\forall x, y \in \mathbb{R}^N$, $\forall \lambda \in \mathbb{R}, \quad (\lambda x|y) = \lambda (x|y)$;

(iv) $\forall x_1, x_2, y \in \mathbb{R}^N, \quad (x_1 + x_2|y) = (x_1|y) + (x_2|y)$.

La seconde de ces propriétés exprime la *symétrie* de $(x,y) \mapsto (x|y)$. Les deux dernières disent que, pour tout $y \in \mathbb{R}^N$ fixé, l'application $x \mapsto (x|y)$ est linéaire. Bien entendu, vu la symétrie, on a aussi que $y \mapsto (x|y)$ est une application linéaire pour tout $x \in \mathbb{R}^N$. Lorsqu'une application $\mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$ est ainsi linéraire par rapport à chacune de ses variables, on dit qu'elle est *bilinéaire*. Une application bilinéraire qui possède la propriété (i) est dite *définie positive*. Dans la suite des considérations sur le produit scalaire, nous utiliserons uniquement ces propriétés et pas la définition (II.1). Nous pourrions donc prendre un *produit scalaire* général, c'est-à-dire une application

$$(\cdot|\cdot) : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R} : (x,y) \mapsto (x|y)$$

qui est bilinéaire, symétrique et définie positive. Dans ce cas, les arguments ci-dessous montrent que la fonction

$$\|\cdot\| : \mathbb{R}^N \to \mathbb{R} : x \mapsto \|x\| := \sqrt{(x|x)}$$

est une norme. Commençons par une conséquence fondamentale de ces propriétés.

Proposition II.5 (Inégalité de Cauchy-Schwarz)**.** $\big|(x|y)\big| \leqslant |x|_2 |y|_2$ *quels que soient* $x, y \in \mathbb{R}^N$.

Démonstration. Soit $x, y \in \mathbb{R}^N$. On peut supposer que $x \neq 0$ et $y \neq 0$ — sinon la thèse est évidente. Le fait que le produit scalaire soit défini positif implique que $(tx+y|tx+y) \geqslant 0$ quel que soit $t \in \mathbb{R}$. En développant cette inégalité grâce à la bilinéarité et la symétrie du produit scalaire, on trouve qu'elle est équivalente à

$$\forall t \in \mathbb{R}, \qquad t^2(x|x) + 2t(x|y) + (y|y) \geqslant 0.$$

Le membre de gauche de l'inégalité est un polynôme du second degré en t (vu que $(x|x) \neq 0$). Le fait qu'il soit positif pour tout t implique que son discriminant soit $\leqslant 0$, c'est-à-dire que $(x|y)^2 - (x|x)(y|y) \leqslant 0$ En faisant passer $(x|x)(y|y)$ dans le membre de droite et en prenant la racine carrée des deux membres, on trouve l'inégalite de Cauchy-Schwarz. □

Proposition II.6. $|\cdot|_2$ *est une norme sur* $\mathbb{R}^N$.

Démonstration. Comme précédemment, il est évident que $\forall x \in \mathbb{R}^N,\ |x|_2 \geqslant 0$ et donc que $|\cdot|_2$ est une fonction de $\mathbb{R}^N$ dans $[0, +\infty[$.

- Supposons $|x|_2 = 0$. De $x_i^2 \leqslant |x|_2^2$ pour tout i, on déduit aisément que $x_i = 0$ quel que soit i et donc que $x = 0$.

- Soit $x \in \mathbb{R}^N$ et $\lambda \in \mathbb{R}$. On a $|\lambda x|_2 = \sqrt{\sum_{i=1}^N (\lambda x_i)^2} = \sqrt{\lambda^2 \sum_{i=1}^N x_i^2} = |\lambda|\,|x|_2$.
- Soit $x, y \in \mathbb{R}^N$. En utilisant l'inégalité de Cauchy-Schwarz, on a

$$\begin{aligned} |x+y|_2^2 &= (x+y|x+y) \\ &= (x|x) + 2(x|y) + (y|y) \\ &\leqslant |x|_2^2 + 2|x|_2|y|_2 + |y|_2^2 = (|x|_2 + |y|_2)^2 \end{aligned}$$

Il suffit alors de prendre la racine carrée des deux membres pour avoir l'inégalité triangulaire. □

À une norme sont naturellement associées des boules qui sont les ensembles des points situés au plus à une certaine distance d'un point donné.

Définition II.7. Soit $\|\cdot\|$ une norme sur $\mathbb{R}^N$. La boule ouverte (resp. fermée) de centre $x \in \mathbb{R}^N$ et de rayon $r \in [0, +\infty[$ est l'ensemble

$$\mathrm{B}_{\|\cdot\|}(x,r) := \{y \in \mathbb{R}^N : \|y-x\| < r\} \qquad (\text{resp. } \mathrm{B}_{\|\cdot\|}[x,r] := \{y \in \mathbb{R}^N : \|y-x\| \leqslant r\}).$$

Lorsque la norme considérée est connue d'après le contexte ou qu'elle n'est pas importante, on se contentera d'écrire $\mathrm{B}(x,r)$ et $\mathrm{B}[x,r]$. La seule différence entre $\mathrm{B}_{\|\cdot\|}(x,r)$ et $\mathrm{B}_{\|\cdot\|}[x,r]$ est que $\mathrm{B}_{\|\cdot\|}[x,r]$ inclut les points y vérifiant $\|y-x\| = r$, c'est-à-dire la frontière de la boule. Toutes les boules peuvent s'obtenir comme dilatation et translation de la boule de rayon unité centrée à l'origine : $\mathrm{B}_{\|\cdot\|}(x,r) = x + r\mathrm{B}_{\|\cdot\|}(0,1)$ et $\mathrm{B}_{\|\cdot\|}[x,r] = x + r\mathrm{B}_{\|\cdot\|}[0,1]$ (voir figure II.5). Il suffit donc de tra-

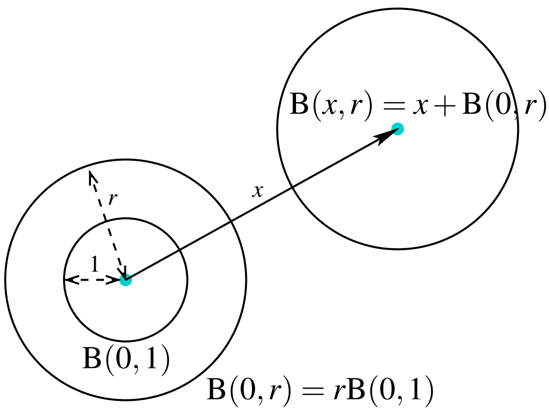


FIGURE II.5 – Translation et dilatation de la boule unité

cer $\mathrm{B}_{\|\cdot\|}(0,1)$ ou $\mathrm{B}_{\|\cdot\|}[0,1]$ pour se rendre compte de la forme de toutes les boules.

Les boules unité ouvertes pour les trois normes $|\cdot|_1$, $|\cdot|_2$ et $|\cdot|_\infty$ sont tracées à la figure II.6 (pouvez-vous expliquer comment construire ces graphiques ?). Les in-

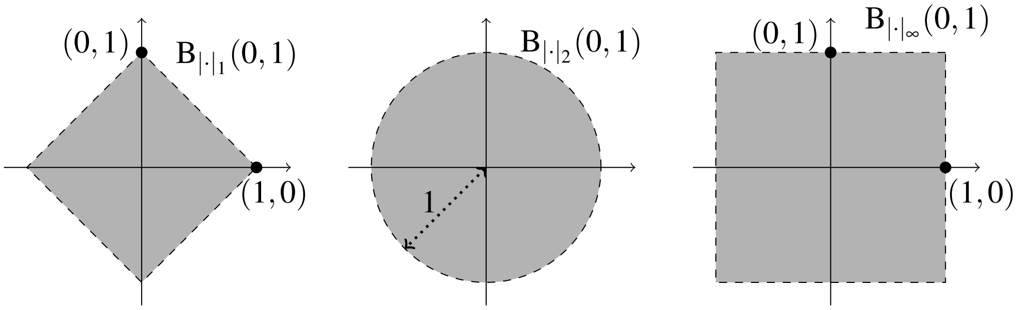


FIGURE II.6 – Boules unité ouvertes

clusions des boules pour une norme dans les boules pour une autre norme traduisent des relations entre ces normes. Nous nous intéresserons uniquement ici à l'équivalence de normes.

Définition II.8. Soit $\|\cdot\|$ et $\|\cdot\|'$ deux normes sur $\mathbb{R}^N$. On dit que $\|\cdot\|$ est *équivalente* à $\|\cdot\|'$ si une des deux propriétés équivalentes suivantes est vérifiée :

(i) $\exists R, R' \in \left]0,+\infty\right[,\ \forall x \in \mathbb{R}^N,\ R\|x\| \leqslant \|x\|' \leqslant R'\|x\|$

(ii) $\exists R, R' \in \left]0,+\infty\right[,\ \mathrm{B}_{\|\cdot\|'}(0,R) \subseteq \mathrm{B}_{\|\cdot\|}(0,1) \subseteq \mathrm{B}_{\|\cdot\|'}(0,R')$

L'équivalence de ces deux définitions est assez facile à montrer — et nous verrons que les quantités R et R' de (i) et (ii) sont les mêmes. Tout d'abord, une inégalité du type $\|x\|' \leqslant R'\|x\|$ implique que $\|x\| < 1 \Rightarrow \|x\|' < R'$ et donc que $\mathrm{B}_{\|\cdot\|}(0,1) \subseteq \mathrm{B}_{\|\cdot\|'(0,R')}$. En utilisant l'inégalité $R\|x\| \leqslant \|x\|'$ de manière similaire, on trouve que $\mathrm{B}_{\|\cdot\|'}(0,R) \subseteq \mathrm{B}_{\|\cdot\|}(0,1)$. Ceci prouve que (i) $\Rightarrow$ (ii). Pour (ii) $\Rightarrow$ (i) il suffit de « retourner » le raisonnement. Supposons que $\mathrm{B}_{\|\cdot\|}(0,1) \subseteq \mathrm{B}_{\|\cdot\|'}(0,R')$. Soit $x \neq 0$ et $\varepsilon \in \left]0,1\right[$. Puisque $(1-\varepsilon)x/\|x\| \in \mathrm{B}_{\|\cdot\|}(0,1)$, il s'ensuit que $(1-\varepsilon)x/\|x\| \in \mathrm{B}_{\|\cdot\|'}(0,R')$ et donc que $(1-\varepsilon)\|x\|' \leqslant R'\|x\|$. Comme c'est vrai pour tout $\varepsilon \in \left]0,1\right[$, on peut passer à la limite $\varepsilon \to 0$ et on obtient $\|x\| \leqslant R'\|x\|$. On a établi cette inégalité pour $x \neq 0$ mais elle est bien sûr valable pour $x = 0$. En procédant de la même manière, on déduit de $\mathrm{B}_{\|\cdot\|'}(0,R) \subseteq \mathrm{B}_{\|\cdot\|}(0,1)$ que $R\|x\| \leqslant \|x\|'$ pour tout $x \in \mathbb{R}^N$. Ceci finit de montrer que (ii) $\Leftrightarrow$ (i).

Comme son nom l'indique, la relation « être équivalent » sur l'ensemble des normes sur $\mathbb{R}^N$ est bien une relation d'équivalence. Nous laissons au lecteur le soin de prouver qu'elle est en effet réflexive, symétrique et transitive.

À la vue des dessins des boules unités (figure II.6), il apparaît immédiatement que $\mathrm{B}_{|\cdot|_1}(0,1) \subseteq \mathrm{B}_{|\cdot|_2}(0,1) \subseteq \mathrm{B}_{|\cdot|_\infty}(0,1) \subseteq \mathrm{B}_{|\cdot|_1}(0,2)$ (voir la figure II.7). Ces inclu-

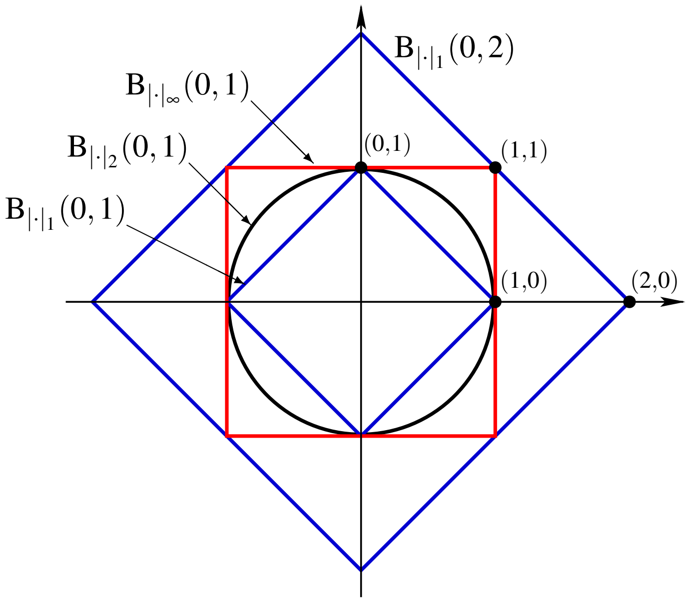


FIGURE II.7 – Inclusion des boules unité

sions traduisent certaines inégalités entre les normes. Le facteur 2 est particulier à la dimension deux. Plus généralement, on a :

Proposition II.9. *Pour tout* $x \in \mathbb{R}^N$, $|x|_\infty \leqslant |x|_2 \leqslant |x|_1 \leqslant N|x|_\infty$. *En particulier, les trois normes* $|\cdot|_\infty$, $|\cdot|_2$ *et* $|\cdot|_1$ *sont équivalentes.*

Démonstration. Les fait que les trois inégalités impliquent que les trois normes sont équivalentes deux à deux est un exercice simple laissé au lecteur. Soit $x = (x_1,\dots,x_N) \in \mathbb{R}^N$.

▪ Montrons $|x|_\infty \leqslant |x|_2$. Par définition, $|x|_\infty = |x_{i^*}|$ pour un certain $i^* \in \{1,\dots,N\}$. Comme $|x_{i^*}|^2 \leqslant \sum_{i=1}^N |x_i|^2$, en prenant la racine carrée des deux membres, on conclut que $|x|_\infty = |x_{i^*}| \leqslant |x|_2$.

▪ Passons à $|x|_2 \leqslant |x|_1$. Puisque ce sont des quantités positives, il est équivalent de montrer que $|x|_2^2 \leqslant |x|_1^2$. On voit clairement que cette inégalité est vraie en développant le carré de la somme :

$$|x|_2^2 = \sum_{i=1}^N |x_i|^2 \leqslant |x|_1^2 = \Big(\sum_{i=1}^N |x_i|\Big)^2 = \sum_{i=1}^N |x_i|^2 + 2 \sum_{\substack{i,j \\ 1 \leqslant i < j \leqslant N}} |x_i|\,|x_j|.$$

▪ Terminons avec $|x|_1 \leqslant N|x|_\infty$. En utilisant le fait que $|x|_\infty$ est plus grand que toutes les composantes $|x_i|$, on déduit $|x|_1 = \sum_{i=1}^N |x_i| \leqslant \sum_{i=1}^N |x|_\infty = N|x|_\infty$. □

Remarque II.10. On dira d'une inégalité du type $\forall x \in \mathbb{R}^N,\ \|x\| \leqslant C\|x\|'$ entre deux normes $\|\cdot\|$ et $\|\cdot\|'$ est *optimale* si le C qui y figure est le plus petit possible. Ce C vaut $\sup\{\|x\| : \|x\|' = 1\}$. Une fois que nous aurons vu la notion de compacité, nous pourrons montrer (en dimension finie) que l'optimalité est équivalente à l'existence d'un x^* tel que $\|x^*\| = C\|x^*\|'$. Au sens ci-dessus, les trois inégalités de la proposition précédente sont optimales. Géométriquement, cela correspond au fait que les frontières des boules se « touchent » (voir figure II.7). Par contre, l'inégalité $|x|_2 \leqslant N|x|_\infty$ qu'on peut déduire de ces trois inégalités n'est pas optimale. L'inégalité optimale est $|x|_2 \leqslant \sqrt{N}|x|_\infty$. (Pouvez-vous la prouver et en donner l'interprétation géométrique ?)

Il y a d'autres normes sur $\mathbb{R}^N$ que les trois que nous avons présentées ci-dessus. Par exemple, pour $p \in]0, +\infty[$, on peut définir

$$|x|_p := \Bigl(\sum_{i=1}^N |x_i|^p\Bigr)^{1/p} \qquad \text{pour tout } x \in \mathbb{R}^N.$$

On peut montrer (voir l'exercice II.18) que, si $p \in [1, +\infty[$, $|\cdot|_p$ est une norme. Ce n'est pas le cas si $p \in]0, 1[$ (exercice II.14). Comme cas particuliers, on retrouve $|\cdot|_1$ et $|\cdot|_2$. En ce qui concerne $|\cdot|_\infty$, il est facile de voir que, pour tout $x \in \mathbb{R}^N$, $|x|_\infty \leqslant |x|_p \leqslant N^{1/p}|x|_\infty$ et donc, en passant à la limite $p \to \infty$, que $|x|_\infty = \lim_{p\to\infty} |x|_p$. Ceci explique la notation $|\cdot|_\infty$. Il y a encore bien d'autres normes possibles sur $\mathbb{R}^N$. Une question naturelle — et particulièrement importante pour la convergence, voir section II — est de savoir si elles sont équivalentes à celles connues ou non. C'est clairement le cas pour $|\cdot|_p$ puisque, comme on l'a dit, $|x|_\infty \leqslant |x|_p \leqslant N^{1/p}|x|_\infty$. En fait, c'est toujours le cas en dimension finie.

Théorème II.11. *Toutes les normes sur $\mathbb{R}^N$ sont équivalentes.*

À ce stade, nous n'avons pas encore les outils nécessaires pour prouver ce théorème. Nous y reviendrons dans le chapitre traitant de la compacité (page 93).

II.2 Convergence des suites vectorielles

Dans cette section, nous allons employer la notion de norme que nous venons de définir pour généraliser le concept de convergence à $\mathbb{R}^N$. Nous verrons que les propriétés prouvées en dimension 1 passent en général à plusieurs dimensions. Comme leurs preuves consistent essentiellement en un simple recopiage de celles en dimension 1 à celà prêt qu'on a remplacé la valeur absolue par une norme, elles seront la plupart du temps laissées au lecteur (qui y verra l'opportunité de tester sa compréhension en faisant les adaptations nécessaires).

Tout d'abord, pour éviter toute confusion, donnons explicitement la définition d'une suite dans $\mathbb{R}^N$.

Définition II.12. Une *suite dans* $\mathbb{R}^N$ est une application $I \to \mathbb{R}^N : n \mapsto x_n$ où $I = \{n \in \mathbb{N} : n \geqslant n_0\}$ pour un certain $n_0 \in \mathbb{N}$. On emploie les notations $(x_n)_{n\in I} \subseteq \mathbb{R}^N$ ou $(x_n : n \in I) \subseteq \mathbb{R}^N$, ou encore, $(x_n) \subseteq \mathbb{R}^N$ ou (x_n) si le contexte supplée aux éléments manquants.

Ayant en notre possession le concept de norme, il est facile de généraliser la définition de convergence.

Définition II.13. Soit $(x_n)_{n\in I} \subseteq \mathbb{R}^N$, $a \in \mathbb{R}^N$ et $\|\cdot\|$ une norme sur $\mathbb{R}^N$. On dit que la suite $(x_n)_{n\in I}$ *converge vers* a *au sens de* $\|\cdot\|$ si

$$\forall \varepsilon > 0,\ \exists n_0 \in \mathbb{N},\ \forall n \in I, \quad n \geqslant n_0 \Rightarrow \|x_n - a\| \leqslant \varepsilon$$

ou, de manière équivalente, [1] si $\|x_n - a\| \xrightarrow[n\to\infty]{\mathbb{R}} 0$. Dans ce cas, on note « $x_n \xrightarrow{\|\cdot\|} a$ si $n \to \infty$ », « $x_n \xrightarrow[n\to\infty]{\|\cdot\|} a$ » ou simplement « $x_n \xrightarrow{\|\cdot\|} a$ ». On appelle a une limite de $(x_n)_{n\in I}$.

A priori, il faut donc faire attention. Une suite pourrait converger pour une norme et pas pour une autre. Le résultat suivant dit quand deux normes induisent le même type de convergence.

Proposition II.14. *Soient* $\|\cdot\|$ *et* $\|\cdot\|'$ *deux normes sur* $\mathbb{R}^N$. *Si* $\|\cdot\|$ *et* $\|\cdot\|'$ *sont équivalentes, alors quels que soient* $(x_n)_{n\in I} \subseteq \mathbb{R}^N$ *et* $a \in \mathbb{R}^N$, *on a*

$$x_n \xrightarrow[n\to\infty]{\|\cdot\|} a \quad \Leftrightarrow \quad x_n \xrightarrow[n\to\infty]{\|\cdot\|'} a.$$

Démonstration. L'équivalence des normes implique qu'il existe une constante $c \in]0, +\infty[$ telle que $\forall x \in \mathbb{R}^N,\ \|x\|' \leqslant C\|x\|$. On va montrer que de cela on peut déduire l'implication « $\Rightarrow$ ». On suppose donc que $\|x_n - a\| \to 0$ et on veut établir que

1. Voyez-vous pourquoi c'est équivalent ?

$\|x_n - a\|' \to 0$. Vu que $\|x_n - a\|' \leqslant C\|x_n - a\|$ pour tout $n \in I$, c'est une simple application de la convergence dominée. L'implication inverse résulte de l'autre inégalité dans la définition de normes équivalentes. □

Étant donné que toutes les normes sont équivalentes dans $\mathbb{R}^N$ (théorème II.11), toutes les définitions de convergence sont équivalentes. On peut donc écrire

$$x_n \xrightarrow[n\to\infty]{} a$$

sans risque de confusion et choisir notre norme favorite lorsqu'on a besoin de l'écrire en termes de ε.

L'unicité de la limite se prouve comme dans le cas unidimensionnel (faites le !). On peut donc employer la notation

$$\lim_{n\to\infty} x_n$$

pour nommer la valeur de la limite de (x_n) — lorsque celle-ci existe.

La somme et la multiplication par un scalaire d'une suite se définissent de manière similaire au cas unidimensionnel et on a l'analogue des règles de calcul.

Proposition II.15. *Soient $(x_n)_{n\in I}$, $(y_n)_{n\in J}$ deux suites de $\mathbb{R}^N$ et $(\lambda_n)_{n\in K}$ une suite de $\mathbb{R}$ telles que $x_n \to a$, $y_n \to b$ et $\lambda_n \to \lambda$ pour certains $a, b \in \mathbb{R}^N$ et $\lambda \in \mathbb{R}$. Alors $x_n + y_n \to a + b$ et $\lambda_n x_n \to \lambda a$. En particulier $\lambda x_n \to \lambda a$ pour tout $\lambda \in \mathbb{R}$.*

Démonstration. Adapter la preuve du résultat à une dimension. □

Comme on n'a pas d'ordre sur $\mathbb{R}^N$, on ne peut évidemment pas généraliser le théorème du sandwich. Cependant, la convergence dominée, elle, passe sans problèmes à $\mathbb{R}^N$. En effet, puisque $x_n \to a$ est équivalent au fait que la suite $\big(\|x_n - a\|\big)_n \subseteq \mathbb{R}$ tende vers zéro, on peut utiliser tous les outils développés pour les suites de nombres réels afin d'établir cette convergence. C'est aussi pratique lorsque $x_n \to a$ est en hypothèse. Par exemple, si $(x_n)_{n\in I} \subseteq \mathbb{R}^N$ et $x \in \mathbb{R}^N$, on a

$$x_n \xrightarrow[n\to\infty]{} x \Rightarrow \|x_n\| \xrightarrow[n\to\infty]{} \|x\|. \tag{II.2}$$

Cela découle simplement de $\big|\|x_n\| - \|x\|\big| \leqslant \|x_n - x\| \to 0$. Il y a une autre manière par laquelle la convergence dans $\mathbb{R}^N$ se réduit à la convergence dans $\mathbb{R}$.

Proposition II.16 (Convergence composante par composante)**.** *Soit $(x_n)_{n\in I}$ une suite de $\mathbb{R}^N$ et $a \in \mathbb{R}^N$. En détaillant leurs composantes, on écrit $x_n = (x_{n,1}, x_{n,2}, x_{n,3}, \dots, x_{n,N})$ et $a = (a_1, a_2, a_3, \dots, a_N)$. On a*

$$x_n \xrightarrow[n\to\infty]{} a \quad \Leftrightarrow \quad \forall i = 1, \dots, N,\ x_{n,i} \xrightarrow[n\to\infty]{} a_i.$$

Démonstration. Condition suffisante. Comme toutes les normes sont équivalentes, nous pouvons choisir celle qui nous convient le mieux. On va établir que $x_n \xrightarrow{|\cdot|_1} a$, c'est-à-dire $|x_n - a|_1 \to 0$. Comme $|x_n - a|_1 = \sum_{i=1}^N |x_{n,i} - a_i|$ et que par hypothèse $|x_{n,i} - a_i| \to 0$ pour tout i, les règles de calcul sur les limites impliquent que $|x_n - a|_1 \to 0$.

Condition nécessaire. Puisque, pour tout i, $|x_{n,i} - a_i| \leqslant |x_n - a|_1 \to 0$, il découle de la convergence dominée que $x_{n,i} \xrightarrow[n\to\infty]{} a_i$. □

On définit une sous-suite d'une suite de $\mathbb{R}^N$ comme pour le cas de $\mathbb{R}$. Les propriétés associées subsistent telles quelles (leurs démonstrations sont facilement adaptées, faites les !).

On peut aussi définir la convergence au sens large pour les suites de $\mathbb{R}^N$. Bien sûr, comme il n'y a pas d'ordre sur $\mathbb{R}^N$, il n'est pas question de parler de $+\infty$ et de $-\infty$. Néanmoins, on peut dire qu'une suite tend vers l'infini quand elle finit par quitter n'importe quelle boule.

Définition II.17. Soit $(x_n)_{n\in I} \subseteq \mathbb{R}^N$ et $\|\cdot\|$ une norme sur $\mathbb{R}^N$. On dit que $(x_n)_{n\in I}$ *converge vers l'infini pour la norme* $\|\cdot\|$ si

$$\forall \rho \in \mathbb{R},\ \exists n_0 \in \mathbb{N},\ \forall n \geqslant n_0,\ \|x_n\| \geqslant \rho$$

ou, de manière équivalente, si $\|x_n\| \xrightarrow[n\to\infty]{} +\infty$. Dans ce cas, on notera $x_n \xrightarrow[n\to\infty]{\|\cdot\|} \infty$.

De nouveau, l'équivalence des normes sur $\mathbb{R}^N$ implique que le fait de tendre vers l'infini ne dépend pas de la norme — les définitions pour différentes normes sont équivalentes. Notez que, pour $N = 1$, « tendre vers l'infini » n'est *pas* équivalent à « tendre vers $+\infty$ ou vers $-\infty$ ». En effet, $(x_n) = \big((-1)^n n\big)_{n\in\mathbb{N}}$ tend vers l'infini puisque $|x_n| = n \to +\infty$ mais ne tend ni vers $+\infty$, ni vers $-\infty$. Par contre, l'implication inverse est vraie : si $x_n \to +\infty$ ou $x_n \to -\infty$, alors (x_n) tend vers l'infini ($|x_n| \to +\infty$). Comme il y a risque de confusion entre $x_n \to \infty$ et $x_n \to \pm\infty$ dans $\mathbb{R}$, on tâchera toujours d'être précis. On laisse au lecteur le soin d'établir la proposition suivante.

Proposition II.18. *Soit* $(x_n)_{n\in I} = \big((x_{n,1},\dots,x_{n,N})\big)_{n\in I} \subseteq \mathbb{R}^N$*. On a*

$$\exists i = 1,\dots,N,\ |x_{n,i}| \xrightarrow[n\to\infty]{} +\infty \quad \Rightarrow \quad x_n \xrightarrow[n\to\infty]{} \infty.$$

Parce qu'on ne dispose pas du signe de l'infini, peu de propriétés subsistent. En voici quelques unes :

Proposition II.19. *Soit* $(x_n)_{n\in I} \subseteq \mathbb{R}^N$ *et* $(\lambda_n)_{n\in J} \subseteq \mathbb{R}$*.*

(i) Si $x_n \to \infty$ *et* $\exists \varepsilon > 0$, $\exists n^* \in \mathbb{N}$, $\forall n \geqslant n^*$, $|\lambda_n| \geqslant \varepsilon$*, alors* $\lambda_n x_n \to \infty$*.*

(ii) Si $\exists \varepsilon > 0$, $\exists n^* \in \mathbb{N}$, $\forall n \geqslant n^*$, $\|x_n\| \geqslant \varepsilon$ *et* $|\lambda_n| \to +\infty$*, alors* $\lambda_n x_n \to \infty$*.*

Démonstration. Ces propriétes découlent directement des règles de calcul pour les limites infinies et de la définition II.17. □

Enfin, examinons le critère de Cauchy et la complétion de $\mathbb{R}^N$. La fait d'être de Cauchy se définit comme sur $\mathbb{R}$.

Définition II.20. Soit $(x_n)_{n\in I} \subseteq \mathbb{R}^N$ et $\|\cdot\|$ une norme sur $\mathbb{R}^N$. La suite $(x_n)_{n\in I}$ est dite *de Cauchy pour la norme* $\|\cdot\|$ si

$$\forall \varepsilon > 0,\ \exists n_0 \in \mathbb{N},\ \forall m,n \in I,\quad (m \geqslant n_0 \wedge n \geqslant n_0) \Rightarrow \|x_m - x_n\| \leqslant \varepsilon.$$

Encore une fois, l'équivalence de toutes les normes fait qu'être de Cauchy au sens d'une norme ou au sens d'une autre est équivalent (pouvez-vous écrire les détails ?). Nous pouvons donc dire « être de Cauchy » sans risque de confusion. Le lecteur prouvera sans peine l'analogue suivant de la proposition I.2.

Proposition II.21. *Toute suite convergente de* $\mathbb{R}^N$ *est de Cauchy.*

Terminons en montrant que la complétude de $\mathbb{R}$ « se transmet » à $\mathbb{R}^N$.

Théorème II.22. $\mathbb{R}^N$ *est complet.*

Démonstration. Soit $(x_n)_{n\in I} \subseteq \mathbb{R}^N$ une suite de Cauchy. Puisqu'on peut choisir la norme, travaillons avec $|\cdot|_1$. Détaillons les composantes de x_n : $x_n = (x_{n,1}, x_{n,2}, \dots, x_{n,N})$. Puisque, pour tout i, $|x_{m,i} - x_{n,i}| \leqslant |x_m - x_n|_1$, il est aisé de montrer que les suites $(x_{n,i})_{n\in I}$, $i = 1,\dots,N$, sont toutes de Cauchy dans $\mathbb{R}$. Grâce à la complétude de $\mathbb{R}$, elles convergent :

$$\forall i = 1,\dots,N,\quad x_{n,i} \xrightarrow[n\to\infty]{} x_i^*$$

pour certains $x_1^*,\dots,x_N^* \in \mathbb{R}$. Posons $x^* = (x_1^*,\dots,x_N^*) \in \mathbb{R}^N$. La proposition II.16 implique que $x_n \to x^*$. □

II.3 Exercices

Exercice II.1. Soient $x = (1,2,3)$ et $y = (4,5,6)$ deux vecteurs de $\mathbb{R}^3$. Calculez $|x|_1$, $|y|_1$, $|x|_2$, $|y|_2$, $|x|_\infty$, $|y|_\infty$ ainsi que la distance entre x et y pour chacune de ces trois normes.

Exercice II.2. Soit $x = (1,2,\dots,N) \in \mathbb{R}^N$ et $y = (1,1,\dots,1) \in \mathbb{R}^N$. Calculez $|x|_p$ et $|y|_p$ pour $p \in \{1,2,\infty\}$.

Exercice II.3. Rappelons que $\mathbb{R}^{2\times 2} = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a,b,c,d \in \mathbb{R} \right\}$. Pour $A \in \mathbb{R}^{2\times 2}$, posons $\|A\| = \max\{|a|+|b|, |c|+|d|\}$. Montrer que $\|\cdot\|$ est une norme sur $\mathbb{R}^{2\times 2}$.

Exercice II.4. Montrez que $\|\cdot\|$ est une norme sur $\mathbb{R}$ si et seulement s'il existe une constante $c \in \mathbb{R}^{>0}$ telle que $\forall x \in \mathbb{R},\ \|x\| = c|x|$.

Exercice II.5. Rappelons que $\mathbb{P}_2$ désigne l'ensemble des fonctions polynomiales de degré 2. Pour $P \in \mathbb{P}_2$, on définit $\|P\| := |a_0| + |a_1| + |a_2|$ où a_0, a_1, a_2 sont les coefficients de P i.e., $P = a_0 + a_1 x + a_2 x^2$. $P \mapsto \|P\|$ est-il une norme sur $\mathbb{P}_2$?

Exercice II.6. Prouvez la convergence des suites ci-dessous grâce à la définition II.13.

- $x_n = \left(\dfrac{n+1}{n}, \dfrac{2n+1}{n} \right)$
- $y_n = \left(\sqrt{\dfrac{1}{n}}, \sqrt{\dfrac{1}{n^3}} \right)$

Exercice II.7. Étudiez la convergence des suites suivantes :

- $x_n = \left(\dfrac{n}{n+1}, \dfrac{2^n}{n!} \right)$
- $y_n = \left(\dfrac{n^4}{3+2n^4}, \sqrt[n]{n}, \dfrac{n^5}{3^n} \right)$
- $z_n = \left(\dfrac{1}{2}, \dfrac{(-1)^n}{n} \right)$

Exercice II.8. Étudiez la convergence des suites de nombres complexes ci-dessous :

(i) $x_n = \dfrac{n+\boldsymbol{i}}{n+1}$ (ii) $x_n = 2^n + \dfrac{\boldsymbol{i}}{n}$

(iii) $x_n = \dfrac{\sqrt{n}}{n+1} + \dfrac{\sqrt{n}+2}{1-\boldsymbol{i}\sqrt{n}}$

(iv) $x_n = \dfrac{\boldsymbol{i}^n}{n}$

(v) $x_n = \dfrac{(1-\boldsymbol{i})^{2n}}{n3^n}$

(vi) $x_n = \dfrac{\sqrt{2}}{2}\boldsymbol{i} + \dfrac{n}{n+1}\boldsymbol{i}^n$

(vii) $x_n = \dfrac{n + (\sqrt{3}\boldsymbol{i}+1)^n}{2^n}$

(viii) $x_n = \dfrac{(n+3)(\sqrt{3}-\boldsymbol{i})^n}{(2n^2+1)2^n}$

Exercice II.9. Soit $\Omega \subseteq \mathbb{R}$. On définit l'ensemble des fonctions bornées de Ω vers $\mathbb{R}$ comme

$$\mathscr{B}(\Omega;\mathbb{R}^M) := \big\{f : \Omega \to \mathbb{R} : \exists R \in \mathbb{R},\ \forall x \in \Omega,\ |f(x)| \leqslant R\big\}$$

Montrez que $\mathscr{B}$ (muni de l'addition et de la multiplication usuelles) est un espace vectoriel. Pour $f \in \mathscr{B}$, on définit $|f|_\infty := \sup_{x\in\Omega}|f(x)|$. Montrez que $|\cdot|_\infty$ est une norme sur $\mathscr{B}$. Calculer la norme de chacune des fonctions suivantes :

- $f(x) = \sin x$ sur $\Omega = [-\pi/2, \pi/2]$;
- $g(x) = |x-1|+1$ sur $\Omega = [-1,2]$;
- $h(x) = \mathrm{e}^{x+1}$ sur $\Omega = [-2,0]$.

Montrez que l'espace $\mathscr{B}$ muni de $|\cdot|_\infty$ est complet (indication : si (f_n) est de Cauchy dans $\mathscr{B}$, alors, quel que soit $x \in \Omega$, $\big(f_n(x)\big)$ est de Cauchy dans $\mathbb{R}$).

Exercice II.10. Montrez que les applications linéraires $L : \mathbb{R}^N \to \mathbb{R}^M$ sont continues.

Exercice II.11 (janvier 2002). Étudiez la convergence des suites ci-dessous et calculez leur limite, si elle existe. Justifiez vos réponses.

- $x_n := 2 + \left(\dfrac{2\boldsymbol{i}-1}{7}\right)^n$
- $y_n := \left(\dfrac{\pi^n}{n!}, \dfrac{(2n+1)^7(n^2+4)}{2n^3(4n^2+3)^3}, \dfrac{\sin n}{n^2}\right)$

Exercice II.12 (janvier 2002). Soient deux normes $\|\bullet\|_1$ et $\|\bullet\|_2$ définies sur $\mathbb{R}^N$. Supposons que $\|\bullet\|_1$ et $\|\bullet\|_2$ sont équivalentes. Montrez que, si une suite $(x_n)_{n\in\mathbb{N}}$ est de Cauchy pour $\|\bullet\|_1$, alors elle l'est également pour $\|\bullet\|_2$.

Exercice II.13 (mars 2003). Soit $a \in \mathbb{R}^N$, $r > 0$, $r' > 0$ et $\|\cdot\|$ une norme sur $\mathbb{R}^N$.

Complétez les égalités et équivalences suivantes afin qu'elles soient vraies.

$$\mathrm{B}_{\|\cdot\|}(a,r) = \{x \qquad : \qquad \}$$

$$\mathrm{B}_{\|\cdot\|}[a,r'] = \{x \qquad : \qquad \}$$

$$x \in \mathrm{B}_{\|\cdot\|}(a,r) \Leftrightarrow$$

$$x \in \mathrm{B}_{\|\cdot\|}[a,r'] \Leftrightarrow$$

Soit $A \subseteq \mathbb{R}^N$. Considérons les deux propriétés suivantes :

(i) $\exists r > 0,\ \mathrm{B}_{\|\cdot\|}(0,r) \subseteq A$

(ii) $\exists r' > 0,\ \mathrm{B}_{\|\cdot\|}[0,r'] \subseteq A$

Montrez, en détaillant votre raisonnement, que (i) $\Leftrightarrow$ (ii).

Exercice II.14. ▪ Soit $\|\cdot\|$ une norme sur $\mathbb{R}^N$, $x \in \mathbb{R}^N$ et $r > 0$. Prouvez que la boule $\mathrm{B}_{\|\cdot\|}[x,r]$ est convexe.

▪ Déduisez-en que $|\cdot|_p$ ne peut être une norme pour $p < 1$. (Pour vous aider, $\mathrm{B}_{|\cdot|_p}[0,1]$ pour $p < 1$ est tracée à la figure II.8.)

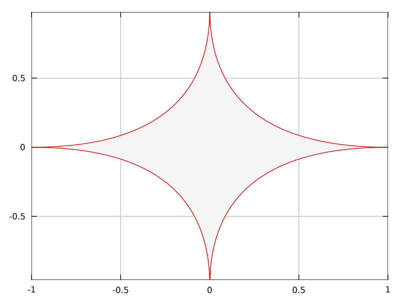


FIGURE II.8 – $\mathrm{B}_{|\cdot|_p}[0,1]$ pour $p <$ 1

Exercice II.15. La moyenne géométrique $x \mapsto \sqrt{\|x\|_1 \|x\|_2}$ de deux normes $\|\bullet\|_1$ et $\|\bullet\|_2$ n'est pas nécessairement une norme. Définissons, par exemple, $f(x) := |x|_1^{1/2} |x|_\infty^{1/2}$. f n'est pas une norme car, bien qu'elle satisfasse les propriétés (i) et (ii) de la définition II.1, la propriété (iii) n'est pas vérifiée. Prouvez ces affirmations. Pour vous aider concernant l'inégalité triangulaire, voyez son lien avec la convexité des boules (exercice II.14) et utilisez la figure II.9 qui représente la boule unité pour f dans $\mathbb{R}^2$.

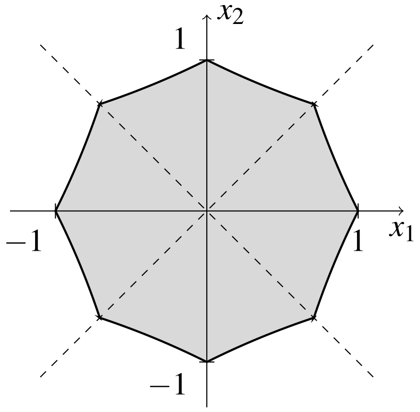


FIGURE II.9 – Boule pour la moyenne géométrique de $|\cdot|_1$ et $|\cdot|_\infty$

Exercice II.16 (Inégalité de Hölder). Soient $p,q \in [1,+\infty]$ tels que $1/p+1/q = 1$. Prouvez que,

$$\forall x,y \in \mathbb{R}^N, \quad \big|(x|y)\big| \leqslant |x|_p |y|_q. \tag{II.3}$$

INDICATION : On peut supposer sans perte de généralité que $1 \leqslant p \leqslant q \leqslant +\infty$. Commencez par le cas $p = 1$, $q = +\infty$. Ensuite, pour $p,q \in\,]1,+\infty[$, prouvez l'inégalité de Young :

$$\forall \xi,\eta \in \mathbb{R}, \quad |\xi\eta| \leqslant \tfrac{1}{p}|\xi|^p + \tfrac{1}{q}|\eta|^q$$

en montrant que la fonction $[0,+\infty[\, \to \mathbb{R} : \xi \mapsto \xi\eta - \xi^p/p$ atteint son maximum en $\eta^{1/(p-1)}$ Pour $x = (x_1,\dots,x_N)$ et $y = (y_1,\dots,y_N)$, appliquez l'inégalité de Young à $\xi = x_i/|x|_p$ et $\eta = y_i/|y|_q$ et sommez sur i.

Exercice II.17. En utilisant l'inégalité de Hölder (II.3), prouvez que, si $p,q \in [1,+\infty]$ sont tels que $1/p+1/q = 1$, on a

$$\forall x \in \mathbb{R}^N, \qquad \max_{|y|_q \leqslant 1} (y|x) = |x|_p$$

Exercice II.18. Prouvez que $|\cdot|_p$ est une norme si $p \geqslant 1$.

INDICATION : Tout d'abord montrez que $|\cdot|_p : \mathbb{R}^N \to [0,+\infty[$ vérifie les deux premières propriétés de la définition II.1 (ce qui est facile). Reste l'inégalité triangulaire aussi appelée dans ce cas ci « inégalité de Minkowski ». Établissez d'abord que

$$\sum_{i=1}^{N} |x_i + y_i|^p \leqslant \sum_{i=1}^{N} |x_i||x_i + y_i|^{p-1} + \sum_{i=1}^{N} |y_i||x_i + y_i|^{p-1}$$

puis appliquez l'inégalité de Hölder aux deux sommes $\sum_{i=1}^{N}$ du membre de droite.

Exercice II.19 (Fractals de Julia et Mandelbrot). On définit la famille de fonctions $f_c : \mathbb{C} \to \mathbb{C} : z \mapsto z^2 + c$ où $c \in \mathbb{C}$ est un paramètre. À un $z_0 \in \mathbb{C}$ on associe son orbite $(z_{c,n})_{n\in\mathbb{N}}$ par f_c définie par

$$\begin{cases} z_{c,0} = z_0 \\ z_{c,n+1} = f_c(z_{c,n}) \end{cases}$$

On s'intéresse à l'ensemble J_c des z_0 pour lesquels la suite $(z_{c,n})_{n\in\mathbb{N}}$ ne converge pas vers ∞ : $J_c := \complement\big\{z_0 \in \mathbb{C} : |z_{c,n}| \xrightarrow[n\to\infty]{} \infty\big\}$. J_c est appelé l'*ensemble plein de Julia* de f_c.

- Montrez que, si $|z_{c,n}| > 2$ pour un certain n, alors $|z_{c,n}| \xrightarrow[n\to\infty]{} \infty$. On a donc $J_c := \big\{z_0 \in \mathbb{C} : \forall n \in \mathbb{N},\ |z_{c,n}| \leqslant 2\big\}$.
- Grâce au résultat précédent, écrivez un algorithme qui permet de tracer J_c aussi précisément que l'on veut.

L'ensemble J_c est un fractal sauf pour $c = -2$ et $c = 0$ (il est facile de voir que $J_0 = B[0,1]$). L'*ensemble de Mandelbrot* est l'ensemble des $c \in \mathbb{C}$ tel que $z_{c,n} \xrightarrow[n\to\infty]{} \infty$ au départ de $z_{c,0} = 0$. Celui-ci est tracé à la figure II.13.

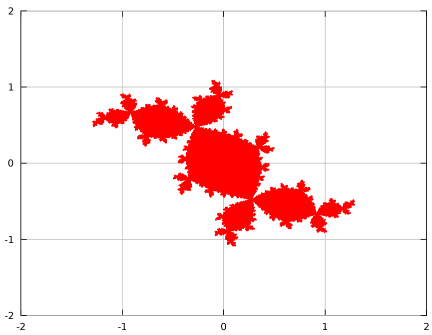


FIGURE II.10 – $J_{-0,123+0,745i}$

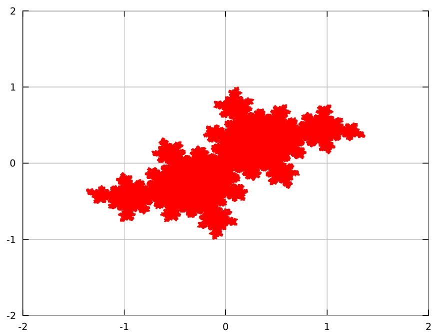


FIGURE II.11 – $J_{-0,391-0.587i}$

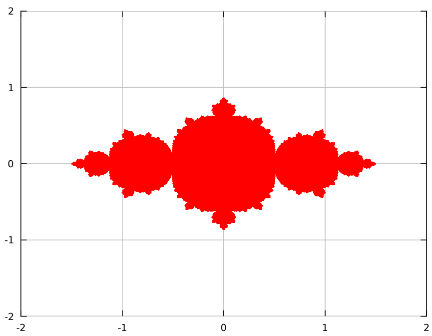


FIGURE II.12 – $J_{-0,75}$

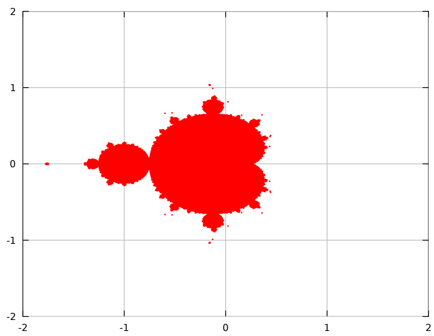


FIGURE II.13 – Ensemble de Mandelbrot

Chapitre III

Notions de topologie

Nous allons nous intéresser ici à la relation entre les ensembles et la convergence des suites. Ce faisant, nous définirons des notions qui ne sont pas liées à une métrique (une manière de mesurer les distances) particulière mais qui expriment des relations liées à une notion abstraite de proximité. D'où le nom de topologie (de *topos* : lieu et *logos* : langage).

III.1 Intérieur, adhérence, ouvert, fermé

Considérons une suite convergente (x_n) dans un intervalle $]a,b[$. Le résultat sur les limites d'inégalités dit que sa limite se trouve dans $]a,b[$ ou dans $\{a,b\}$ — qui est le bord de $]a,b[$. Ceci est-il juste un cas particulier ou au contraire un exemple représentatif ? Et comment définir le bord d'un ensemble de $\mathbb{R}^N$? Si l'ensemble est simple, comme par exemple $\mathrm{B}_{\|\cdot\|}(0,1)$, c'est facile : son bord est la sphère unité $\mathbb{S}_{\|\cdot\|} := \{x \in \mathbb{R}^N : \|x\| = 1\}$. Mais cela a-t-il encore un sens de parler du bord d'un ensemble plus complexe, voire très irrégulier ? Si l'on se base sur l'exemple de l'intervalle, il est plus facile de définir l'ensemble « avec son bord » que le bord tout seul — $[a,b]$ est l'ensemble des limites des suites convergentes de $]a,b[$. C'est ce que nous allons faire. Si A est un sous-ensemble de $\mathbb{R}^N$, nous définissons « A avec son bord », ou techniquement l'adhérence de A, comme l'ensemble des limites des suites convergentes de A (voir figure III.1), c'est-à-dire l'ensemble A lui-même auquel on a ajouté tous les points qui « collent » à A.

Intéressons nous maintenant à un concept apparenté : l'intérieur d'un ensemble. On a envie de dire qu'on est à l'intérieur d'un ensemble si on n'est pas sur son bord, c'est-à-dire si on a un peu d'espace autour de soi. Par exemple, on a envie de dire que x est à l'intérieur de A sur la figure III.2 car non seulement c'est un point de A mais toute la zone grisée autour de lui est aussi dans A. Qu'est-ce que cela a à voir avec la notion de convergence ? La figure III.2 le montre : un point x est à l'intérieur d'un ensemble si n'importe quelle suite convergeant vers x finit par entrer dans l'ensemble. Autrement dit, si $x_n \to x$, alors $x_n \in A$ lorsque n est grand. Au contraire, le x' de la figure III.2 est sur le bord et non pas à l'intérieur

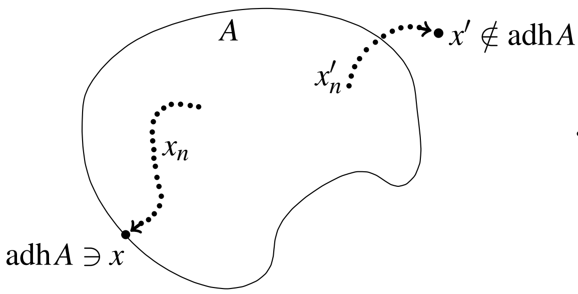


FIGURE III.1 – Adhérence

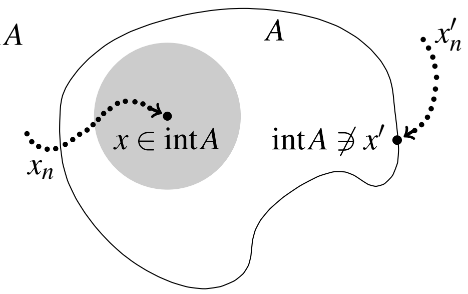


FIGURE III.2 – Intérieur

de A et il existe une suite (x'_n) qui converge vers lui sans jamais entrer dans A. Les définitions formelles qui suivent devraient maintenant vous paraître naturelles.

Définition III.1. Soit $A \subseteq \mathbb{R}^N$. L'*intérieur* de A, noté $\text{int}A$, et l'*adhérence* de A, noté $\text{adh}A$, sont les sous-ensembles de $\mathbb{R}^N$ définis par

$$\text{int}A := \{x \in \mathbb{R}^N : \forall (x_n)_{n\in I} \subseteq \mathbb{R}^N,\ x_n \to x \Rightarrow \exists n_0 \in \mathbb{N},\ \forall n \geqslant n_0,\ x_n \in A\}$$
$$\text{adh}A := \{x \in \mathbb{R}^N : \exists (x_n)_{n\in I} \subseteq A,\ x_n \to x\}$$

D'autres notations répandues sont $\mathring{A}$ pour l'intérieur de A et $\overline{A}$ pour l'adhérence de A. Comme on s'y attend d'après les intuitions données, on a

$$\text{int}A \subseteq A \subseteq \text{adh}A.$$

Pour la première inclusion, étant donné $x \in \text{int}A$, on peut considérer la suite constante $(x_n) = (x)_{n\in\mathbb{N}} \subseteq \mathbb{R}^N$ qui converge vers x et qui, selon la définition d'intérieur, doit vérifier $x_n \in A$ pour n grand, d'où $x = x_n \in A$. La seconde inclusion se montre de manière similaire : si $x \in A$, la suite constante $(x)_{n\in\mathbb{N}} \subseteq A$ converge vers x et donc $x \in \text{adh}A$.

D'après les dessins, le bord d'un ensemble est ce qu'il faut ajouter à l'intérieur pour avoir l'adhérence. C'est ce que nous allons prendre comme définition.

Définition III.2. Le *bord* ou la *frontière* d'un ensemble $A \subseteq \mathbb{R}^N$ est l'ensemble $(\operatorname{adh} A) \setminus (\operatorname{int} A)$.

Les inclusions $\operatorname{int} A \subseteq A \subseteq \operatorname{adh} A$ situent l'ensemble A par rapport à $\operatorname{int} A$ qui ne comprend aucun point du bord, et $\operatorname{adh} A$ qui les inclus tous. Les deux cas d'égalité sont importants.

Définition III.3. Soit $A \subseteq \mathbb{R}^N$. On dit que A est *ouvert* si $A = \operatorname{int} A$ et que A est *fermé* si $A = \operatorname{adh} A$.

Ainsi un ensemble ouvert est un ensemble qui n'a pas de bord et donc dont tous les points ont un peu d'espace autour d'eux (de taille variable selon le point). Un ensemble fermé contient toutes les limites des suites convergentes qui sont dans cet ensemble. Autrement dit, les limites ne peuvent pas « s'échapper » d'un ensemble fermé.

La définition de l'intérieur traduisait en termes de suites le fait qu'un point x intérieur à un ensemble avait « de la place » autour de lui. Mais on pourrait vouloir exprimer cela directement en disant qu'il y a une petite boule $\mathrm{B}(x,r)$ centrée en x et qui est dans A (voir figure III.4). De même, le fait qu'un point x soit dans l'adhérence d'un ensemble A, qu'il « colle » à A, peut s'exprimer en termes de boules en disant que toutes les boules $\mathrm{B}(x,r)$ centrées en x, même les très petites, intersectent A (voir figure III.3). Ceci nous mène naturellement à prouver :

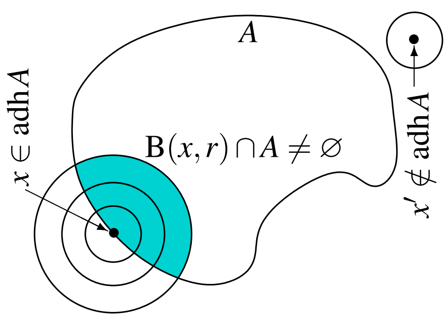


FIGURE III.3 – Adhérence

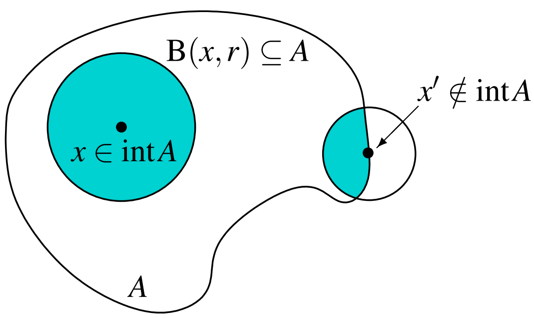


FIGURE III.4 – Intérieur

Proposition III.4. *Soit $A \subseteq \mathbb{R}^N$, $x \in \mathbb{R}^N$ et $\|\cdot\|$ une norme sur $\mathbb{R}^N$. Les équivalences suivantes sont vraies :*

- $x \in \operatorname{int} A \Leftrightarrow \exists r > 0,\ \mathrm{B}_{\|\cdot\|}(x,r) \subseteq A$ *;*

- $x \in \operatorname{adh} A \Leftrightarrow \forall r > 0,\ \mathrm{B}_{\|\cdot\|}(x,r) \cap A \neq \varnothing$.

Remarquez qu'à priori, les propriétés à droite des signes « $\Leftrightarrow$ » dépendent de la norme qu'on considère. Comme elles sont équivalentes à celles de gauche, ce n'est pas le cas. En fait, puisque toutes les normes sur $\mathbb{R}^N$ sont équivalentes et puisque l'équivalence de normes se traduit en inclusions de boules, il est facile de montrer directement (faites le !) que les propriétés à droite des symboles « $\Leftrightarrow$ » sont équivalentes pour toutes les normes sur $\mathbb{R}^N$. D'un point de vue pratique, cela permet de choisir la norme, et donc la forme de la boule, la mieux adaptée à la situation.

Démonstration. Il faut montrer deux équivalences, c'est-à-dire quatre implications.

- Supposons que $x \in \operatorname{int} A$. Nous voulons montrer qu'il existe un $r > 0$ tel que $\mathrm{B}_{\|\cdot\|}(x,r) \subseteq A$. Supposons au contraire que ce ne soit pas possible, c'est-à-dire que

$$\forall r > 0,\ \exists y \notin A,\ \|y - x\| < r$$

(la non-inclusion de $\mathrm{B}_{\|\cdot\|}(x,r)$ dans A dit qu'il existe un y qui appartient à la boule mais pas à A). En choisissant $r = 1/n$, $n \in \mathbb{N} \setminus \{0\}$, on trouve un $y_n \notin A$ tel que $\|y_n - x\| < 1/n$. Comme n est quelconque, on a en fait une suite $(y_n)_{n \geqslant 1} \subseteq \mathbb{R}^N$. Puisque $\|y_n - x\| < 1/n \to 0$, on a $y_n \to x$. Mais le fait que $x \in \operatorname{int} A$ implique, par définition, qu'il existe un $n_0 \in \mathbb{N}$ tel que, si $n \geqslant n_0$ et en particulier pour $n = n_0$, $y_n \in A$. Ceci contredit le fait que, par construction, aucun des y_n n'appartient à A.

- Supposons que $\mathrm{B}_{\|\cdot\|}(x,r) \subseteq A$ pour un certain $r > 0$ et montrons que $x \in \operatorname{int} A$. Soit $(x_n)_{n \in I}$ une suite qui converge vers x. On veut prouver qu'elle finit par rentrer dans A. Par définition de $x_n \to x$, on a

$$\exists n_0 \in \mathbb{N},\ \forall n \geqslant n_0,\ \|x_n - x\| \leqslant r/2 < r.$$

Comme $\|x_n - x\| < r$ implique que $x_n \in \mathrm{B}_{\|\cdot\|}(x,r) \subseteq A$, on a fini.

- Supposons maintenant $x \in \operatorname{adh} A$. Soit $r > 0$. Montrons que $\mathrm{B}_{\|\cdot\|}(x,r) \cap A \neq \varnothing$. Par définition de l'adhérence, il existe une suite $(x_n)_{n \in I} \subseteq A$ telle que $x_n \to x$. De la définition de convergence, on déduit que, pour n suffisament grand, $\|x_n - x\| \leqslant r/2 < r$ et donc $x_n \in \mathrm{B}(x,r)$. Dès lors $x_n \in \mathrm{B}(x,r) \cap A$ et cet ensemble ne peut être vide.

- Pour finir, supposons que $\forall r > 0,\ \mathrm{B}_{\|\cdot\|}(x,r) \cap A \neq \varnothing$ et déduisons que $x \in \operatorname{adh} A$. En prenant $r = 1/n$, $n \geqslant 1$, dans l'hypothèse, on construit une suite $(x_n)_{n\geqslant 1} \subseteq A$ telle que $x_n \in \mathrm{B}_{\|\cdot\|}(x, 1/n)$ pour tout n. Dès lors $\|x_n - x\| < 1/n \to 0$ ce qui implique que $x_n \to x$ et donc x satisfait la définition de $x \in \operatorname{adh} A$. □

Exemple III.5. Nous allons montrer que

$$\operatorname{adh} \mathrm{B}_{\|\cdot\|}(x,r) = \mathrm{B}_{\|\cdot\|}[x,r] \quad \text{et} \quad \operatorname{int} \mathrm{B}_{\|\cdot\|}[x,r] = \mathrm{B}_{\|\cdot\|}(x,r).$$

Si $(y_n) \subseteq \mathrm{B}(x,r)$ est une suite convergeant vers y, alors $\|y_n - x\| \to \|y - x\|$ et, puisque $\|y_n - x\| < r$ pour tout n, il s'ensuit que $\|y - x\| \leqslant r$. Nous venons de montrer que $\operatorname{adh} \mathrm{B}_{\|\cdot\|}(x,r) \subseteq \mathrm{B}_{\|\cdot\|}[x,r]$. Pour l'inclusion inverse, prenons $y \in \mathrm{B}[x,r]$ et considérons $y_n := x/n + (1 - 1/n)y$ (faites un dessin et visualisez où se trouve y_n). Vu que $\|y_n - x\| = (1-1/n)\|y - x\| < \|y - x\| \leqslant r$, on a que $y_n \in \mathrm{B}(x,r)$. Bien sûr, $y_n \xrightarrow[n\to\infty]{} y$.

Soit $y \in \operatorname{int} \mathrm{B}[x,r]$ et montrons que $y \in \mathrm{B}(x,r)$. Si $y = x$, il n'y a rien à faire. Reste le cas $y \neq x$. Le fait que y soit à l'intérieur implique que $\mathrm{B}(y,\rho) \subseteq \mathrm{B}[x,r]$ pour un certain $\rho > 0$. On a $y + \delta(y - x) \in B(y,\rho) \subseteq \mathrm{B}[x,r]$, avec $\delta := \frac{\rho}{2\|y-x\|}$, et donc $(1+\delta)\|y - x\| \leqslant r$. Comme $1 < 1+\delta$, cela implique $\|y - x\| < r$ comme désiré. Inversement, si $y \in \mathrm{B}(x,r)$, alors $\mathrm{B}(y,\rho) \subseteq \mathrm{B}[x,r]$ avec $\rho := r - \|y - x\|$, ce qui montre bien que $y \in \operatorname{int} \mathrm{B}[x,r]$.

Corollaire III.6. *Soit A un sous-ensemble de $\mathbb{R}^N$ et $\|\cdot\|$ une norme sur $\mathbb{R}^N$. Les équivalences suivantes sont vraies :*

- *A est ouvert $\Leftrightarrow \forall x \in A,\ \exists r > 0,\ \mathrm{B}_{\|\cdot\|}(x,r) \subseteq A$;*
- *A est fermé $\Leftrightarrow \forall x \in \mathbb{R}^N,\ \big(\forall r > 0,\ \mathrm{B}_{\|\cdot\|}(x,r) \cap A \neq \varnothing\big) \Rightarrow x \in A$.*

Démonstration. Les formules à droite des symboles d'équivalence expriment respectivement que $A \subseteq \operatorname{int} A$ et $\operatorname{adh} A \subseteq A$. □

Exemple III.7. $\mathrm{B}_{\|\cdot\|}(x,r)$ est ouverte et $\mathrm{B}_{\|\cdot\|}[x,r]$ est fermée. Ceci explique les noms donnés à ces boules.

$\mathrm{B}(x,r)$ est ouverte car, si $y \in \mathrm{B}(x,r)$, alors $\mathrm{B}(y,\rho) \subseteq \mathrm{B}(x,r)$ où $\rho := (r - \|y - x\|)/2$. En effet, si $z \in \mathrm{B}(y,\rho)$, alors $\|z - x\| \leqslant \|z - y\| + \|y - x\| \leqslant \rho + \|y - x\| = (r + \|y - x\|)/2 < r$.

Pour que $\mathrm{B}[x,r]$ soit fermé, il suffit que $\operatorname{adh} \mathrm{B}[x,r] \subseteq \mathrm{B}[x,r]$. Soit $y \in \mathbb{R}^N$ pour lequel il existe une suite $(y_n) \subseteq \mathrm{B}[x,r]$ telle que $y_n \to y$. Il faut voir que $y \in \mathrm{B}[x,r]$.

Par hypothèse, $\|y_n - x\| \leqslant r$ pour tout n. En passant à la limite, on trouve $\|y-x\| = \lim_{n\to\infty}\|y_n - x\| \leqslant r$.

Exemple III.8. Comme cas particuliers de ce que nous avons fait ci-dessus, $]a,b[$ est un sous-ensemble ouvert de $\mathbb{R}$, $[a,b]$ est un sous-ensemble fermé, $\operatorname{adh}]a,b[= [a,b]$ et $\operatorname{int}[a,b] =]a,b[$. En effet, $]a,b[= \mathrm{B}_{|\cdot|}\big((a+b)/2, |b-a|/2\big)$ et $[a,b] = \mathrm{B}_{|\cdot|}\big[(a+b)/2, |b-a|/2\big]$.

Lemme III.9. *Soient A et B deux sous-ensembles de $\mathbb{R}^N$. Si $A \subseteq B$ alors $\operatorname{int} A \subseteq \operatorname{int} B$ et $\operatorname{adh} A \subseteq \operatorname{adh} B$.*

Démonstration. Si $x \in \operatorname{int} A$, cela veut dire qu'il existe un $r > 0$ tel que $\mathrm{B}(x,r) \subseteq A$ et donc $\mathrm{B}(x,r) \subseteq B$, ce qui implique que $x \in \operatorname{int} B$.

Soit $x \in \operatorname{adh} A$. Pour montrer que $x \in \operatorname{adh} B$, prenons un $r > 0$ arbitraire et établissons que $\mathrm{B}(x,r) \cap B \neq \varnothing$. C'est le cas car, par hypothèse, $\mathrm{B}(x,r) \cap A \neq \varnothing$ et $\mathrm{B}(x,r) \cap A \subseteq \mathrm{B}(x,r) \cap B$. □

Avant d'aller plus loin, établissons une forme de dualité antre l'intérieur et l'adhérence qui nous permettra de déduire d'un énoncé sur l'un des deux, un énoncé sur l'autre.

Proposition III.10. *Pour tout $A \subseteq \mathbb{R}^N$, $\operatorname{adh} A = \complement \operatorname{int} \complement A$ et $\operatorname{int} A = \complement \operatorname{adh} \complement A$.*

La seconde égalité est une conséquence de la première : celle-ci avec $\complement A$ au lieu de A implique $\operatorname{adh} \complement A = \complement \operatorname{int} \complement\complement A = \complement \operatorname{int} A$ et il suffit de prendre le complémentaire des deux membres de l'égalité.

Réécrivons $\operatorname{adh} A = \complement \operatorname{int} \complement A$ comme $\complement \operatorname{adh} A = \operatorname{int} \complement A$. Cette égalité est intuitivement vraie. En effet, si un point x n'est pas dans l'adhérence de A, s'il ne colle pas à A, c'est qu'il a un peu d'espace autour de lui dans le complémentaire de A, c'est-à-dire qu'il est dans l'intérieur de $\complement A$ (voir figure III.5).

Démonstration. Nous allons utiliser les caractérisations équivalentes de l'intérieur et de l'adhérence données par la proposition III.4. Quel que soit $x \in \mathbb{R}^N$,

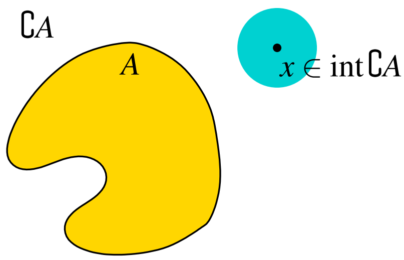


FIGURE III.5 – $\complement \operatorname{adh} A = \operatorname{int} \complement A$

on a

$$\begin{aligned} x \in \complement \operatorname{int} \complement A &\Leftrightarrow \neg\big(x \in \operatorname{int} \complement A\big) \\ &\Leftrightarrow \neg\big(\exists r > 0,\ \mathrm{B}_{\|\cdot\|}(x,r) \subseteq \complement A\big) \\ &\Leftrightarrow \neg\big(\exists r > 0,\ \forall y \in \mathrm{B}_{\|\cdot\|}(x,r),\ y \notin A\big) \\ &\Leftrightarrow \forall r > 0,\ \exists y \in \mathrm{B}_{\|\cdot\|}(x,r),\ y \in A \\ &\Leftrightarrow \forall r > 0,\ \mathrm{B}_{\|\cdot\|}(x,r) \cap A \neq \varnothing \\ &\Leftrightarrow x \in \operatorname{adh} A \end{aligned}$$ □

Cette dualité existe aussi entre les ensembles ouverts et fermés.

Corollaire III.11. *Soit $A \subseteq \mathbb{R}^N$. A est ouvert si et seulement si $\complement A$ est fermé.*

Bien évidemment, en remplaçant A par $\complement A$, on a aussi que A est fermé si et seulement si $\complement A$ est ouvert.

Démonstration. C'est une conséquence directe de la proposition III.10. □

La proposition suivante montre que les opérateurs « int » et « adh » sont idempotents.

Proposition III.12. *Soit $A \subseteq \mathbb{R}^N$.* $\operatorname{int}(\operatorname{int} A) = \operatorname{int} A$ *et* $\operatorname{adh}(\operatorname{adh} A) = \operatorname{adh} A$.

Démonstration. Montrons que $\operatorname{int}(\operatorname{int} A) = \operatorname{int} A$. Comme $\operatorname{int}(\operatorname{int} A) \subseteq \operatorname{int} A$, il suffit de prouver l'inclusion inverse. Soit $x \in \operatorname{int} A$. Il faut montrer qu'on peut trouver un $r > 0$ tel que

$$\mathrm{B}(x,r) \subseteq \operatorname{int} A.$$

Comme $x \in \operatorname{int} A$, il existe un $\rho > 0$ tel que $\mathrm{B}(x,\rho) \subseteq A$. Prenons $r := \rho/2$. Soit $y \in \mathrm{B}(x,r)$. Il faut prouver que $y \in \operatorname{int} A$. Ce sera le cas si on établit que $\mathrm{B}(y,\rho/2) \subseteq A$.

Soit $z \in \mathrm{B}(y, \rho/2)$. Puisque $\|z-x\| \leqslant \|z-y\| + \|y-x\| < \rho/2 + r = \rho$, on a bien $z \in \mathrm{B}(x,\rho) \subseteq A$ (voir figure III.6).

Le cas de l'adhérence se déduit par dualité. En effet, l'égalité $\operatorname{adh}(\operatorname{adh} A) = \operatorname{adh} A$ est équivalente à $\complement \operatorname{adh}(\operatorname{adh} A) = \complement \operatorname{adh} A$, c'est-à-dire $\operatorname{int}(\operatorname{int} \complement A) = \operatorname{int} \complement A$ ce qui est vrai d'après la première partie. $\square$

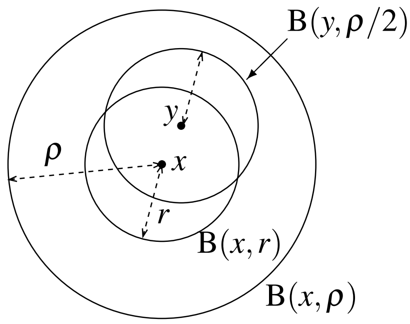


FIGURE III.6 – $\operatorname{int}(\operatorname{int} A) = \operatorname{int} A$

Les autres combinaisons des opérateurs « int » et « adh » peuvent donner de nouveaux ensembles. Par exemple, $\operatorname{int} A \subseteq \operatorname{int}(\operatorname{adh} A)$ — en appliquant le lemme III.9 à $A \subseteq \operatorname{adh} A$ — mais on n'a pas nécessairement l'égalité. Considérons $A = \left]-1,1\right[\setminus \{0\}$. Il est aisé de voir que cet ensemble est ouvert, c'est-à-dire que $\operatorname{int} A = A$. Montrons que $\operatorname{adh} A = [-1,1]$. Clairement, comme $A \subseteq [-1,1]$ et que $[-1,1]$ est fermé, $\operatorname{adh} A \subseteq [-1,1]$. D'autre part, comme $A \subseteq \operatorname{adh} A$, il reste à montrer que $-1, 0, 1 \in \operatorname{adh} A$. C'est le vas car ces trois nombres sont les limites respectives des suites $(-1+1/n)_{n\geqslant 2}$, $(1/n)_{n\geqslant 2}$ et $(1-1/n)_{n\geqslant 2}$ qui sont dans A. En conclusion $A = \operatorname{int} A \subsetneq \operatorname{int}(\operatorname{adh} A) = \left]-1,1\right[$.

III.2 Union et intersection

Intéressons nous maintenant au comportement de ces opérations vis à vis de l'union et de l'intersection. Commençons par préciser la terminologie.

Définition III.13. Une *famille* de sous-ensembles de $\mathbb{R}^N$ est une fonction

$$A \to \mathscr{P}(\mathbb{R}^N) : \alpha \mapsto B_\alpha$$

où A est un ensemble et $\mathscr{P}(\mathbb{R}^N)$ est l'ensemble des parties de $\mathbb{R}^N$. On emploiera souvent la notation $(B_\alpha)_{\alpha\in A}$. On dira que la famille est *finie* si A est fini.

Proposition III.14. *Soit $(B_\alpha)_{\alpha\in A}$ une famille de sous-ensembles de $\mathbb{R}^N$. Alors*

- $\operatorname{int}\Big(\bigcup_{\alpha\in A} B_\alpha\Big) \supseteq \bigcup_{\alpha\in A} \operatorname{int} B_\alpha$ *;*
- $\operatorname{int}\Big(\bigcap_{\alpha\in A} B_\alpha\Big) \subseteq \bigcap_{\alpha\in A} \operatorname{int} B_\alpha$ *et on a l'égalité si A est fini ;*
- $\operatorname{adh}\Big(\bigcap_{\alpha\in A} B_\alpha\Big) \subseteq \bigcap_{\alpha\in A} \operatorname{adh} B_\alpha$ *;*
- $\operatorname{adh}\Big(\bigcup_{\alpha\in A} B_\alpha\Big) \supseteq \bigcup_{\alpha\in A} \operatorname{adh} B_\alpha$ *et on a l'égalité si A est fini.*

Démonstration. ▪ Si $x \in \bigcup_{\alpha\in A} \operatorname{int} B_\alpha$, cela veut dire que $x \in \operatorname{int} B_{\alpha^*}$ pour un certain $\alpha^* \in A$ et donc il existe un $r > 0$ tel que $\mathrm{B}(x,r) \subseteq B_{\alpha^*}$. Dès lors $\mathrm{B}(x,r) \subseteq \bigcup_{\alpha\in A} B_\alpha$ et par conséquent $x \in \operatorname{int}\big(\bigcup_{\alpha\in A} B_\alpha\big)$.

▪ Si $x \in \operatorname{int}\big(\bigcap_{\alpha\in A} B_\alpha\big)$, il existe un $r > 0$ tel que $\mathrm{B}(x,r) \subseteq \bigcap_{\alpha\in A} B_\alpha$. Dès lors, quel que soit $\alpha \in A$, $\mathrm{B}(x,r) \subseteq B_\alpha$ et donc $x \in \operatorname{int} B_\alpha$. Par conséquent, $x \in \bigcap_{\alpha\in A} \operatorname{int} B_\alpha$.

Supposons maintenant que $x \in \bigcap_{\alpha\in A} \operatorname{int} B_\alpha$ et que A soit fini. Par définition, quel que soit $\alpha \in A$, il existe un $r_\alpha > 0$ tel que $\mathrm{B}(x,r_\alpha) \subseteq B_\alpha$. Posons $r := \min\{r_\alpha : \alpha \in A\}$. Puisque $\mathrm{B}(x,r) \subseteq \mathrm{B}(x,r_\alpha) \subseteq B_\alpha$ pour tout α, on déduit que $\mathrm{B}(x,r) \subseteq \bigcap_{\alpha\in A} B_\alpha$ et donc que $x \in \operatorname{int}\big(\bigcap_{\alpha\in A} B_\alpha\big)$.

▪ Les deux autres affirmations se déduisent des premières par dualité. Détaillons celle qui concerne l'intersection. Vu les équivalences

$$\begin{aligned}
\operatorname{adh}\Big(\bigcap_{\alpha\in A} B_\alpha\Big) \subseteq \bigcap_{\alpha\in A} \operatorname{adh} B_\alpha &\Leftrightarrow \complement \operatorname{adh}\Big(\bigcap_{\alpha\in A} B_\alpha\Big) \supseteq \complement \bigcap_{\alpha\in A} \operatorname{adh} B_\alpha \\
&\Leftrightarrow \operatorname{int}\Big(\complement \bigcap_{\alpha\in A} B_\alpha\Big) \supseteq \bigcup_{\alpha\in A} \complement \operatorname{adh} B_\alpha \\
&\Leftrightarrow \operatorname{int}\Big(\bigcup_{\alpha\in A} \complement B_\alpha\Big) \supseteq \bigcup_{\alpha\in A} \operatorname{int} \complement B_\alpha
\end{aligned}$$

et le fait que les dernière formule est vraie par les propriétés de l'intérieur, la première formule de la chaine d'équivalences est aussi vraie. □

Remarque III.15. Cette proposition est optimale au sens où les inclusions non énoncées ne sont pas toujours vraies.

Pour l'intérieur de l'union, considérons $B_1 = [-1,0]$ et $B_2 = [0,1]$. On a sans peine (faites les justifications !) que

$$]-1,1[\;=\; \operatorname{int}(B_1 \cup B_2) \supsetneq \operatorname{int} B_1 \cup \operatorname{int} B_2 \;=\;]-1,0[\cup]0,1[\;=\;]-1,1[\setminus \{0\}.$$

Pour l'intérieur de l'intersection, il faut montrer qu'on n'a pas nécessairement l'égalite si A est infini. Considérons $B_n :=]-1/n, 1/n[$, $n \in A := \mathbb{N} \setminus \{0\}$. Clairement $\bigcap_{n\in A} B_n = \{0\}$. En effet, $0 \in B_n$ pour tout n, d'où $0 \in \bigcap B_n$ et, si $x \in \bigcap B_n$, alors $|x| < 1/n$ pour tout n d'où, en passant à la limite $n \to +\infty$, on trouve que $x = 0$. On a donc $\varnothing = \operatorname{int}\big(\bigcap_{n\in A} B_n\big) \subsetneq \bigcap_{n\in A} \operatorname{int} B_n = \bigcap_{n\in A} B_n = \{0\}$.

Des exemples qui montrent qu'on n'a pas toujours l'égalité pour les deux dernières inclusions se trouvent par dualité de ceux ci-dessus. Nous invitons le lecteur à écrire le détail de cet argument de dualité et/ou a chercher ses propres contre-exemples.

Comme d'habitude, ce qui est prouvé sur l'intérieur et l'adhérence a des conséquences pour les ouverts et les fermés.

Corollaire III.16. *Si $(O_\alpha)_{\alpha\in A}$ est une famille d'ouverts de $\mathbb{R}^N$, alors*

- $\bigcup_{\alpha\in A} O_\alpha$ *est ouvert ;*
- $\bigcap_{\alpha\in A} O_\alpha$ *est ouvert si A est fini.*

Si $(F_\alpha)_{\alpha\in A}$ est une famille de fermés de $\mathbb{R}^N$, alors

- $\bigcap_{\alpha\in A} F_\alpha$ *est fermé ;*
- $\bigcup_{\alpha\in A} F_\alpha$ *est fermé si A est fini.*

Une famille d'ouverts (resp. de fermés) est bien entendu une famille de sous-ensembles qui sont tous ouverts (resp. fermés).

Démonstration. La proposition précédente implique que $\operatorname{int}(\bigcup O_\alpha) \supseteq \bigcup \operatorname{int} O_\alpha = \bigcup O_\alpha$. Comme par ailleurs l'intérieur d'un ensemble est toujours inclus à cet ensemble, on a $\operatorname{int}(\bigcup O_\alpha) = \bigcup O_\alpha$.

Pour l'intersection d'ouverts, la proposition précédente nous donne de suite la réponse : $\operatorname{int}(\bigcap O_\alpha) = \bigcap \operatorname{int} O_\alpha = \bigcap O_\alpha$.

Les affirmations sur les fermés se déduisent par dualité. (Le lecteur est invité à en imaginer une preuve directe.) □

III.3 Densité

Terminons cette introduction à la topologie en introduisant quelques notions d'usage fréquent.

Tout d'abord, parlons de la densité. Un ensemble A est dit dense dans B s'il est « presque partout » présent dans B, c'est-à-dire si tout point de B peut être « bien approximé » par des points de A. En d'autres mots, pour tout $b \in B$, il existe une suite $(a_n) \subseteq A$ telle que $a_n \to b$. Plus succinctement, on peut définir la densité comme suit.

Définition III.17. Soient A et B deux sous-ensembles de $\mathbb{R}^N$. On dit que A est *dense* dans B si et seulement si $A \subseteq B \subseteq \operatorname{adh} A$.

Par exemple, la proposition I.39 dit que $\mathbb{Q}$ est dense dans $\mathbb{R}$. Le lecteur pourra facilement vérifier que $\mathbb{Q}^N$ est dense dans $\mathbb{R}^N$. L'intérêt de la densité est que si une propriété ou une fonction est définie sur A, que celle-ci est suffisament « continue » et que A est dense dans B, alors on peut en général l'étendre à B. Cette démarche a été abondamment utilisée à la section I.5 où on a étendu diverses propriétés et fonctions de $\mathbb{Q}$ à $\mathbb{R}$. Elle le sera de nouveau dans des cours d'Analyse plus avancés.

III.4 Voisinages

Finalement, abordons la notion de voisinage. D'après le terme, si V est un voisinage d'un point x, c'est que x a un peu d'espace autour de lui. Plus précisément :

Définition III.18. Soit $V \subseteq \mathbb{R}^N$, $x \in \mathbb{R}^N$ et $\|\cdot\|$ une norme sur $\mathbb{R}^N$. On dit que V est un *voisinage* de x si il existe un $r > 0$ tel que $\mathrm{B}_{\|\cdot\|}(x,r) \subseteq V$.

Une fois de plus, l'équivalence de toutes les normes sur $\mathbb{R}^N$ fait que la notion de voisinage est indépendante de la norme. D'ailleurs, V est un voisinage de x si et seulement si $x \in \operatorname{int} V$. On peut alors dire qu'un ouvert est un ensemble qui est un voisinage de chacun de ses points.

Les voisinages sont fort flexibles : à part $x \in \operatorname{int} V$, on n'impose pas de forme à V — e.g., que V soit une boule — ni de propriétés du type V ouvert, fermé,... Les voisinages offrent néanmoins un langage naturel pour exprimer les propriétés topologiques. Bien souvent les boules peuvent être remplacées par des voisinages. L'avantage de ceux-ci est qu'ils ne dépendent pas d'une norme sous-jacente. Pour appuyer l'affirmation ci-dessus, intéressons nous aux deux propositions suivantes.

Proposition III.19. *Soit $(x_n)_{n\in I} \subseteq \mathbb{R}^N$ et $x \in \mathbb{R}^N$. Alors $x_n \to x$ si et seulement si*

$$\forall V \text{ voisinage de } x,\ \exists n_0 \in \mathbb{N},\ \forall n \geqslant n_0,\ x_n \in V. \tag{III.1}$$

Démonstration. Condition nécessaire. Supposons que $x_n \to x$ et montrons (III.1). Soit V un voisinage de x. Par définition, il existe un $r > 0$ tel que $B_{\|\cdot\|}(x,r) \subseteq V$. Par définition de $x_n \to x$, il existe un $n_1 \in \mathbb{N}$ tel que $\forall n \geqslant n_1, \|x_n - x\| \leqslant r/2$. Prenons $n_0 := n_1$. Si $n \geqslant n_0$, $\|x_n - x\| < r$ ce qui implique $x_n \in B_{\|\cdot\|}(x,r) \subseteq V$.

Condition suffisante. C'est évident car dans la définition de convergence $\|x_n - x\| \leqslant r$ peut être interprété comme $x_n \in B_{\|\cdot\|}[x,r]$ et $B_{\|\cdot\|}[x,r]$ est un voisinage de x. □

Proposition III.20. *Soit $A \subseteq \mathbb{R}^N$ et $x \in \mathbb{R}^N$. On a les équivalences suivantes :*

- $x \in \operatorname{int} A \Leftrightarrow \exists V$ *voisinage de* x, $V \subseteq A$ *;*
- $x \in \operatorname{adh} A \Leftrightarrow \forall V$ *voisinage de* x, $A \cap V \neq \varnothing$.

Démonstration. Laissée au lecteur. □

Dans cette section, nous sommes partis de la notion de convergence pour définir l'intérieur et l'adhérence. Ces propositions montrent qu'on aurait pu très bien commencer avec la notion de voisinage.

De manière équivalente, on aurait pu commencer avec le concept d'ouvert. C'est d'ailleurs ce qu'on fait lorsqu'on aborde la topologie de manière générale. On commence par se donner les ouverts d'un espace — qui doivent vérifier les propriétés du corollaire III.16 — et, à partir de ceux-ci, on définit les fermés en prenant leur complémentaire, les notions d'intérieur, d'adhérence, de convergence,... Pour nous persuader que c'est possible, examinons la proposition suivante.

Proposition III.21. *Soit $A \subseteq \mathbb{R}^N$. L'intérieur de A est le plus grand ouvert O inclus à A au sens où*

- $O \subseteq A$ *et*
- *si O' est un ouvert inclus à A, alors $O' \subseteq O$.*

L'adhérence de A est le plus petit fermé F contenant A au sens où

- $F \supseteq A$ *et*
- *si F' est un fermé contenant A, $F' \supseteq F$.*

Remarque III.22. Le plus grand ouvert inclus à A existe : il suffit de prendre l'union de tous les ouverts inclus à A,

$$O := \bigcup \{O' : O' \text{ ouvert et } O' \subseteq A\}.$$

C'est encore un ouvert en vertu du corollaire III.16.

Démonstration. Montrons que $O := \operatorname{int} A$ est le plus grand ouvert de A. Tout d'abord, $O \subseteq A$. Ensuite, si $O' \subseteq A$ est un ouvert, le lemme III.9 implique que $O' = \operatorname{int} O' \subseteq \operatorname{int} A = O$.

Les affirmations sur les fermés se déduisent par dualité. □

III.5 Exercices

Exercice III.1.
- $\{2\}$ est-il un ensemble ouvert ? Fermé ? Justifiez.
- $[0,1]$ est-il un ensemble ouvert ? Fermé ? Justifiez.
- $]0,3]$ est-il un ensemble ouvert ? Fermé ? Justifiez.

Exercice III.2.
- Montrez que $]a,b[$, avec $a,b \in \mathbb{R} \cup \{-\infty, +\infty\}$, est ouvert.
- Montrez que $[a,b]$, avec $a,b \in \mathbb{R}$, est fermé.

Exercice III.3. Montrez que $\operatorname{adh}]-3,5] = [-3,5]$ et $\operatorname{int}]-3,5] =]-3,5[$.

Exercice III.4. Les ensembles $\varnothing$, $\mathbb{R}$, $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ sont-ils ouverts ? Fermés ? Justifiez.

Exercice III.5. L'ensemble $\{(a,b)\}$ où $a,b \in \mathbb{R}$ est-il ouvert ? Fermé ?

Exercice III.6. Les ensembles suivants sont-ils ouverts ? Fermés ?
- $A_1 = \{x \in \mathbb{R} : x^2 < x\}$
- $A_2 = \{(x,y) \in \mathbb{R}^2 : 4x - y = 5\}$
- $A_3 = \{(x,y) \in \mathbb{R}^2 : 0 \leqslant x < y\}$
- $A_4 = \{(x,x^3) \in \mathbb{R}^2 : x \in \mathbb{R}\}$
- $A_5 = \{(x,y) \in \mathbb{R} : |x| \leqslant 1\}$
- $A_6 = \{\frac{1}{n} : n \in \mathbb{N} \setminus \{0\}\}$

Exercice III.7. Montrez que $[-1,2] \times [-4,-3]$ est un ensemble fermé.

Exercice III.8. Montrez que $\mathrm{B}_{|\cdot|_1}[(-2,1),3]$ est un ensemble fermé.

Exercice III.9. Montrez que $\mathrm{B}_{|\cdot|_\infty}\big((-2,1),3\big)$ est un ensemble ouvert.

Exercice III.10. Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction continue. Posons $E := \{x \in \mathbb{R} : f(x) \leqslant 4\}$. Montrez que E est un ensemble fermé.

Exercice III.11.
- Si O est un ouvert de $\mathbb{R}^N$ et $x_0 \in \mathbb{R}^N$, alors $O \setminus \{x_0\}$ est encore un ouvert.
- Si O est un ouvert et F un fermé de $\mathbb{R}^N$, alors $O \setminus F$ est encore un ouvert et $F \setminus O$ est encore un fermé de $\mathbb{R}^N$.

Exercice III.12.
- Prouvez que $\pi \cdot \mathbb{Q}$ est dense dans $\mathbb{R}$.
- Prouvez que $\{(x,y) \in \mathbb{R}^2 : \exists \alpha \in \mathbb{Q},\ y = \alpha x\}$ est dense dans $\mathbb{R}^2$. (Interprétez géométriquement ce fait.)

Exercice III.13. Soit $p(x) = a_0 + a_1 x + \cdots + a_n x^n$ avec $a_n \neq 0$ et $\forall i = 0, \dots, n$, $a_i \in \mathbb{R}$. Posons $E = \{x \in \mathbb{R} : p(x) = 0\}$, $F = \{x \in \mathbb{R} : \sin x \neq 0\}$ et

$$G = \{x \in \mathbb{R} : p(x) \cdot \sin x \neq 0\}.$$

G est-il ouvert ?

Exercice III.14. Soient $a, b \in \mathbb{R}$ avec $a \neq b$. Montrez qu'il existe une fonction $f : [0,1] \to [a,b]$ continue, bijective, et telle que son inverse f^{-1} soit continue (une telle fonction s'appelle un *homéomorphisme*).

Exercice III.15 (juin 2007). Soit $A \subseteq \mathbb{R}^N$. Montrez que « A est borné » est équivalent à

$$\exists \rho > 0,\ \forall x \in A,\ A \subseteq B_{\|\cdot\|}[x, \rho] \tag{III.2}$$

Exercice III.16 (juin 2007). Soit $(x_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}$ une suite bornée. Montrez $\exists \delta > 0,\ \forall n \in \mathbb{N},\ |x_n| \leqslant x_n + \delta$.

Exercice III.17 (août 2007). Montrer que $A \subseteq \mathbb{R}^N$ est borné si et seulement si

$$\forall (x_n)_{n \in \mathbb{N}} \subseteq A,\ \exists R > 0,\ \forall n \in \mathbb{N},\ \|x_n\| \leqslant R \tag{III.3}$$

Exercice III.18. Nous savons qu'une intersection finie d'ouverts est encore ouverte. Montrez que ce n'est plus vrai en général pour une intersection infinie. (Indication : pensez à faire simple, en particulier à ce qui se passe en dimension un pour commencer.)

Exercice III.19. Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction continue. Supposons que

$$\forall t \in \mathbb{R},\ \exists k \in \mathbb{Z},\ f(t) \in \left]k, k+1\right[.$$

Prouvez que $\exists k \in \mathbb{Z},\ \forall t \in \mathbb{R},\ f(t) \in \left]k, k+1\right[$.

Exercice III.20. Soit $A \subseteq \mathbb{R}^N$. On définit la distance de x à A comme le nombre

$$\operatorname{dist}(x, A) := \inf\{|x - a| : a \in A\}.$$

- Démontrez que si $A \neq \varnothing$, $\operatorname{dist}(x, A) \in [0, +\infty[$.
- Montrez que $A_{<\varepsilon} := \{x \in \mathbb{R}^N : \operatorname{dist}(x, A) < \varepsilon\}$ est un ensemble ouvert contenant A quel que soit $\varepsilon > 0$.
- Prouvez que $x \mapsto \operatorname{dist}(x, A)$ est une fonction continue.
- Prouvez que $A_{\leqslant\varepsilon} := \{x \in \mathbb{R}^N : \operatorname{dist}(x, A) \leqslant \varepsilon\}$ est un ensemble fermé contenant A.
- Justifiez le fait que : $\operatorname{dist}(x, a) = 0 \Leftrightarrow x \in \operatorname{adh}(A)$.
- Établissez que $\bigcap_{\varepsilon>0} A_{<\varepsilon} = \bigcap_{\varepsilon>0} A_{\leqslant\varepsilon} = \operatorname{adh}(A)$.

Exercice III.21 (août 2006). Soit $A \subseteq \mathbb{R}$. Établissez que

$$\big(\exists r > 0,\ \forall x \in A,\ \mathrm{B}(x, r) \subseteq A\big) \quad \Leftrightarrow \quad A = \varnothing \text{ ou } A = \mathbb{R}$$

Exercice III.22 (mars 2008). On considère les trois ensembles suivants :

$$E = \big\{(x, y) \in \mathbb{R}^2 \bigm| x - 2y = 3\big\}, \qquad F = B_{|\bullet|_2}\big((1,1), 2\big), \qquad G = \left]-1, 1\right[.$$

En utilisant la définition III.1 et éventuellement la proposition III.4, dites si

(a) Vrai : ☐ Faux : ☐ $(5, 1)$ est un point intérieur à E.

(b) Vrai : ☐ Faux : ☐ $(1, 1/2)$ est un point intérieur à F.

(c) Vrai : ☐ Faux : ☐ 1 est un point adhérent à G.

Justifiez vos réponses et détaillez vos calculs.

Exercice III.23. On considère les deux ensembles suivants :

$$A = \big\{x \in \mathbb{R} \bigm| x^2 + x + 1 \leqslant 1\big\}, \quad B = \big\{(x, y) \in \mathbb{R}^2 \bigm| x^4 + y^4 - \sin(xy) < 17\big\}.$$

Pour chacun d'entre-eux, dites s'il s'agit d'un ensemble ouvert, fermé ou borné. Veillez à justifier vos affirmations avec suffisament de détails.

Chapitre IV

Compacité

La compacité est un concept clé qui se retrouve dans de nombreuses branches des mathématiques. De manière très grossière, mais qui illustre bien son importance, on peut dire que la compacité permet de ramener l'étude d'une infinité d'objets ou de cas possibles à un nombre fini d'entre eux. Ce chapitre la décrit dans le cadre de la topologie usuelle sur $\mathbb{R}^N$.

Nous allons commencer par donner différentes visions, à priori disparates, de cette notion. Nous formaliserons ces approches et en prouverons l'équivalence. Nous montrerons ensuite comment cela permet de résoudre diverses questions importantes.

IV.1 Introduction

§1. Comme nous l'avons vu à la section I, la complétude de $\mathbb{R}$ est équivalente au fait qu'une suite d'intervalles fermés non-vides emboîtés possède une intersection non-vide. L'impression qu'on peut avoir est que ce résultat reste vrai si ces intervalles ne sont plus emboîtés. Bien sûr, il ne faut pas prendre des intervalles disjoints. Plus précisément donc, si $(I_n)_{n\in\mathbb{N}}$ est une suite d'intervalles telle que toute intersection *finie* de I_n n'est pas vide, alors en « passant à la limite », l'intersection de *tous* les I_n est non-vide. Cette affirmation est supportée par la figure IV.1. Cette généralisation est vraie car on peut considérer la suite d'intervalles $(J_n)_{n\in\mathbb{N}}$ définie par $J_n := \bigcap_{m\leqslant n} I_m$ qui est une suite d'intervalles fermés (comme intersection d'intervalles fermés), non-vide (par hypothèse, puisqu'il s'agit d'une intersection finie de I_m), emboîtés (par définition) et possède donc une intersection non-vide.

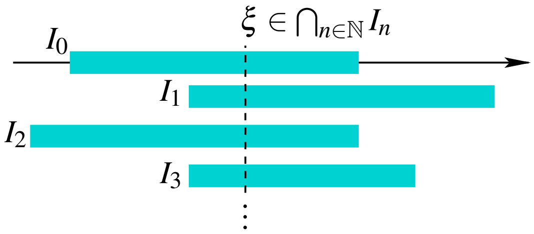


FIGURE IV.1 – Intersection d'intervalles

Il suffit alors de remarquer que $\bigcap_{n\in\mathbb{N}} J_n = \bigcap_{n\in\mathbb{N}} I_n$ (voyez-vous pourquoi ?).

Peut-on généraliser encore la situation et prendre des intervalles I_n ouverts et/ou (semi-)infinis ? Malheureusement non. En effet, considérons les suites d'intervalles $(I_n)_{n\in\mathbb{N}}$ et $(I'_n)_{n\in\mathbb{N}}$ définies par

$$I_n := \left]0, \frac{1}{n+1}\right[\qquad \text{et} \qquad I'_n := \left]-\infty, -n\right]$$

(voir aussi les figures IV.2 et IV.3). Les intersections finies de I_n ou de I'_n sont non-vides (pouvez-vous le montrer ?). De plus, ces intervalles sont emboîtés : $I_{n+1} \subseteq I_n$ et $I'_{n+1} \subseteq I'_n$. Néanmoins $\bigcap_{n\in\mathbb{N}} I_n = \varnothing$ et $\bigcap_{n\in\mathbb{N}} I'_n = \varnothing$. Dans le premier cas, le point d'intersection « s'échappe » par le bord gauche des intervalles tandis que dans le second, il « s'échappe » à l'infini. Pour éviter ce genre de situations, on va avoir

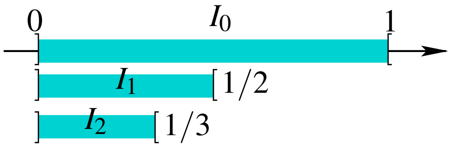


FIGURE IV.2 – $\bigcap_{n\in\mathbb{N}} I_n = \varnothing$

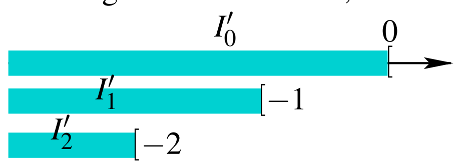


FIGURE IV.3 – $\bigcap_{n\in\mathbb{N}} I'_n = \varnothing$

besoin de mieux contrôler les intersections. Nous allons demander qu'elles aient lieu dans un ensemble C. De plus, nous allons nous restreindre aux fermés (mais pas nécessairement des intervalles) pour éviter le problème de la première suite ci-dessus.

Définition IV.1. Nous dirons qu'un ensemble $C \subseteq \mathbb{R}^N$ possède la *propriété des intersections finies* (PIF) si, quelle que soit la famille de fermés $(F_\alpha)_{\alpha\in A}$ dont les intersections finies sont non-vides au-dessus de C i.e., si

$$\forall B \subseteq A,\ B \text{ fini}, \qquad \Big(\bigcap_{\alpha\in B} F_\alpha\Big) \cap C \neq \varnothing,$$

alors l'intersection de tous ces fermés est non-vide au dessus de C :

$$\Big(\bigcap_{\alpha\in A} F_\alpha\Big) \cap C \neq \varnothing.$$

Tous les ensembles C n'ont pas la PIF. D'après les exemples ci-dessus, on se dit qu'il est nécessaire que C soit borné pour éviter que le point d'intersection ne s'échappe à l'infini et qu'il soit fermé afin que le point d'intersection ne s'échappe pas par le bord de C. Nous verrons que ces conditions — C fermé et borné — sont exactement celles qui caractérisent les ensembles ayant la PIF.

§2. Nous allons maintenant présenter une approche différente qui mène à la même notion. Il arrive souvent en analyse que les propriétés ne soient pas seulement ponctuelles mais *locales*. Par exemple, prenons une fonction $f : \mathbb{R} \to \mathbb{R}$ telle que $f(x) > 0$ pour tout $x \in \mathbb{R}$. On peut dire que :

$$\forall x \in \mathbb{R},\ \exists c_x > 0,\ f(x) \geqslant c_x. \tag{IV.1}$$

En effet, il suffit de prendre $c_x = f(x)$! Comme indiqué par son indice, c_x peut dépendre de x. La contrainte que c_x doit satisfaire, $f(x) \geqslant c_x$, ne dépend que de la valeur de f en x. Pour cette raison, on dira que cette propriété est *ponctuelle*. L'analogue local de (IV.1) est le suivant :

$$\forall x \in \mathbb{R},\ \exists c_x > 0,\ \exists V \text{ voisinage de } x,\ \forall y \in V,\ f(y) \geqslant c_x. \tag{IV.2}$$

De nouveau, c_x peut dépendre de x — mais uniquement de x, pas de V ni de y. Le c_x choisi doit satisfaire une propriété du même type qu'avant, $f(y) \geqslant c_x$, mais, cette fois ci, plus seulement pour $y = x$ mais pour tout y « proche » de x, c'est-à-dire pour tout y dans un (petit) voisinage V de x. Pour cette raison, on dit que c_x doit satisfaire une propriété *locale*.

De manière générale, si $P(x)$ est une propriété qui dépend de $x \in \mathbb{R}^N$, la propriété locale correspondante est

$$P_{\text{loc}}(x) := \exists V_x \text{ voisinage de } x,\ \forall y \in V_x,\ P(y)$$

Revenons à notre exemple : (IV.2) est vraie si f est continue sur $\mathbb{R}$. En effet, étant donné $x \in \mathbb{R}$, la définition de continuité avec $\varepsilon = f(x)/2 > 0$ implique qu'il existe un $\delta > 0$ tel que

$$\forall y \in \mathrm{B}(x, \delta),\ |f(y) - f(x)| \leqslant \varepsilon.$$

Posons $c_x = f(x)/2$ et $V_x = \mathrm{B}(x, \delta)$. Il est facile de déduire (faites le !) de la formule ci-dessus que

$$\forall y \in V_x,\ f(y) \geqslant c_x. \tag{IV.3}$$

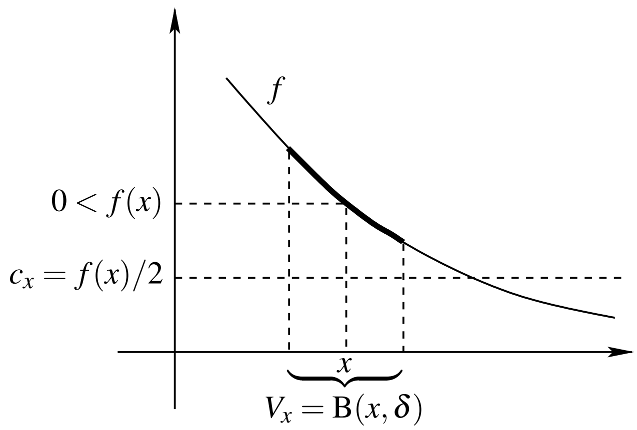


FIGURE IV.4 – $f \geqslant c_x$ localement

Ce raisonnement est aussi représenté sur la figure IV.4. En résumé, la continuité nous permet de passer d'une propriété ponctuelle, $f(x) \geqslant c_x > 0$, à une propriété locale : $f(y) \geqslant c_x > 0$ pour tout y proche de x.

Voulant continuer à étendre le domaine de validité de cette proposition, il est naturel de se demander si on peut trouver un $c > 0$ tel que $f(y) \geqslant c$ soit vrai pour des y pas uniquement proches d'un x fixé. Plus précisément, on voudrait savoir pour quels $C \subseteq \mathbb{R}$ on peut avoir

$$\exists c > 0,\ \forall y \in C,\ f(y) \geqslant c. \tag{IV.4}$$

Ce n'est pas possible quels que soient f et C. En effet, si on prend $C = \mathbb{R}$ et $f(x) = \mathrm{e}^x$, on a que $f(x) > 0$ pour tout $x \in \mathbb{R}$ et pourtant il est faux qu'il existe un $c > 0$ tel que $\forall x \in \mathbb{R},\ f(x) \geqslant c$. S'il en existait un, on aurait $0 = \lim_{x \to -\infty} f(x) \geqslant c$ et donc la contradiction $0 \geqslant c > 0$.

Si on sait que C peut être recouvert par un nombre *fini* de voisinages $V_{x_1}, \dots, V_{x_k}$, alors on peut prendre $c := \min\{c_{x_i} : 1 \leqslant i \leqslant k\}$. En effet, si $y \in C$, alors, puisque $C \subseteq \bigcup_{i=1}^{k} V_{x_i}$, il existe un i tel que $y \in V_{x_i}$ et par (IV.3),

$$f(y) \geqslant c_{x_i} \geqslant c.$$

On peut toujours recouvrir C par un nombre *infini* de V_x : vu que $x \in V_x$, il suffit de prendre tous les V_x pour $x \in C$:

$$C \subseteq \bigcup_{x \in C} V_x.$$

Cependant, pour un nombre infini de V_x, on ne peut reproduire l'argument ci-dessus. Le problème vient du fait que $\min\{c_x : x \in C\}$ pourrait ne pas exister. Si on cherche à contourner la difficulté en définissant $c := \inf\{c_x : x \in C\}$, il se peut que $c = 0$ bien que tous les c_x soient strictement positifs. En vérité, on ne peut espérer adapter l'argument au cas d'une infinité de V_x puisqu'on a vu que bien que $\mathbb{R}$ soit recouvrable par une infinité de V_x vérifiant (IV.3), (IV.4) n'est pas vrai pour $f(x) = \mathrm{e}^x$.

C'est donc que les ensembles C pour lesquels on peut étendre l'argument ont une propriété particulière. Cette propriété c'est justement que d'un recouvrement infini on puisse extraire un sous-recouvrement fini. De tels ensembles sont appelés compacts.

Définition IV.2. Un ensemble $C \subseteq \mathbb{R}^N$ est dit *compact* si de tout recouvrement de C par une famille d'ouverts $(O_\alpha)_{\alpha \in A}$, on peut extraire un sous-recouvrement fini $O_{\alpha_1}, \dots, O_{\alpha_k}$ pour certains $\alpha_1, \dots, \alpha_k \in A$ bien choisis.

Plus symboliquement, si $C \subseteq \bigcup_{\alpha \in A} O_\alpha$ où les O_α sont des ouverts, alors il est possible de trouver un nombre fini $k \in \mathbb{N}$ d'indices $\alpha_1, \dots, \alpha_k \in A$ tels que $C \subseteq \bigcup_{i=1}^k O_{\alpha_i}$.

IV.2 Définitions équivalentes

Dans cette section nous allons montrer que les deux définitions précédentes sont équivalentes entre elles ainsi qu'à deux autres propriétés, la première faisant le lien avec les suites et la seconde permettant de facilement les vérifier en pratique.

Théorème IV.3 (Définitions équivalentes de compacité)**.** *Soit C un sous-ensemble de $\mathbb{R}^N$. Les propriétes suivantes sont équivalentes :*

(i) C possède la propriété des intersections finies ;

(ii) C est compact ;

(iii) C est séquentiellement compact, *c'est-à-dire que, de toute suite $(x_n)_{n \in I} \subseteq C$, on peut extraire une sous-suite $(x'_n)_{n \in I'}$ telle que (x'_n) converge vers un certain $x^* \in C$.*

(iv) C est fermé et borné.

Remarque IV.4. La caractérisation de la compacité par la propriété (iv) n'est valable qu'en dimension finie. Cette caractérisation est appelé le *théorème de Heine–Borel*. En dimension infinie, un compact est nécessairement fermé et borné mais la réciproque n'est pas vraie.

Démonstration. Commençons par montrer que (i) $\Leftrightarrow$ (ii).

(i) $\Rightarrow$ (ii). Supposons que C ait la propriété des intersections finies et montrons qu'il est compact. Soit donc $(O_\alpha)_{\alpha\in A}$ un recouvrement ouvert de C. Argumentons par contradiction et supposons au contraire qu'on ne puisse trouver de sous-recouvrement fini adéquat, c'est-à-dire que

$$\forall B \subseteq A,\ B \text{ fini}, \quad \bigcup_{\alpha\in B} O_\alpha \text{ ne recouvre pas } C.$$

Le fait que l'union ne recouvre pas C veut dire qu'au moins un point de C n'est pas dans l'union. La propriété précédente est donc équivalente à

$$\forall B \subseteq A,\ B \text{ fini},\ \exists x \in C, \quad x \notin \bigcup_{\alpha\in B} O_\alpha.$$

En passant au complémentaire, on a une famille de fermés $(\complement O_\alpha)_{\alpha\in A}$ telle que

$$\forall B \subseteq A,\ B \text{ fini},\ \exists x \in C, \quad x \notin \bigcup_{\alpha\in B} \complement\complement O_\alpha = \complement \bigcap_{\alpha\in B} \complement O_\alpha.$$

ou encore, vu que la non-appartenance au complémentaire équivaut à l'appartenance à l'ensemble,

$$\forall B \subseteq A,\ B \text{ fini}, \quad C \cap \bigcap_{\alpha\in B} \complement O_\alpha \neq \varnothing.$$

Autrement dit, les intersections finies des $(\complement O_\alpha)_{\alpha\in A}$ sont non-vides au dessus de C et (i) implique que

$$C \cap \bigcap_{\alpha\in A} \complement O_\alpha \neq \varnothing$$

ou encore que

$$\exists x \in C, \quad x \in \bigcap_{\alpha\in A} \complement O_\alpha = \complement \bigcup_{\alpha\in A} O_\alpha.$$

Dès lors, ce $x \in C$ ne peut appartenir à $\bigcup_{\alpha\in A} O_\alpha$ ce qui veut dire que $(O_\alpha)_{\alpha\in A}$ n'est pas un recouvrement de C contrairement à l'hypothèse. C'est la contradiction recherchée.

(ii) $\Rightarrow$ (i). Le principe — argument par contradiction et passage au complémentaire — est le même que pour (i) $\Rightarrow$ (ii). Les détails sont laissés au lecteur.

Nous allons maintenant prouver que (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv) $\Rightarrow$ (ii)

(ii) $\Rightarrow$ (iii). Supposons que C soit compact et montrons qu'il est séquentiellement compact. Soit $(x_n)_{n\in I} \subseteq C$. Supposons par contradiction qu'aucune sous-suite de $(x_n)_{n\in I}$ ne converge vers un élément de C, c'est-à-dire que, quel que soit $x^* \in C$, il ne peut être la limite d'aucune sous-suite de $(x_n)_{n\in I}$:

$$\forall x^* \in C,\ \forall (x'_n) \subseteq (x_n), \quad x'_n \not\to x^*. \tag{IV.5}$$

Montrons que cela implique que (x_n) reste ultimement à une certaine distance de x^* :

$$\forall x^* \in C,\ \exists \varepsilon = \varepsilon(x^*) > 0,\ \exists n_0 = n_0(x^*),\ \forall n \geqslant n_0, \quad x_n \notin \mathrm{B}(x^*, \varepsilon). \tag{IV.6}$$

En effet, soit $x^* \in C$ et supposons au contraire que

$$\forall \varepsilon > 0,\ \forall n_0,\ \exists n \geqslant n_0, \quad x_n \in \mathrm{B}(x^*, \varepsilon).$$

En prenant $\varepsilon = 1$ et $n_0 = 0$, on trouve un n_1 tel que $x_{n_1} \in \mathrm{B}(x^*, 1)$. On considère ensuite $\varepsilon = 1/2$ et $n_0 = n_1 + 1$ et on trouve un $n_2 > n_1$ tel que $x_{n_2} \in \mathrm{B}(x^*, 1/2)$. On continuant de la sorte, on trouve $n_1 < n_2 < n_3 < \cdots < n_k < \cdots$ tels que $x_{n_k} \in \mathrm{B}(x^*, 1/k)$ pour tout $k \geqslant 1$. Autrement dit, $(x_{n_k})_{k\geqslant 1}$ est une sous-suite de $(x_n)_{n\in I}$ telle que $\|x_{n_k} - x^*\| < 1/k$. Par la convergence dominée, $x_{n_k} \xrightarrow[k\to\infty]{} x^*$ ce qui contredit (IV.5). L'affirmation (IV.6) est donc bien vraie.

En utilisant (IV.6) et la compacité de C, nous allons maintenant aboutir à une contradiction, ce qui montrera qu'on ne pouvait supposer (IV.5). Considérons $\big(\mathrm{B}(x^*, \varepsilon(x^*)) : x^* \in C\big)$. Il s'agit d'une famille d'ouverts. Elle recouvre C car, si $x \in C$, alors $x \in \mathrm{B}(x, \varepsilon(x)) \subseteq \bigcup_{x^*\in C} \mathrm{B}(x^*, \varepsilon(x^*))$. Puisque C est compact, un nombre fini des ces ouverts suffit à le recouvrir :

$$C \subseteq \mathrm{B}(x_1^*, \varepsilon(x_1^*)) \cup \cdots \cup \mathrm{B}(x_k^*, \varepsilon(x_k^*)). \tag{IV.7}$$

La propriété (IV.6) appliquée à ces x_j^* donne :

$$\forall j = 1, \ldots, k,\ \forall n \geqslant n_0(x_j^*), \quad x_n \notin \mathrm{B}(x_j^*, \varepsilon(x_j^*)).$$

Posons $n^* := \max\{n_0(x_j^*) : j = 1, \dots, k\}$. Puisque $n \geqslant n^* \Rightarrow n \geqslant n_0(x_j^*)$, l'affirmation précédente implique que

$$\forall j = 1, \dots, k,\ \forall n \geqslant n^*, \quad x_n \notin \mathrm{B}(x_j^*, \varepsilon(x_j^*))$$

ou encore que

$$\forall n \geqslant n^*, \quad x_n \notin \bigcup_{j=1}^{k} \mathrm{B}(x_j^*, \varepsilon(x_j^*)).$$

En particulier, x_{n^*} ne peut appartenir à l'union des boules et donc pas non plus à C au vu de (IV.7). Ceci contredit le fait qu'on avait choisi une suite $(x_n) \subseteq C$.

(iii) $\Rightarrow$ (iv). Supposons que C soit séquentiellement compact et montrons qu'il est fermé et qu'il est borné.

Pour voir que C est fermé, prenons une suite $(x_n) \subseteq C$ qui converge vers un $a \in \mathbb{R}^N$ et prouvons que $a \in C$. Vu que C est séquentiellement compact, il existe une sous-suite $(x'_n) \subseteq (x_n)$ et $x^* \in C$ tel que $x'_n \to x^*$. Vu que $x_n \to a$, on a $x'_n \to a$. Dès lors, de l'unicité de la limite, on déduit $a = x^* \in C$.

Pour montrer que C est borné, procédons par l'absurde. Si C n'est pas borné, on a

$$\forall R > 0,\ \exists x \in C, \quad \|x\| \geqslant R.$$

En particulier, en prenant successivement $R = 1, 2, \dots$, on obtient une suite $(x_n)_{n \geqslant 1}$ dans C telle que

$$\forall n \in \mathbb{N}^{\geqslant 1}, \quad \|x_n\| \geqslant n. \tag{IV.8}$$

Puisque C est séquentiellement compact, il existe une sous-suite $(x_{n_k})_k \subseteq (x_n)$ et $x^* \in C$ tels que $x_{n_k} \xrightarrow[k\to\infty]{} x^*$. Par (IV.8), on a

$$\|x_{n_k}\| \geqslant n_k \xrightarrow[k\to\infty]{} +\infty.$$

ce qui contredit la convergence de (x_{n_k}) vers $x^* \in \mathbb{R}^N$. □

Avant de nous attaquer à la dernière partie de la preuve, à savoir (iv) $\Rightarrow$ (ii), prouvons un lemme qui nous sera nécessaire.

Lemme IV.5. *Soit $x \in \mathbb{R}^N$ et $R > 0$. On peut trouver un ensemble fini $P \subseteq \mathrm{B}_\infty[x, R]$ tel que*

$$\mathrm{B}_\infty[x, R] = \bigcup_{y \in P} \mathrm{B}_\infty[y, R/2].$$

En fait on peut choisir P pour qu'il ait 2^N éléments.

Graphiquement, l'idée est très simple. La figure IV.5 en témoigne à deux dimensions : la boule $B_\infty[x,R]$ est simplement un carré de coté $2R$ centré en x et on peut le recouvrir à l'aide de 4 carrés de coté R, c'est-à-dire 2^2 boules $B_\infty[y,R/2]$ pour certains y bien choisis.

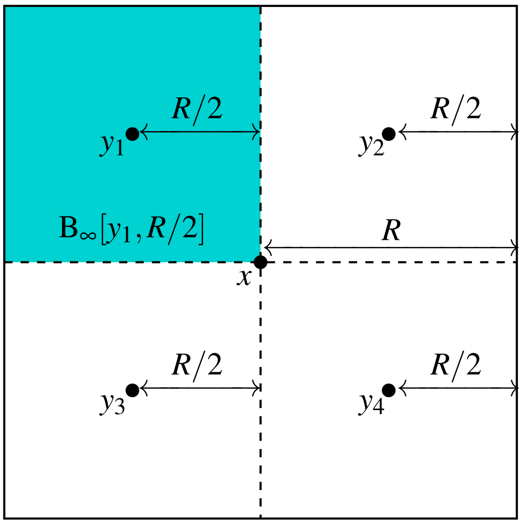


FIGURE IV.5 – Recouvrement de $B_\infty[x,R]$ par des $B_\infty[y,R/2]$

Démonstration. Considérons l'ensemble

$$P := \big\{x+(\sigma_1,\dots,\sigma_N)R/2 : \sigma_i \in \{-1,+1\} \text{ pour } i=1,\dots,N\big\}.$$

Cet ensemble a 2^N éléments car il y a deux choix possibles pour σ_1, deux choix pour σ_2,..., et deux choix pour σ_N. D'autre part, si $y \in P$, alors

$$|y-x|_\infty = \big|(\sigma_1,\dots,\sigma_N)R/2\big|_\infty = \max_{1\leqslant i\leqslant N} |\sigma_i R/2| = R/2.$$

Dès lors, si $z \in B_\infty[y,R/2]$ pour un $y \in P$, $|z-x|_\infty \leqslant |z-y|_\infty + |y-x|_\infty \leqslant R/2 + R/2 = R$. Ceci montre que $B_\infty[y,R/2] \subseteq B_\infty[x,R]$ quel que soit $y \in P$ et donc que

$$\bigcup_{y\in P} B_\infty[y,R/2] \subseteq B_\infty[x,R].$$

Reste à voir l'inclusion inverse. Soit $z \in B_\infty[x,R]$. Il faut montrer qu'il existe un $y \in P$ tel que $z \in B_\infty[y,R/2]$. Choisissons $y = x+(\sigma_1,\dots,\sigma_N)R/2$ avec les σ_i

déterminés à l'aide de la règle suivante :

$$\sigma_i = \begin{cases} +1 & \text{si } z_i - x_i \geqslant 0 \\ -1 & \text{si } z_i - x_i < 0 \end{cases}$$

où z_i et x_i désignent la i^{e} composante de z et x respectivement. (Pour comprendre ceci, prenez divers z sur la figure IV.5.) On doit montrer que $|z-y|_\infty \leqslant R/2$, c'est-à-dire $|z_i - y_i| \leqslant R/2$ pour tout $i = 1, \dots, N$. Soit i arbitraire. Par hypothèse $z \in \mathrm{B}_\infty[x,R]$ et donc $|z_i - x_i| \leqslant R$ ou encore $-R \leqslant z_i - x_i \leqslant R$. Distinguons deux cas :

- $z_i - x_i \geqslant 0$ auquel cas $\sigma_i = +1$, d'où $y_i = x_i + R/2$ et $0 \leqslant z_i - x_i \leqslant R$. Alors

  $$-R/2 \leqslant z_i - y_i = z_i - x_i - R/2 \leqslant R/2$$

  ou encore $|z_i - y_i| \leqslant R/2$.
- $z_i - x_i < 0$ auquel cas $\sigma_i = -1$, d'où $y_i = x_i - R/2$ et $-R \leqslant z_i - x_i < 0$. Dès lors

  $$-R/2 \leqslant z_i - y_i = z_i - x_i + R/2 < R/2$$

  ce qui implique $|z_i - y_i| \leqslant R/2$. □

Démonstration du théorème IV.3 (suite). La preuve de (iv) ⇒ (ii) s'effectue en deux étapes. Nous allons d'abord montrer que les boules $\mathrm{B}_\infty[x,R]$ sont compactes et ensuite en déduire (iv) ⇒ (ii).

$\mathrm{B}_\infty[x,R]$ *est compact.* Soit $(O_\alpha)_{\alpha \in A}$ un recouvrement ouvert de $\mathrm{B}_\infty[x,R]$. Il faut montrer qu'on peut en extraire un sous-recouvrement fini. Supposons au contraire que ce ne soit pas le cas, c'est-à-dire que si on prend un nombre fini de O_α, on ne recouvre jamais $\mathrm{B}_\infty[x,R]$. Par le lemme IV.5, on peut écrire

$$\mathrm{B}_\infty[x,R] = \bigcup_{y \in P_1} \mathrm{B}_\infty[y, R/2]$$

pour un certain ensemble fini $P_1 \subseteq \mathrm{B}_\infty[x,R]$. Au moins une des boules $\mathrm{B}_\infty[y,R/2]$ ne peut être recouverte par un nombre fini de O_α. En effet, si chacune des boules $\mathrm{B}_\infty[y,R/2]$, $y \in P_1$, pouvait être recouverte par un nombre fini de O_α, en prenant tous ces O_α (qui sont en nombre fini puisqu'il y en a un nombre fini par boule et un nombre fini de boules), on recouvrirait $\bigcup_{y \in P_1} \mathrm{B}_\infty[y,R/2] = \mathrm{B}_\infty[x,R]$, ce qu'on avait supposé ne pas pouvoir faire. Notons $\mathrm{B}_\infty[y_1, R/2]$ une des boules non-recouvrables par un nombre fini de O_α.

En recommençant le même raisonnement avec $\mathrm{B}_\infty[y_1, R/2] = \bigcup_{y\in P_2} \mathrm{B}_\infty[y, R/4]$ au lieu de $\mathrm{B}_\infty[x,R]$, on trouve un $y_2 \in P_2$ tel que $\mathrm{B}_\infty[y_2, R/4] \subseteq \mathrm{B}_\infty[y_1, R/2]$ ne peut être recouvert par un nombre fini de O_α.

On continue par récurrence. Étant donné $\mathrm{B}_\infty[y_n, R/2^n]$ qui n'est pas recouvrable par un nombre fini de O_α, on trouve un y_{n+1} tel que $\mathrm{B}_\infty[y_{n+1}, R/2^{n+1}] \subseteq \mathrm{B}_\infty[y_n, R/2^n]$ ne soit pas recouvrable par un nombre fini de O_α.

Montrons que la suite $(y_n)_{n\geqslant 1}$ est de Cauchy. Soit $\varepsilon > 0$. De la convergence $R/2^n \to 0$, on déduit qu'il existe un $n_0 \in \mathbb{N}$ tel que $\forall n \geqslant n_0,\ R/2^n \leqslant \varepsilon$. Montrons que ce n_0 convient dans la définition de suite de Cauchy, c'est-à-dire que, si $m \geqslant n \geqslant n_0$, alors $|y_m - y_n| \leqslant \varepsilon$. Soit $m \geqslant n \geqslant n_0$. Vu que $y_m \in \mathrm{B}_\infty[y_m, R/2^m] \subseteq \mathrm{B}_\infty[y_{m-1}, R/2^{m-1}] \subseteq \cdots \subseteq \mathrm{B}_\infty[y_n, R/2^n]$, on a que $|y_m - y_n|_\infty \leqslant R/2^n \leqslant \varepsilon$.

Donc $(y_n)_{n\geqslant 1}$ est de Cauchy et, puisque $\mathbb{R}^N$ est complet, il existe un $y^* \in \mathbb{R}^N$ tel que $y_n \to y^*$. Comme on vient de le voir, si $m \geqslant n \geqslant 1$, alors $y_m \in \mathrm{B}_\infty[y_n, R/2^n]$. Donc la sous-suite $(y_m)_{m\geqslant n} \subseteq (y_n)_{n\geqslant 1}$ est dans $\mathrm{B}_\infty[y_n, R/2^n]$. Cette sous-suite convergeant vers y^* et $\mathrm{B}_\infty[y_n, R/2^n]$ étant fermé, on a $y^* \in \mathrm{B}_\infty[y_n, R/2^n]$. Ce raisonnement est valable quel que soit $n \geqslant 1$ et par conséquent

$$y^* \in \bigcap_{n\geqslant 1} \mathrm{B}_\infty[y_n, R/2^n].$$

En particulier, $y^* \in \mathrm{B}_\infty[y_1, R/2] \subseteq \mathrm{B}_\infty[x,R]$. Puisque $(O_\alpha)_{\alpha\in A}$ est un recouvrement de $\mathrm{B}_\infty[x,R]$, il existe au moins un $\alpha^* \in A$ tel que

$$y^* \in O_{\alpha^*}.$$

L'ensemble O_{α^*}étant ouvert, il existe un $\delta > 0$ tel que

$$\mathrm{B}_\infty(y^*, \delta) \subseteq O_{\alpha^*}.$$

Comme $y_n \to y^*$, dès que n est assez grand, disons $n \geqslant n_1$, on a $|y_n - y^*|_\infty \leqslant \delta/3$. Comme $R/2^n \to 0$, on sait qu'il existe un n_2 tel que, si $n \geqslant n_2$, $R/2^n \leqslant \delta/3$. Dès lors, si $n \geqslant \max\{n_1, n_2\}$,

$$\mathrm{B}_\infty[y_n, R/2^n] \subseteq \mathrm{B}_\infty(y^*, \delta).$$

En effet, si $z \in \mathrm{B}[y_n, R/2^n]$, alors $|z - y^*|_\infty \leqslant |z - y_n|_\infty + |y_n - y^*|_\infty \leqslant R/2^n + \delta/3 \leqslant \delta/3 + \delta/3 < \delta$. Mais alors, $\mathrm{B}_\infty[y_n, R/2^n] \subseteq O_{\alpha^*}$ ce qui veut dire que $\mathrm{B}_\infty[y_n, R/2^n]$ est recouvrable par un seul ouvert O_{α^*}. Ceci est en contradiction avec la manière dont on a construit $\mathrm{B}_\infty[y_n, R/2^n]$ qui garantissait qu'il n'était pas recouvrable par un nombre fini de O_α.

En conclusion, on ne pouvait pas supposer qu'il était impossible de recouvrir $\mathrm{B}_\infty[x,R]$ par un nombre fini de O_α. Ceci termine l'argument établissant la compacité de $\mathrm{B}_\infty[x,R]$.

(iv) ⇒ (ii). Montrons que si C est fermé et borné, alors C est compact. Soit $(O_\alpha)_{\alpha\in A}$ un recouvrement de C. Puisque C est borné, il existe un $R > 0$ tel que $C \subseteq \mathrm{B}_\infty[x,R]$. Posons $O' := \mathbb{R}^N \setminus C = \complement C$. C'est un ouvert. De plus

$$\mathrm{B}_\infty[x,R] \subseteq O' \cup \bigcup_{\alpha\in A} O_\alpha$$

puisque, si $y \in \mathrm{B}_\infty[x,R]$, soit $y \in C$ auquel cas il est dans $\bigcup_{\alpha\in A} O_\alpha$, soit $y \notin C$ et donc $y \in O'$. Comme $\mathrm{B}_\infty[x,R]$ est compact, on peut extraire un sous-recouvrement fini de $(O', O_\alpha : \alpha \in A)$, c'est-à-dire qu'il existe $\alpha_1, \dots, \alpha_k \in A$ tels que

$$\mathrm{B}_\infty[x,R] \subseteq O_{\alpha_1} \cup \dots \cup O_{\alpha_k} \quad \text{ou} \quad \mathrm{B}_\infty[x,R] \subseteq O' \cup O_{\alpha_1} \cup \dots \cup O_{\alpha_k}.$$

Dans le premier cas, on a fini puisque $C \subseteq \mathrm{B}_\infty[x,R]$. Dans le second, on va montrer que

$$C \subseteq O_{\alpha_1} \cup \dots \cup O_{\alpha_k}.$$

Quel que soit $y \in C$, on a $y \in \mathrm{B}_\infty[x,R]$ et donc $y \in O_{\alpha_i}$ pour un i, ce qui est ce qu'on veut, ou $y \in O' = \complement C$, ce qui est impossible vu que $y \in C$. □

IV.3 Théorème des bornes atteintes

Dans cette section, nous allons montrer qu'une application continue sur un compact possède une valeur maximale et une valeur minimale. Nous allons donner plusieurs démonstrations de ce fait afin de voir les diverses formulations de la compacité en œuvre.

Théorème IV.6 (Théorème des bornes atteintes)**.** *Soit $C \subseteq \mathbb{R}^N$ un compact non-vide et $f : C \to \mathbb{R}$ une application continue. Alors f atteint ses bornes, c'est-à-dire qu'il existe au moins un $x_{\min} \in C$ et un $x_{\max} \in C$ tels que, pour tout $x \in C$, $f(x_{\min}) \leqslant f(x) \leqslant f(x_{\max})$.*

Remarque IV.7. ▪ On peut de manière équivalente dire qu'une application $f: C \to \mathbb{R}$ atteint ses bornes si $\min f(C)$ et $\max f(C)$ existent — et en fait, avec les notations du théorème, on a $f(x_{\min}) = \min f(C)$ et $f(x_{\max}) = \max f(C)$.

- L'hypothèse de compacité ne peut être enlevée. En effet, considérons l'application $f:]0,1] \to \mathbb{R} : x \mapsto 1/x$. Cette application est continue et pourtant elle n'atteint pas son maximum vu que $\sup f(]0,1]) = \sup[1,+\infty[= +\infty$. (Atteint-elle son minimum ?)

Toutes les démonstrations qui suivent vont seulement montrer le fait que f atteint son maximum. Les démonstrations pour le minimum sont similaires — on peut aussi utiliser la relation $\min f = -\max(-f)$.

Démonstration utilisant la PIF. Notons $s := \sup f(C) \in [-\infty, +\infty]$. Puisque $C \neq \varnothing$, $f(C) \neq \varnothing$ et donc $s \neq -\infty$. Pour $n \in \mathbb{N} \setminus \{0\}$, posons

$$\rho_n := \begin{cases} s - 1/n & \text{si } s \in \mathbb{R} \\ n & \text{si } s = +\infty \end{cases}$$

$$F_n := \{x \in C : f(x) \geqslant \rho_n\}.$$

Les ensembles F_n sont fermés. En effet, si $(x_k) \subseteq F_n$ converge vers x^*, alors en passant à la limite sur $\forall k,\ f(x_k) \geqslant \rho_n$ et en utilisant la continuité de f, on obtient $f(x^*) = f(\lim x_k) = \lim f(x_k) \geqslant \rho_n$ et donc $x^* \in F_n$.

De plus, aucun F_n n'est vide. En effet, puisque $\rho_n < \sup f(C)$, on sait qu'il existe un élément de $f(C)$, c'est-à-dire un $f(x)$ pour un $x \in C$, tel que $f(x) \geqslant \rho_n$ (propositions I.16 (iii) et I.18 (iii)). Ce x appartient à F_n.

Par ailleurs, comme les ρ_n sont croissants, les ensembles F_n sont décroissants : $F_1 \supseteq F_2 \supseteq F_3 \supseteq \cdots$ Si on fait l'intersection d'un nombre fini d'entre eux, disons $F_{n_1}, F_{n_2}, \ldots, F_{n_k}$, on a

$$\bigcap_{i=1}^{k} F_{n_i} = F_{\max\{n_1,\ldots,n_k\}}$$

(voyez-vous pourquoi ? pouvez-vous le montrer ?). Par conséquent, les intersections finies des F_n sont non-vides au dessus de C (puisque $F_{\max\{n_1,\ldots,n_k\}} \subseteq C$).

En vertu de la PIF, l'intersection de tous les F_n est non-vide au dessus de C, c'est-à-dire qu'il existe un $x_{\max} \in C$ tel que

$$x_{\max} \in \bigcap_{n=1}^{\infty} F_n.$$

Reste à montrer que $f(x_{\max})$ est bien le maximum de $f(C)$. Comme $x_{\max} \in F_n$ veut dire que $f(x_{\max}) \geqslant \rho_n$, en passant à la limite on a $f(x_{\max}) \geqslant \lim \rho_n = \sup f(C)$. Donc $f(x_{\max}) = \sup f(C)$ et c'est bien le maximum (proposition I.21). □

Démonstration utilisant la définition de compacité. Supposons au contraire que f n'atteigne pas son maximum sur C, c'est-à-dire que

$$\forall x \in C,\ \exists y \in C, \quad f(x) < f(y) \tag{IV.9}$$

Montrons que le y ne marche pas seulement pour x mais pour un voisinage de x, c'est-à-dire :

$$\forall x \in C,\ \exists y \in C,\ \exists V_x \text{ voisinage ouvert de } x,\ \forall x' \in V_x, \quad f(x') < f(y). \tag{IV.10}$$

En effet, soit $x \in C$. Prenons un $y \in C$ dont l'existence est affirmée par (IV.9). En considérant $\varepsilon = (f(y) - f(x))/2 > 0$ dans la définition de continuité de f au point x, on trouve qu'il existe un $\delta > 0$ tel que

$$\forall x' \in \mathrm{B}(x,\delta), \quad |f(x') - f(x)| \leqslant \varepsilon.$$

Posons $V_x := \mathrm{B}(x,\delta)$ qui est un voisinage ouvert de x. Si $x' \in V_x$, alors $f(x') \leqslant f(x) + \varepsilon = (f(x) + f(y))/2 < f(y)$ et on a bien prouvé (IV.10).

La famille $(V_x)_{x\in C}$ est un recouvrement ouvert de C. En effet, quel que soit $x \in C$, $x \in V_x$ et donc

$$C \subseteq \bigcup_{x\in C} V_x.$$

La compacité de C implique alors qu'il suffit d'un nombre fini d'ouverts pour recouvrir C i.e.,

$$C \subseteq V_{x_1} \cup \cdots \cup V_{x_n}$$

pour certains $x_1, \ldots, x_n \in C$. Puisque $C \neq \varnothing$, on a $n \geqslant 1$. Notons $y_1, \ldots, y_n$ des y correspondant à chacun de ces x_i par (IV.10). Parmi les $f(y_i)$, $1 \leqslant i \leqslant n$, il y en a un qui est le plus grand, disons que c'est $f(y_j)$:

$$f(y_j) = \max_{1\leqslant i\leqslant n} f(y_i).$$

Ce y_j appartient à C qui est recouvert par les V_{x_i}, $1 \leqslant i \leqslant n$, donc $y_j \in V_{x_i}$ pour un certain i. Mais ce fait combiné avec la propriété (IV.10) implique que

$$f(y_j) < f(y_i) \leqslant \max_{1\leqslant i\leqslant n} f(y_i) = f(y_j).$$

Cette contradiction montre qu'on ne pouvait pas supposer que f n'atteignait pas son maximum. $\square$

La démonstration suivante porte le nom de méthode directe du calcul des variations parce qu'on recherche le maximum directement, comme limite d'une suite bien construite.

Démonstration basée sur la compacité séquentielle. Posons $s := \sup f(C)$. Comme $C \neq \varnothing$, on a $f(C) \neq \varnothing$ et donc $s \neq -\infty$. Les propriétés du supremum (propositions I.16 (ii) et I.18 (ii)) impliquent qu'il existe une suite $(s_n)_{n\in I} \subseteq f(C)$ telle que $s_n \to s$. Par définition de $f(C)$, chaque élément s_n s'écrit comme $f(x_n)$ pour un certain $x_n \in C$. Puisque C est séquentiellement compact, il existe une sous-suite (x'_n) de (x_n) qui converge vers un $x^* \in C$. Comme f est continue, $f(x'_n) \to f(x^*)$. Puisque $\big(f(x'_n)\big)$ est une sous-suite de $\big(f(x_n)\big) = (s_n)$, on a $f(x'_n) \to s$. L'unicité de la limite implique alors

$$f(x^*) = s = \sup f(C)$$

et donc, vu que $f(x^*) \in f(C)$, on a $f(x^*) = \max f(C)$. $\square$

IV.4 Équivalence des normes en dimension finie

Nous avons affirmé à la page 51 que toutes les normes sur $\mathbb{R}^N$ étaient équivalentes. Nous avons maintenant les outils pour le démontrer. Heureusement, jusqu'à présent nous n'avons pas utilisé le fait que *toutes* les normes sur $\mathbb{R}^N$ étaient équivalentes. Ce que nous avons dit est que certaines propriétés ne dépendaient pas de la norme tant qu'on se restreignait à des normes équivalentes et nous avons à plusieurs reprises particularisé la norme à $|\cdot|_1$, $|\cdot|_2$ ou $|\cdot|_\infty$. En d'autres termes, nous avons jusqu'à présent travaillé avec toutes les normes équivalentes à $|\cdot|_1$, $|\cdot|_2$ et $|\cdot|_\infty$ (dont on sait qu'elles sont équivalentes, cf. proposition II.9). Ce que nous allons maintenant prouver est que, si $\|\cdot\|$ est une norme quelconque sur $\mathbb{R}^N$ dont nous ne savons pas *à priori* si elle est équivalente à $|\cdot|_1$, $|\cdot|_2$ et $|\cdot|_\infty$ et donc pour laquelle nous ne sommes pas sûrs que les théorèmes précédents soient valides, cette norme $\|\cdot\|$ n'a en réalité pas d'autre choix que d'être équivalente à $|\cdot|_1$, $|\cdot|_2$ et $|\cdot|_\infty$ — ce qui nous permet de dire *à posteriori* que les théorèmes prouvés pour les normes équivalentes à $|\cdot|_1$, $|\cdot|_2$ et $|\cdot|_\infty$ sont valables pour toutes les normes.

Théorème IV.8 (Équivalence des normes en dimension finie)**.** *Toutes les normes sur* $\mathbb{R}^N$ *sont équivalentes.*

Démonstration. Soit $\|\cdot\| : \mathbb{R}^N \to [0, +\infty[$ une norme arbitraire sur $\mathbb{R}^N$. Nous allons montrer que $\|\cdot\|$ est équivalente à $|\cdot|_1$, c'est-à-dire qu'il existe des constantes $C \geqslant 0$ et $C' \geqslant 0$ telles que

$$\forall x \in \mathbb{R}^N, \qquad \|x\| \leqslant C|x|_1 \quad \text{et} \quad |x|_1 \leqslant C'\|x\|. \tag{IV.11}$$

La première inégalité est facile à prouver. Notons $(e_1, \ldots, e_N)$ la base canonique de $\mathbb{R}^N$, c'est-à-dire que $e_i \in \mathbb{R}^N$ est le vecteur dont toutes les composantes sont nulles sauf la i^e qui vaut 1. Si $x = (x_1, \ldots, x_N) \in \mathbb{R}^N$, on peut écrire $x = \sum_{i=1}^N x_i e_i$ et donc en utilisant les propriétés des normes :

$$\|x\| \leqslant \sum_{i=1}^N |x_i|\, \|e_i\| \leqslant C|x|_1$$

où $C := \max\{\|e_i\| : 1 \leqslant i \leqslant N\}$. Ceci établit aussi que $\|\cdot\|$ est continue (par rapport à $|\cdot|_1$) : quelle que soit $(x_n) \subseteq \mathbb{R}^N$ et $x^* \in \mathbb{R}^N$,

$$x_n \xrightarrow[n\to\infty]{|\cdot|_1} x^* \;\Rightarrow\; \|x_n\| \xrightarrow[n\to\infty]{\mathbb{R}} \|x^*\|. \tag{IV.12}$$

En effet, $\big|\|x_n\| - \|x^*\|\big| \leqslant \|x_n - x^*\| \leqslant C|x_n - x^*|_1 \to 0$ et une simple application de la convergence dominée donne le résultat.

Passons maintenant à la preuve de la seconde inégalité de (IV.11). Procédons par l'absurde et supposons qu'il n'existe pas de C' qui marche pour tous les x, c'est-à-dire :

$$\forall C' \geqslant 0,\ \exists x \in \mathbb{R}^N, \quad |x|_1 > C'\|x\|.$$

Pour chaque $n \in \mathbb{N}$, en considérant $C' = n$ dans cette proposition, on trouve qu'il existe un $x_n \in \mathbb{R}^N$ tel que $|x_n|_1 > n\|x_n\|$. Posons $\xi_n := x_n/|x_n|_1$. Nous avons que

$$|\xi_n|_1 = 1 \qquad \text{et} \qquad \|\xi_n\| = \frac{\|x_n\|}{|x_n|_1} < \frac{1}{n}.$$

Comme la sphère unité $\mathbb{S}_{|\cdot|_1}(0,1) := \{x \in \mathbb{R}^N : |x|_1 = 1\}$ est compacte (on vérifie facilement qu'elle est fermée est bornée), on peut trouver une sous-suite (ξ'_n) de (ξ_n) et un $\xi^* \in \mathbb{S}_{|\cdot|_1}(0,1)$ tels que $\xi'_n \xrightarrow{|\cdot|_1} \xi^*$. En particulier $|\xi^*|_1 = 1$. Par ailleurs, (IV.12) implique que

$$\|\xi'_n\| \to \|\xi^*\|.$$

Mais de $\|\xi_n\| < 1/n \to 0$ on déduit que $\xi_n \to 0$ et donc $\xi'_n \to 0$. Par unicité de la limite $\|\xi^*\| = 0$. Cela implique que $\xi^* = 0$ et contredit la conclusion antérieure que $|\xi^*|_1 = 1$. Ceci est la contradiction recherchée. □

IV.5 Exercices

Exercice IV.1. Prouvez que tout sous-ensemble fini de $\mathbb{R}$ est compact.

Exercice IV.2. Si C est un compact et F est un fermé, alors $C \cap F$ est un compact.

Exercice IV.3. Montrez que $[0,1] \times [0,1]$ est un compact de $\mathbb{R}^2$. Plus généralement, si $C_1 \subseteq \mathbb{R}^{N_1}$ et $C_2 \subseteq \mathbb{R}^{N_*2}$ sont deux compacts alors $C_1 \times C_2 \subseteq \mathbb{R}^{N_1} \times \mathbb{R}^{N_2}$ est compact.

Exercice IV.4. Si C_1 et C_2 sont deux compacts de $\mathbb{R}^N$, alors $C_1 \cup C_2$, $C_1 \cap C_2$ et $C_1 + C_2 := \{x_1 + x_2 : x_1 \in C_1,\ x_2 \in C_2\}$ sont encore des compacts de $\mathbb{R}^N$.

Exercice IV.5. Soit A un sous-ensemble fini de $\mathbb{R}$. Prouvez qu'il est impossible que

$$\forall y \in A,\ \exists \eta \in A, \quad \eta > y.$$

Exercice IV.6 (août 2007). Soient $a, b \in \mathbb{R}$ et $f : [a,b] \to \mathbb{R}$ une fonction continue. Montrez que

$$M := \Big\{x \in [a,b] : f(x) = \max_{\xi \in [a,b]} f(\xi)\Big\}$$

est un ensemble compact.

Exercice IV.7. Soit $f : [0,1] \to \mathbb{R}$. Prouvez que, si f est continue, il est impossible que

$$\forall x \in [0,1],\ \exists \xi \in [0,1], \quad f(\xi) > f(x).$$

Cela reste-t-il vrai pour $f :]0,1[\to \mathbb{R}$?

Exercice IV.8. Si $f : \mathbb{R}^N \to \mathbb{R}^M$ est une fonction continue et $C \subseteq \mathbb{R}^N$ est compact, alors $f(C)$ est compact.

Exercice IV.9. Montrez par un contre-exemple que l'image inverse d'un compact par une fonction, même continue, n'est pas nécessairement compact.

Exercice IV.10. Soit $f : [0,1] \to \mathbb{R}$ une fonction continue.

- Montrez qu'il existe deux nombres $a \leqslant b$ tels que, pour tout $x \in [0,1]$, $a \leqslant f(x) \leqslant b$.
- Donnez une condition nécessaire et suffisante sur f (et l'interpréter en français) pour qu'on puisse choisir $a = b$.
- Aurait-on pu prouver les mêmes choses pour $f :]0,1[\to \mathbb{R}$? (Preuve ou contre-exemple.)

Exercice IV.11. Soit $f : C \subseteq \mathbb{R}^N \to \mathbb{R}$ une fonction continue définie sur un compact C. Prouvez que, si $\forall x \in C, f(x) > 0$, alors il existe un $c > 0$, tel que $\forall x \in C,\ f(x) \geqslant c$. Trouvez des contre-exemples qui montrent que ce n'est pas vrai si on omet l'hypothèse de continuité ou de compacité.

Exercice IV.12. Soient $A \subseteq \mathbb{R}^N$ et $B \subseteq \mathbb{R}^M$ deux ensembles fermés. Désignons par $p : \mathbb{R}^N \times \mathbb{R}^M \to \mathbb{R}^N : (x,y) \mapsto x$ la projection sur la première composante du produit $\mathbb{R}^N \times \mathbb{R}^M$. On suppose que A est assez riche dans le sens où il existe $X \subseteq A$ tel que X est non-fermé. Prouvez que

$$\text{pour tout ensemble fermé } C \subseteq A \times B,\ p(C) \text{ est fermé} \quad \Leftrightarrow \quad B \text{ est compact.}$$

INDICATIONS : Pour ($\Rightarrow$), prouvez la contraposée. Il est aussi utile d'argumenter séparément que si $(z_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}^P$ est une suite telle que $\|z_n\| \to +\infty$, alors $\{z_n \mid n \in \mathbb{N}\}$ est un ensemble fermé.

Exercice IV.13. Soit $f : \mathbb{R}^2 \to \mathbb{R} : (x,y) \mapsto y$ la projection sur la deuxième composante.

(i) Prouvez que f est une fonction continue.

(ii) Soit $A \subseteq \mathbb{R}^2$. Complétez la définition suivante :

$$\begin{aligned} f(A) &= \{y \in \mathbb{R} : \ \ldots\ldots\ldots\ldots\ldots\ldots \} \quad \text{(pour un } f : \mathbb{R}^2 \to \mathbb{R} \text{ général)} \\ &= \{y \in \mathbb{R} : \ \ldots\ldots\ldots\ldots\ldots\ldots \} \quad \text{(particularisé au } f \text{ ci-dessus).} \end{aligned}$$

(iii) Montrez que, si $A \subseteq \mathbb{R}^2$ est fermé et borné, alors $f(A)$ est fermé.

(iv) Trouvez un exemple de fermé $A \subseteq \mathbb{R}^2$ tel que $f(A)$ ne soit *pas* fermé. (Indication : Essayez par exemple d'avoir $f(A) =]0,1]$. Le point (vii) vous indique dans quelle direction ne *pas* chercher.)

(v) Prouvez que, quel que soit $\rho > 0$, $f(A) \cap [-\rho, \rho] = f\big(A \cap (\mathbb{R} \times [-\rho, \rho])\big)$.

(vi) Pour un ensemble arbitraire $B \subseteq \mathbb{R}^N$, montrez que

$$B \text{ est fermé} \quad \Leftrightarrow \quad \forall \rho > 0,\ B \cap \mathrm{B}[0, \rho] \text{ est compact.}$$

(vii) Prouvez que si $A \subseteq \mathbb{R}^2$ est fermé et s'il existe un $R > 0$ tel que $A \subseteq [-R, R] \times \mathbb{R}$, alors $f(A)$ est fermé. (Indication : Un dessin de la situation peut vous aider.)

Exercice IV.14. On dit qu'une fonction $f : A \subseteq \mathbb{R}^N \to \mathbb{R}^M$ est uniformément continue sur A si on peut choisir un δ dans la définition de continuité qui marche pour tous les points i.e., si

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall x \in A,\ \forall x' \in \mathrm{B}(x, \delta),\ \|f(x') - f(x)\| \leqslant \varepsilon.$$

Montrez que si $f : A \subseteq \mathbb{R}^N \to \mathbb{R}^M$ est continue et que A est compact, alors f est uniformément continue sur A.

Exercice IV.15 (août 2008). Soient $f : \mathbb{R} \to \mathbb{R}$ une fonction continue arbitraire sur $\mathbb{R}$ et deux réels $a < b$. Considérons les ensembles

$$N := \Big\{ x \in \,]a, b[\ \Big|\ f(x) = \inf_{x \in]a,b[} f(x) \Big\}$$

et

$$M := \Big\{ x \in [a, b]\ \Big|\ f(x) = \sup_{x \in [a,b]} f(x) \Big\}$$

(l'infimum et le supremum sont considérés au sens large).

(i) Peut-on affirmer que N est non-vide ? Que M est non-vide ? Justifiez par une preuve ou un contre-exemple.

(ii) Peut-on affirmer que N est compact ? Que M est compact ? Justifiez par une preuve ou un contre-exemple.

Exercice IV.16. On considère la propriété suivante à propos d'une application $u : [a, b] \to \mathbb{R}$:

$$\forall \alpha, \beta \in [a, b],\ \forall x \in \,]\alpha, \beta[,\ u(x) > \min\{u(\alpha), u(\beta)\}. \tag{IV.13}$$

(i) Montrez que si u est strictement croissante ou strictement décroissante, alors u satisfait (IV.13).

(ii) Montrez que si u est continue et satisfait (IV.13) avec $a < b$, alors il existe un unique $c \in [a,b]$ tel que u est strictement croissant sur $[a,c]$ et strictement décroissant sur $[c,b]$. (*Indication :* Regarder le maximum de u.)

Exercice IV.17. Soit $f \in \mathscr{C}([a,b])$ et (f_n) une suite de fonctions de $\mathscr{C}([a,b])$. On dit que (f_n) *converge uniformément* vers f si

$$|f_n - f|_\infty := \sup_{x\in[a,b]} |f_n(x) - f(x)| \xrightarrow[n\to\infty]{} 0.$$

(i) Montrez que si (f_n) converge uniformément vers f et si (x_n) est une suite de $[a,b]$ qui converge vers x^*, alors $f_n(x_n) \to f(x^*)$.

(ii) Montrez que $\min_{[a,b]} f_n \to \min_{[a,b]} f$.

(iii) Montrez que le point (ii) est faux si on suppose que (f_n) converge ponctuellement vers f, c'est-à-dire que $\forall x \in [a,b],\ f_n(x) \to f(x)$.

Chapitre V

Différentielle totale

V.1 Définition et interprétations

Définition V.1. Soit Ω un ouvert de $\mathbb{R}^N$, $f : \Omega \to \mathbb{R}^M$ une fonction, $a \in \Omega$ et $d \in \mathbb{R}^N$. On dit que f est *dérivable en a dans la direction d* si la limite suivante existe :

$$\lim_{t \to 0} \frac{f(a+td) - f(a)}{t} \in \mathbb{R}^M.$$

Dans ce cas, la valeur de cette limite est appelée la *dérivée directionnelle de f dans la direction d* est est notée $f'(a;d)$.

Notons que $f'(a;d)$ est simplement la dérivée de la fonction $\mathbb{R} \to \mathbb{R}^M : t \mapsto f(a+td)$ en $t = 0$. Il résulte facilement des règles de calcul des dérivées de fonctions d'une variable réelle que $f'(a;\alpha d) = \alpha f'(a;d)$ pour tout $\alpha \in \mathbb{R}$. Lorsque $\|d\| = 1$, on appelle $f'(a;d)$ la *pente de f en a dans la direction d*.

Parmi toutes les directions possibles, celles de la base canonique sont privilégiées.

Définition V.2. Soit Ω un ouvert de $\mathbb{R}^N$, $f : \Omega \to \mathbb{R}^M$ une fonction, $a \in \Omega$ et $(e_1, \dots, e_N)$ la base canonique [1] de $\mathbb{R}^N$. On appelle *dérivée partielle $k^{\text{ième}}$ de f en a* la dérivée dans la direction e_k de f en a ($1 \leqslant k \leqslant N$). On la note $\partial_k f(a)$, $\partial_{x_k} f(a)$ ou $\dfrac{\partial f}{\partial x_k}(a)$.

En particularisant ce qui a été dit pour les dérivées directionnelles, on voit que la $k^{\text{ième}}$ dérivée partielle de f en a est simplement la dérivée de la fonction $x_k \mapsto$

1. Pour rappel, $e_k = (0, \dots, 0, 1, 0, \dots, 0)$ où le 1 est en $k^{\text{ième}}$ position ($1 \leqslant k \leqslant N$).

$f(a_1,\dots,a_{k-1},x_k,a_{k+1},\dots,a_N)$ en $x_k = a_k$. Autrement dit, prendre la $k^{\text{ième}}$ dérivée partielle d'une fonction revient à considérer f uniquement comme fonction de la variable réelle x_k, les autres variables étant fixées, et à dériver cette fonction.

Définition V.3. Soit Ω un ouvert de $\mathbb{R}^N$, $f : \Omega \to \mathbb{R}^M$ une fonction et $a \in \Omega$. On dit que f est *dérivable au sens de Gateau en a* si f est dérivable en a dans toutes les directions $d \in \mathbb{R}^N$ et si la fonction $\mathbb{R}^N \to \mathbb{R}^M : d \mapsto f'(a;d)$ est linéaire. Cette fonction, $f'(a;\cdot)$, est appelée la *dérivée de Gateau* de f en a.

La demande de linéarité de $f'(a;\cdot)$ est équivalente à ce que le graphe de $x \mapsto f(a) + f'(a;x-a) : \mathbb{R}^N \to \mathbb{R}^M$ soit le translaté d'un sous-espace vectoriel de $\mathbb{R}^N \times \mathbb{R}^M$. Ce graphe est l'union des droites tangentes au graphe de f dans chaque direction. En effet, la coupe du graphe de f dans la direction d est paramétrisée par $\gamma_d : \mathbb{R} \to \mathbb{R}^N \times \mathbb{R}^M : t \mapsto \big(a+td, f(a+td)\big)$ qui passe par $\big(a, f(a)\big)$ en $t = 0$. Un vecteur tangent en $\big(a, f(a)\big)$ est donc donné par $\partial_t \gamma_d(0) = \big(d, f'(a;d)\big)$ et en conséquence la droite tangente est donnée par l'équation paramétrique $(x,y) = \gamma_d(0) + \lambda \partial \gamma_d(0)$, $\lambda \in \mathbb{R}$. L'union des droites tangentes est donc

$$\begin{aligned}
&\big\{\gamma_d(0) + \lambda \partial \gamma_d(0) : \lambda \in \mathbb{R},\ d \in \mathbb{R}^N\big\}\\
&\quad = \big\{\big(a+\lambda d, f(a) + \lambda f'(a;d)\big) : \lambda \in \mathbb{R},\ d \in \mathbb{R}^N\big\}\\
&\quad = \big\{\big(a+\lambda d, f(a) + f'(a;a+\lambda d - a)\big) : \lambda \in \mathbb{R},\ d \in \mathbb{R}^N\big\}\\
&\quad = \big\{\big(x, f(a) + f'(a;x-a)\big) : x \in \mathbb{R}^N\big\}
\end{aligned}$$

La dérivabilité au sens de Gateau est une contrainte assez faible : une fonction peut être dérivable au sens de Gateau en un point a sans pour autant être continue en a. La manière « correcte » de définir la différentiabilité de f en a est d'exprimer que f soit bien approchée par un espace tangent au voisinage de a. Pour une fonction d'une variable, c'est l'équation $f(x) = f(a) + b(x-a) + \mathrm{o}(x-a)$ qui traduit ce fait. À plusieurs dimensions le terme $b(x-a)$ avec $b \in \mathbb{R}$ doit être remplacé par $B(x-a)$ où $B : \mathbb{R}^N \to \mathbb{R}^M$ est une application linéaire. On arrive donc à la définition suivante.

Définition V.4 (Dérivabilité)**.** Soit Ω un ouvert de $\mathbb{R}^N$, $f : \Omega \to \mathbb{R}^M$ une fonction et $a \in \Omega$. On dit que f est *dérivable au sens de Fréchet en a* ou encore *différentiable en a* s'il existe une application linéaire $B : \mathbb{R}^N \to \mathbb{R}^M$ telle que

$$f(x) = f(a) + B(x-a) + \mathrm{o}(x-a) \quad \text{lorsque } x \to a. \tag{V.1}$$

Dans cette définition, la notation $\mathrm{o}(x-a)$ doit être comprise comme représentant une fonction g qui vérifie

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall x \in \mathrm{B}(a,\delta) \cap \operatorname{Dom} f,\ \|g(x)\| \leqslant \varepsilon \|x-a\| \tag{V.2}$$

ou, de manière équivalente (les détails sont laissés au lecteur),

$$g(a) = 0 \text{ (si}^{2}\ a \in \operatorname{Dom} g) \quad \text{et} \quad \frac{g(x)}{\|x-a\|} \to 0 \text{ lorsque } x \xrightarrow{\neq} a.$$

Le B de la définition précédente est unique.

Lemme V.5 (Unicité)**.** *Sous les hypothèses de la définition V.4, si B_1 et B_2 satisfont* (V.1)*, alors $B_1 = B_2$.*

Démonstration. En soustrayant membre à membre les deux égalités de type (V.1) pour B_1 et B_2 on a $0 = B_1(x-a) - B_2(x-a) + \mathrm{o}(x-a)$. Posons $B := B_1 - B_2$. On a

$$B(x-a) = \mathrm{o}(x-a) \quad \text{lorsque } x \to a. \tag{V.3}$$

On veut montrer que $B = 0$, c'est-à-dire que pour un $d \in \mathbb{R}^N$ arbitraire, on a $B(d) = 0$. Puisque $B : \mathbb{R}^N \to \mathbb{R}^M$ est une application linéaire, on sait que $B(td) = tB(d)$. En prenant $x = a + td$ avec $t \to 0$ dans (V.3), on trouve que $B(td) = tB(d) = \mathrm{o}(td)$. Par conséquent $B(d) = \mathrm{o}(td)/t \to 0$ et donc $B(d) = 0$. □

Définition V.6 (Dérivée de Fréchet)**.** Soit Ω un ouvert de $\mathbb{R}^N$, $f : \Omega \to \mathbb{R}^M$ une fonction f est dérivable au sens de Fréchet en $a \in \Omega$. L'unique application linéaire B de la définition V.4 est appelée la *dérivée de Fréchet* ou la *différentielle* de f en a et on la note $\partial f(a)$, $\partial_x f(a)$ (si on veut insister sur le fait qu'on dérive par rapport à la variable x), $Df(a)$, $\mathrm{d}f(a)$, ou f'_a.

Pour distinguer visuellement le point a auquel on prend la dérivée de la direction d en laquelle on veut calculer celle-ci, nous utiliserons des crochets autour de l'argument v. On écrit donc l'application linéraire $\partial f(a)$ comme

$$\partial f(a) : \mathbb{R}^N \to \mathbb{R}^M : d \mapsto \partial f(a)[d].$$

2. C'est le cas dans le cadre de la définition V.4. Nous l'avons écrit de la sorte pour que l'équivalence avec la propriété (V.2) soit valable pour un g général.

Lemme V.7. *Soit Ω un ouvert de $\mathbb{R}^N$, $f : \Omega \to \mathbb{R}^M$ une fonction et $a \in \Omega$. Si f est dérivable au sens de Fréchet en a, alors f est continue en a.*

Démonstration. De (V.1), il vient $f(x) - f(a) = B(x-a) + \mathrm{o}(x-a)$ avec $B = \partial f(a)$. Lorsque $x \to a$, on a $B(x-a) \to B(0) = 0$ (vu que les applications linéaires sont continues) et $\mathrm{o}(x-a) \to 0$. On a donc bien $f(x) - f(a) \to 0$. □

Lemme V.8. *Soit Ω un ouvert de $\mathbb{R}^N$, $f : \Omega \to \mathbb{R}^M$ et $a \in \Omega$. Si f est dérivable au sens de Fréchet en a, alors f est dérivable au sens de Gateau en a. De plus $f'(a;d) = \partial f(a)[d]$. En particulier, toutes les dérivées partielles de f en a existent et, pour tout $k = 1,\dots,N$, $\partial_k f(a) = \partial f(a)[e_k]$ où $(e_1,\dots,e_N)$ est la base canonique de $\mathbb{R}^N$.*

Démonstration. Soit $d \in \mathbb{R}^N$. En utilisant (V.1), on peut écrire

$$\begin{aligned}\frac{f(a+td)-f(a)}{t} &= \frac{\partial f(a)[td] + \mathrm{o}(td)}{t}\\ &= \frac{t\partial f(a)[d]}{t} + \frac{\mathrm{o}(td)}{t} = \partial f(a)[d] + \mathrm{o}(1) \xrightarrow[t\to 0]{} \partial f(a)[d]\end{aligned}$$

Ceci montre que $f'(a;d)$, la dérivée dans la direction d, existe et qu'elle vaut $\partial f(a)[d]$. □

Ce dernier résultat nous permet de calculer la dérivée de Fréchet en fonction des dérivées partielles.[3] Soit $d = (d_1,\dots,d_N) \in \mathbb{R}^N$. On écrit d dans la base canonique $(e_1,\dots,e_N)$ de $\mathbb{R}^N$ comme $d = \sum_{k=1}^N d_k e_k$. Étant donné que $\partial f(a)$ est une application linéaire, on a

$$\partial f(a)[d] = \partial f(a)\left[\sum_{k=1}^N d_k e_k\right] = \sum_{k=1}^N d_k\, \partial f(a)[e_k] = \sum_{k=1}^N d_k\, \partial_k f(a) \qquad \text{(V.4)}$$

Dans ce contexte, il est de coutume d'écrire $\mathrm{d}x_k$ pour la fonction projection sur la $k^{\text{ième}}$ composante : $(d_1,\dots,d_N) \mapsto d_k$. L'égalité précédente se réécrit avec cette notation comme $\partial f(a)[d] = \sum_{k=1}^N \partial_k f(a)\, \mathrm{d}x_k(d)$, ou encore, comme elle est valable pour tout d,

$$\partial f(a) = \sum_{k=1}^N \partial_k f(a)\, \mathrm{d}x_k$$

3. Une fonction dérivable au sens de Gateau en a ne l'est pas forcément au sens de Fréchet (vous verrez dans d'autres cours des critères sur les dérivées de Gateau qui permettent de conclure à l'existence de la dérivée de Fréchet). L'équation (V.4) ne donne donc qu'un candidat pour la dérivée de Fréchet $\partial f(a)$ mais ne prouve pas que celui-ci vérifie (V.1).

En fait ce calcul se généralise au cas où les x_k ne sont pas des scalaires mais des « sous-vecteurs » de x. Plus précisément, si $f : \mathbb{R}^{N_1} \times \cdots \times \mathbb{R}^{N_p} \to \mathbb{R}^M : (x_1,\dots,x_p) \mapsto f(x_1,\dots,x_p)$ est dérivable en $a = (a_1,\dots,a_p) \in \mathbb{R}^{N_1} \times \cdots \times \mathbb{R}^{N_p}$, on a pour tout $d = (d_1,\dots,d_p) \in \mathbb{R}^{N_1} \times \cdots \times \mathbb{R}^{N_p}$,

$$\partial f(a)[(d_1,\dots,d_p)] = \sum_{k=1}^{p} \partial_{x_k} f(a)[d_k] \tag{V.5}$$

Ici $\partial_{x_k} f(a) : \mathbb{R}^{N_k} \to \mathbb{R}^M$ est la dérivée de Fréchet de la fonction f uniquement considérée comme fonction de x_k, c'est-à-dire de

$$x_k \mapsto f(a_1,\dots,a_{k-1},x_k,a_{k+1},\dots,a_p) : \mathbb{R}^{N_k} \to \mathbb{R}^M,$$

en $x_k = a_k$.

Définition V.9. Soit Ω un ouvert de $\mathbb{R}^N$, $f : \Omega \to \mathbb{R}^M$ une fonction dérivable au sens de Fréchet en $a \in \Omega$. On appelle *matrice Jacobienne de f en a* la matrice de l'application linéaire $\partial f(a) : \mathbb{R}^N \to \mathbb{R}^M$ dans les bases canoniques de $\mathbb{R}^N$ et $\mathbb{R}^M$. On la note $\dfrac{\partial f}{\partial x}(a)$.

Si on note $(f_1(x),\dots,f_M(x))$ les composantes de $f(x)$, on a

$$\frac{\partial f}{\partial x}(a) = \begin{pmatrix} \partial_1 f_1(a) & \cdots & \partial_N f_1(a) \\ \vdots & & \vdots \\ \partial_1 f_M(a) & \cdots & \partial_N f_M(a) \end{pmatrix} \in \mathbb{R}^{M\times N}.$$

V.2 Règles de calcul

Lemme V.10. *Soit $A : \mathbb{R}^N \to \mathbb{R}^M$ une application linéaire. Quel que soit $a \in \mathbb{R}^N$, A est dérivable au sens de Fréchet en a et $\partial A(a) = A$.*

Démonstration. On vérifie aisément que l'équation (V.1) est satisfaite car $A(x) = A(a) + A(x-a)$. □

Lemme V.11. *Soit $A : \mathbb{R}^{N_1} \times \cdots \times \mathbb{R}^{N_p} \to \mathbb{R}^M$ une application multilinéaire et $a = (a_1,\dots,a_p) \in \mathbb{R}^{N_1} \times \cdots \times \mathbb{R}^{N_p}$. Alors A est dérivable au sens de Fréchet en a et, pour tout $d = (d_1,\dots,d_p) \in \mathbb{R}^{N_1} \times \cdots \times \mathbb{R}^{N_p}$, on a*

$$\partial A(a)[d] = \sum_{k=1}^{p} A(a_1,\dots,a_{k-1},d_k,a_{k+1},\dots,a_p)$$

Théorème V.12 (Dérivée des fonctions composées)**.** *Soient Ω un ouvert de $\mathbb{R}^N$, $f : \Omega \to \mathbb{R}^M$, Ψ un ouvert et $\mathbb{R}^M$ et $g : \Psi \to \mathbb{R}^P$. Soit $a \in \Omega$ tel que $f(a) \in \Psi$, f est dérivable au sens de Fréchet en a et g est dérivable au sens de Fréchet en $f(a)$. Alors $g \circ f$ est dérivable au sens de Fréchet en a et*

$$\partial(g \circ f)(a) = \partial g(f(a)) \circ \partial f(a) \tag{V.6}$$

Le symbole « $\circ$ » du membre de droite est la composition de deux applications linéaires. En termes matriciels, cela se traduit par la multiplication des matrices. Si on note x la variable de f et y celle de g, cela donne

$$\frac{\partial(g \circ f)}{\partial x}(a) = \frac{\partial g}{\partial y}(f(a)) \cdot \frac{\partial f}{\partial x}(a)$$

Si on détaille cette égalité en termes de dérivées partielles, on obtient

$$\partial_{x_k}(g \circ f)(a) = \sum_{i=1}^{M} \partial_{y_i} g(f(a))\, \partial_{x_k} f_i(a)$$

Corollaire V.13. *Soit Ω un ouvert de $\mathbb{R}^N$ et f et g deux fonctions de Ω vers $\mathbb{R}^M$ dérivables au sens de Fréchet en $a \in \Omega$. Alors la fonction $f + g : \Omega \to \mathbb{R}^M$ est dérivable au sens de Fréchet en a et $\partial(f+g)(a) = \partial f(a) + \partial g(a)$.*

Démonstration. Ceci résulte du fait que $f + g$ est la composée de $(f, g) : \Omega \to \mathbb{R}^M \times \mathbb{R}^M : x \mapsto \big(f(x), g(x)\big)$ et de l'application linéaire $\mathbb{R}^M \times \mathbb{R}^M \to \mathbb{R}^M : (y_1, y_2) \mapsto y_1 + y_2$. □

V.3 Exercices

Exercice V.1. Soit $f : \mathbb{R}^2 \to \mathbb{R} : (x_1, x_2) \mapsto x_1^2 - x_2^2$.

- Calculez la dérivée totale de f au point $(1, 1)$.
- Calculez la Jacobienne $\dfrac{\partial f}{\partial(x_1, x_2)}$ au point $(1, 1)$.
- Calculez la dérivée directionnelle de f en $(1, 1)$ dans la direction $\Big(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\Big)$.

Exercice V.2 (Examen du 5 juin 2001). Soit $f : \mathbb{R}^2 \to \mathbb{R}^2$ la fonction définie par

$$f(x, y) = \big((x-y)^2, x + 4xy\big).$$

- Calculez la matrice Jacobienne $\dfrac{\partial f}{\partial(x,y)}$ au point $(1,1)$.
- Donnez $\partial f(1,1)[h]$ où $h \in \mathbb{R}^2$ est le vecteur unitaire faisant un angle de 45° avec l'axe des x.

Exercice V.3. Soit $f : \mathbb{R}^3 \to \mathbb{R}^2$ une fonction différentiable telle que

$$f(0,0,0) = (-1,1) \qquad \text{et} \qquad \frac{\partial f}{\partial(x_1,x_2,x_3)}(0,0,0) = \begin{pmatrix} -1 & 3 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

Posons $g(x,y) = (2x-3y-1, -x^2y+2)$. Donnez $\partial(g \circ f)(0,0,0)$.

Exercice V.4. Soient $f : \mathbb{R}^4 \to \mathbb{R} : (u,v,s,t) \mapsto -2u^2v + 4s - 5u/t$ et $g : \mathbb{R} \to \mathbb{R}^4 : x \mapsto \big(x^2, -x+1, 0, (x+1)^3\big)$. Donnez $\partial(g \circ f)(0,1,0,1)$.

Exercice V.5. Soit $f : \mathbb{R}^2 \to \mathbb{R}$. Donnez l'équation cartésienne du plan tangent en (x_0,y_0). Appliquez à $f(x,y) = x^2 + y^2$ en $(1,2)$.

Exercice V.6. Soit $f : \mathbb{R}^3 \to \mathbb{R} : (r,s,t) \mapsto f(r,s,t) = F(u(r,s), v(s,t))$ où $F : \mathbb{R}^2 \to \mathbb{R}$ et $u,v : \mathbb{R}^2 \to \mathbb{R}$ sont des fonctions dérivables. Calculez $\partial_r f$ et $\partial_s f$.

Exercice V.7. Si $u(x,y) = f(x-y, y-x)$ où $f : \mathbb{R}^2 \to \mathbb{R}$ est une fonction différentiable, montrez que $\partial_x u + \partial_y u = 0$.

Exercice V.8. Si $u(x,y,z) = x^3 f(y/x, z/x)$ où $f : \mathbb{R}^3 \to \mathbb{R}$ est une fonction différentiable, montrez que $x\partial_x u + y\partial_y u + z\partial_z u = 3u$.

Exercice V.9. Soient $f : \mathbb{R}^2 \to \mathbb{R}$, $g : \mathbb{R} \to \mathbb{R}$ et $h : \mathbb{R} \to \mathbb{R}$ tels que

$$\begin{array}{lll} f(1,2) = 4 & g(0) = 1 & h(0) = 2 \\ \partial_x f(1,2) = 2 & \partial_t g(0) = 4 & \partial_t h(0) = 8 \\ \partial_y f(1,2) = \pi & & \end{array}$$

Calculez $\partial_t\big(f(g(t),h(t))\big)\big|_{t=0}$.

Exercice V.10 (juin 2001). Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction de classe $\mathscr{C}^1$. Appelons $z : \mathbb{R}^2 \to \mathbb{R}$ la fonction définie par

$$z(x,y) := yf(x^2 - y^2).$$

Montrez que

$$y\frac{\partial z}{\partial x} + x\frac{\partial z}{\partial y} = \frac{xz}{y}\ .$$

Exercice V.11 (juin 2002). Soit la fonction $w : \mathbb{R}^3 \to \mathbb{R}^N$ la fonction définie par

$$w(t,x,y) = f\big(u(t,x), v(t,y)\big)$$

pour $t \in \mathbb{R}$ où u et v sont les fonctions de $\mathbb{R}^2$ dans $\mathbb{R}$ définies par

$$u(t,x) = x + at \quad \text{et} \quad v(t,y) = y + bt.$$

Montrez que

$$\frac{\partial w}{\partial t} = a\frac{\partial w}{\partial x} + b\frac{\partial w}{\partial y}\ .$$

Exercice V.12 (août 2006). Calculez la dérivée totale et la matrice Jacobienne de la fonction

$$f : \mathbb{R}^3 \to \mathbb{R}^2 : (u,v,t) \mapsto \Big(\cos\sqrt{u^2+v^3-t}, \ln\frac{u}{v+t}\Big)$$

au point $(1,1,1)$.

Exercice V.13 (juin 2007). Calculez la dérivée totale et la matrice Jacobienne de la fonction

$$f : \mathbb{R}^2 \to \mathbb{R}^2 : (u,v) \mapsto \Big((1+u)\cos\sqrt{uv^2},\ \mathrm{e}^{\pi u^2\sqrt{v^3}}\Big)$$

au point $(\pi^2, 1)$.

Exercice V.14 (août 2007). Calculez la dérivée totale et la matrice Jacobienne de la fonction

$$f : \mathbb{R}^2 \to \mathbb{R}^2 : (u,v) \mapsto \Big((1+v)\mathrm{e}^{-v\sqrt[3]{u^2}},\ (1-u)^2\sqrt{\mathrm{tg}(uv)}\Big)$$

au point $(\pi/2, \pi/2)$.

Exercice V.15. Montrez que $\det(\mathbb{1} + A) = 1 + \mathrm{tr}A + \mathrm{o}(A)$ lorsque $A \to 0$ dans $\mathbb{R}^{N\times N}$.

Indication : Utilisez la formule de Leibniz du déterminant (c'est celle basée sur les permutations).

Chapitre VI

Introduction à l'intégration de fonctions de plusieurs variables

La théorie de l'intégration occupe une place importante en Analyse mathématique. C'est elle qui permet de calculer des longueurs, des aires, des volumes, l'énergie d'un système, le travail effectué par un objet en déplacement,... Nous ne présenterons dans ce cours qu'une introduction au sujet.

L'approche que nous avons choisie est celle de présenter les concepts et les théorèmes à un niveau intuitif — une approche plus complète vous sera offerte dans d'autres cours. En particulier, ce chapitre ne comporte pas de démonstrations ni d'ailleurs de définition précise de l'intégrale. Nous présenterons néanmoins quelques argumentations visant à vous convaincre que les formules données sont « naturelles ».

VI.1 Intégrale de fonctions d'une variable réelle

Commençons par l'intégrale des fonctions d'une variable. Celle-ci calcule l'aire « en dessous » du graphe d'une fonction. Plus précisément, si $f : [a,b] \to \mathbb{R} : x \mapsto f(x)$ est une fonction, on note

$$\int_{[a,b]} f(x)\,\mathrm{d}x \quad \text{ou simplement} \quad \int_{[a,b]} f$$

l'aire *signée* comprise entre le graphe de f est l'axe des x. Le fait que l'aire soit signée signifie que les parties au-dessus de l'axe des x y contribuent positivement tandis que celles en dessous y contribuent négativement. Ceci est illustré à la figure VI.1. Par exemple, $\int_{[0,3]} 1 - \frac{1}{2}x \, \mathrm{d}x = 3/4 = 1 - \frac{1}{4}$ car le triangle au dessus

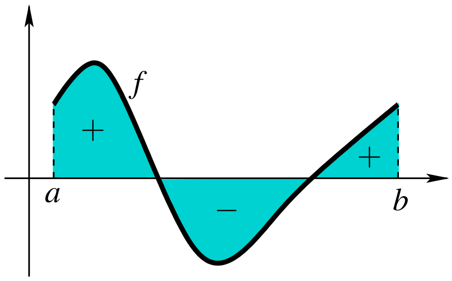


FIGURE VI.1 – Intégrale de f

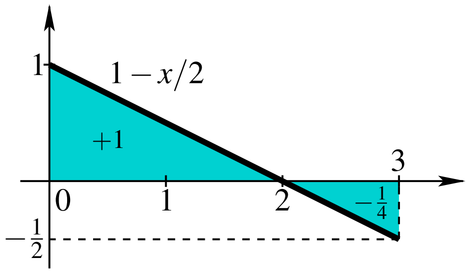


FIGURE VI.2 – $\int_{[0,3]} 1 - \frac{1}{2}x \, \mathrm{d}x$

de l'axe des x a une aire de $\frac{1}{2} \cdot 2 \cdot 1 = 1$ et celui en dessous possède une aire de $\frac{1}{2} \cdot 1 \cdot \frac{1}{2} = \frac{1}{4}$. Nous verrons ci-après des manières plus puissantes pour calculer des intégrales.

Rappelons que dans ce cours, $[a,b]$ représente juste l'ensemble des points entre a et b et donc que $[a,b] = [b,a]$, d'où $\int_{[a,b]} f = \int_{[b,a]} f$. Cependant, il est commode pour certaines formules de prendre une notation qui distingue les deux cas. Posons

$$\int_a^b f(x) \, \mathrm{d}x = \int_a^b f := \begin{cases} \int_{[a,b]} f & \text{si } a \leqslant b \\ -\int_{[a,b]} f & \text{si } a \geqslant b \end{cases}$$

Dès lors $\int_a^b f = -\int_b^a f$. Remarquons que $a = b$ se retrouve dans les deux cas sans engendrer de problème car $\int_{[a,a]} f = 0$.

On a ramené ci-avant la définition d'intégrale à la notion d'aire. Si ceci est satisfaisant pour notre intuition — nous voyons bien ce qu'est une aire — elle l'est beaucoup moins si on se pose des questions du type « l'aire a-t-elle un sens pour des ensembles compliqués ? », « quelles sont les fonctions pour lesquelles l'intégrale existe ? », « comment établir des propriétés précises de l'intégrale ? »,... Nous allons donc proposer une définition un peu plus précise de la notion d'intégrale.

Comme premier pas vers la définition de $\int_{[a,b]} f$, il faut remarquer qu'on a bien peu de chance de définir cette quantité « directement », par une « formule » comprenant f. En effet, en général l'aire entre le graphe de f et l'axe des x ne sera pas décomposable en morceaux pour lesquels on connait des formules exactes des aires (carrés, triangles, secteurs de disques,...). Comme d'habitude en analyse, on va d'abord définir l'intégrale de manière approchée puis on raffinera cette approximation. La vraie valeur s'obtiendra par passage à la limite.

L'approche que nous avons adoptée ci-dessous a l'avantage de la simplicité mais il existe des façons plus subtiles de procéder qui permettent d'intégrer plus de fonctions. Soit $a < b$. Pour approcher $\int_{[a,b]} f$, nous allons diviser l'intervalle $[a,b]$ en n sous-intervalles $[x_0,x_1]$, $[x_1,x_2],\dots,[x_{n-1},x_n]$. Dans chacun de ces intervalles $[x_i,x_{i+1}]$, $i = 0,\dots,n-1$, nous allons prendre un point ξ_i. Posons

$$I_{x_0,\dots,x_n}(f) := \sum_{0\leqslant i<n} f(\xi_i)(x_{i+1}-x_i) \tag{VI.1}$$

Cette quantité représente l'aire signée des rectangles ayant comme base les intervalles $[x_i,x_{i+1}]$ et comme hauteurs respectives $f(\xi_i)$ (voir figure VI.3) et approxime donc l'aire signée entre le graphe de f et l'axe des x. Cette approximation est d'autant meilleure que la base des rectangles est petite car on espère alors que le profil en escalier se rapproche de celui de la fonction f (voir figure VI.4). L'in-

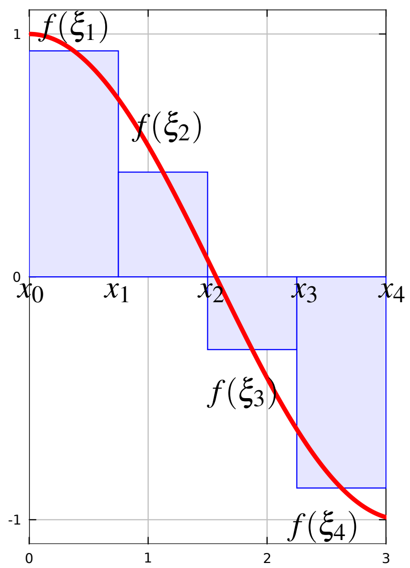


FIGURE VI.3 – $I_{x_0,\dots,x_n}(f)$

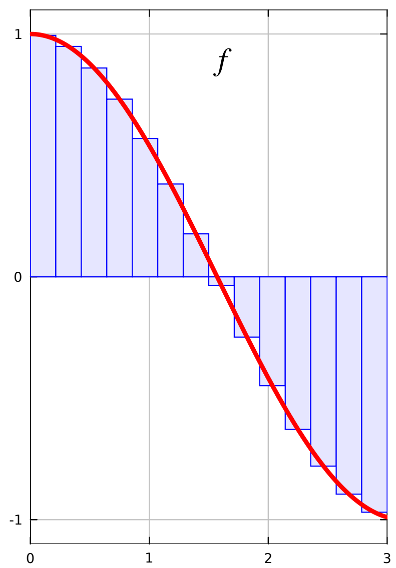


FIGURE VI.4 – $I_{x_0,\dots,x_n}(f)$

tégrale de f sera alors la limite des valeurs $I_{x_0,\dots,x_n}(f)$ lorsque toutes les longueurs des sous-intervalles tendent vers 0, c'est-à-dire lorsque $\max\{|x_{i+1}-x_i| : 0 \leqslant i < n\} \to 0$. Cela conduit à la définition suivante.

Définition VI.1 (Intégrale de Riemann)**.** Une fonction $f : [a,b] \to \mathbb{R}$ est dite *intégrable* s'il existe un $c \in \mathbb{R}$, tel que, quelle que soit la suite de divisions $\big(a = x_0^{(k)} < \cdots < x_{n(k)}^{(k)} : k \in \mathbb{N}\big)$, telle que

$$\max\big\{|x_{i+1}^{(k)} - x_i^{(k)}| : 0 \leqslant i < n(k)\big\} \xrightarrow[k\to\infty]{} 0,$$

la suite $\big(I_{x_0^{(k)},\ldots,x_{n(k)}^{(k)}}(f)\big)_{k\in\mathbb{N}}$ converge vers c. Dans ce cas, la valeur c est appelée l'intégrale de f sur $[a,b]$ et est notée $\int_{[a,b]} f$.

Voici quelques propriétés de l'intégrale avec quelques intuitions ainsi que des commentaires sur le fait que certaine de ces propriétés requièrent une intégrale plus forte que celle définie ci-dessus.

(i) L'ensemble des fonctions intégrables est un espace vectoriel et l'intégrale est une application linéaire de cet espace vers $\mathbb{R}$. Autrement dit, si $f, g : [a,b] \to \mathbb{R}$ sont des fonctions intégrables et $\alpha, \beta \in \mathbb{R}$, alors $\alpha f + \beta g$ est intégrable et

$$\int_{[a,b]} \alpha f(x) + \beta g(x)\,\mathrm{d}x = \alpha \int_{[a,b]} f(x)\,\mathrm{d}x + \beta \int_{[a,b]} g(x)\,\mathrm{d}x$$

Cette propriété est assez naturelle puisque

$$I_{x_0,\ldots,x_n}(\alpha f + \beta g) = \alpha I_{x_0,\ldots,x_n}(f) + \beta I_{x_0,\ldots,x_n}(g)$$

et que l'intégrale s'obtient par un passage à la limite sur $I_{x_0,\ldots,x_n}$.

(ii) Si $f, g : [a,b] \to \mathbb{R}$ sont deux fonctions telles que

$$\forall x \in [a,b], \quad |f(x)| \leqslant g(x)$$

et que g est intégrable sur $[a,b]$, alors f est également intégrable sur $[a,b]$. Un tel résultat peut se comprendre par le fait qur l'inégalité implique que l'aire entre f et l'axe des x est bornée par celle entre g et l'axe des x. Si cette dernière est finie, alors la première doit l'être aussi. C'est analogue à la convergence dominée pour les séries. Malheureusement ce résultat n'est pas vrai pour l'intégrale de Riemann (voir le point suivant pour un exemple). Il faut passer à une intégrale plus puissante (c'est-à-dire qui peut intégrer plus de fonctions), appelée intégrale de Lebesgue, et supposer que

f est *mesurable*. La notion de fonction mesurable sera définie dans un autre cours mais il suffit de savoir ici que toutes les fonctions « usuelles » le sont. De plus une fonction intégrable est nécessairement mesurable.

(iii) Soit $f : [a,b] \to \mathbb{R}$ une fonction. Alors

$$f \text{ est intégrable} \Leftrightarrow |f| \text{ est intégrable.} \tag{VI.2}$$

L'implication « $\Leftarrow$ » n'est pas valable pour l'intégrale de Riemann. Elle l'est pour l'intégrale de Lebesgue car on peut la voir comme un cas particulier de la propriété ci-dessus avec $g = |f|$.

Pour comprendre pourquoi cela ne marche pas pour l'intégrale de Riemann, prenons la fonction $f : [0,1] \to \mathbb{R}$ définie par

$$f(x) = \begin{cases} +1 & \text{si } x \in \mathbb{Q} \\ -1 & \text{si } x \notin \mathbb{Q} \end{cases}$$

Comme $|f|$ est la fonction constante 1, elle est intégrable sur $[0,1]$. Par contre, on peut prendre des points de division $0 = x_0 < \cdots < x_n = 1$ et $\xi_i = (x_{i+1} + x_i)/2$ avec $\max|x_{i+1} - x_i|$ aussi petit que l'on veut tels que l'une ou l'autre des deux situations suivantes (au choix) soit vraie :

- tous les x_i soient rationnels et donc aussi les ξ_i, ce qui implique que $f(\xi_i) = 1$ pour tout i et donc $I_{x_0,\dots,x_n}(f) = 1$;
- les x_i sont choisis tels que $x_1 - x_0 = x_1 \notin \mathbb{Q}$ et $x_{i+1} - x_i \in \mathbb{Q}$ pour $i = 1, \dots, n-2$, ce qui implique que tous les ξ_i sont irrationnels d'où $f(\xi_i) = -1$ et donc $I_{x_0,\dots,x_n}(f) = -1$.

En conclusion, certaines suites de $I_{x_0,\dots,x_n}(f)$ vont conveger vers 1 et d'autres vers -1. Ceci montre qu'aucune valeur $c \in \mathbb{R}$ de l'intégrale ne peut satisfaire la définition VI.1.

(iv) Si $f : [a,b] \to \mathbb{R}$ est continue, alors elle est intégrable.

Dans le cadre de l'intégrale de Lebesgue, cela peut être vu comme une conséquence du fait que les fonctions constantes sont intégrables et de l'inégalité

$$\forall x \in [a,b], \quad |f(x)| \leqslant c \qquad \text{avec } c := \sup_{x \in [a,b]} |f(x)| \in \mathbb{R}$$

où l'appartenance de c à $\mathbb{R}$ résulte du fait que les fonctions continues, en l'occurence $x \mapsto |f(x)|$, atteignent leurs bornes sur des compacts (théorème IV.6).

Une preuve directe de cette même propriété peut être faite pour l'intégrale de Riemann.

(v) Si $f : [a,b] \to \mathbb{R}$ est une fonction intégrable et $[c,d] \subseteq [a,b]$, alors f est intégrable sur $[c,d]$.

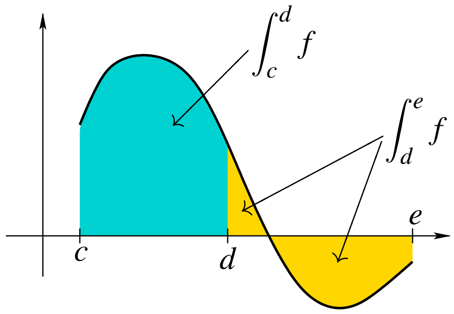
Ceci est assez naturel puisque l'aire « au-dessus » de $[c,d]$ est forcément plus petite que l'aire au-dessus de $[a,b]$.

(vi) Si $f : [a,b] \to \mathbb{R}$ est une fonction intégrable et $c,d,e \in [a,b]$, alors

$$\int_c^d f + \int_d^e f = \int_c^e f. \tag{VI.3}$$

Il suffit de montrer cette inégalité dans le cas $c < d < e$. Les autres cas s'en déduisent. Par exemple, si $c < e < d$, on peut écrire $\int_c^e f + \int_e^d f = \int_c^d f$ et donc (VI.3) en faisant passer $\int_e^d f = -\int_d^e f$ dans le membre de droite. Il en va de même pour les autres possibilités (quelles sont-elles et comment les résout-on ?).

Pour $c < d < e$, l'égalité (VI.3) est représentée à la figure VI.5.



$$\int_c^d f + \int_d^e f = \int_c^e f$$

FIGURE VI.5 – Additivité de l'intégrale

(vii) Si $f,g : [a,b] \to \mathbb{R}$ sont deux fonctions intégrables et $f(x) \leqslant g(x)$ pour $x \in [a,b]$, alors

$$\int_{[a,b]} f \leqslant \int_{[a,b]} g$$

Graphiquement, on peut voir que l'aire signée sous le graphe de f est « plus petite » ou « plus négative » que celle de g (voir figure VI.6). Ceci

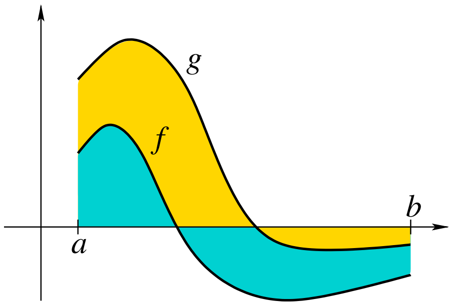


FIGURE VI.6 – Croissance de l'intégrale

découle aussi d'un passage à la limite sur l'inégalité (facile à établir) : $I_{x_0,\dots,x_n}(f) \leqslant I_{x_0,\dots,x_n}(g)$.

De l'intuition en termes d'aires, il vient immédiatement que $\int_{[a,b]} 1 \, \mathrm{d}x = |b-a|$ ce qui est la longueur de $[a,b]$ (voir figure VI.7). De manière générale, étant donné

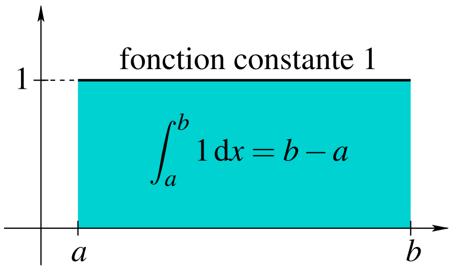


FIGURE VI.7 – Longueur d'un intervalle

un ensemble borné $A \subseteq \mathbb{R}$, on définit la mesure de A par

$$\operatorname{mes}(A) := \int_{[a,b]} \chi_A$$

où $a, b \in \mathbb{R}$ sont tels que $A \subseteq [a,b]$ (ils existent puisque A est borné) et où χ_A est la *fonction caractéristique* de l'ensemble A, c'est-à-dire

$$\chi_A : \mathbb{R} \to \mathbb{R} : x \mapsto \chi_A(x) := \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{sinon.} \end{cases}$$

(La quantité $\operatorname{mes}(A)$ n'est pas définie pour tout ensemble borné A car χ_A n'est pas forcément intégrable. Cette question sera abordée dans vos cours ultérieurs.) En utilisant la propriété (vi) ci-dessus, il est aisé de prouver que $\operatorname{mes}(A)$ ne dépend pas du choix de a et b.

Plus généralement, si A est un ensemble borné et $f : A \to \mathbb{R}$, on dira que f est *intégrable* sur A si la fonction

$$\tilde{f} : [a,b] \to \mathbb{R} : x \mapsto \tilde{f}(x) := \begin{cases} f(x) & \text{si } x \in A \\ 0 & \text{sinon} \end{cases}$$

est intégrable sur $[a,b]$ où a et b sont tels que $A \subseteq [a,b]$. L'intégrale de f sur A, notée $\int_A f$, est alors définie comme

$$\int_A f := \int_{[a,b]} \tilde{f}.$$

Comme précédemment, il est facile de montrer que ces deux dernières définitions ne dépendent pas des valeurs particulières de a et b (tant que $[a,b]$ contient A).

VI.2 Intégrale de fonctions de plusieurs variables réelles

Les idées de la section précédente peuvent être étendues aux fonctions de plusieurs variables. On appelle un *pavé* un produit d'intervalles fermés de même dimension que l'espace. Donc, si P est un pavé de $\mathbb{R}^N$, il existe certains réels $a_1, \dots, a_N, b_1, \dots, b_N$ tels que $P = \prod_{i=1}^N [a_i, b_i] = [a_1, b_1] \times \cdots \times [a_N, b_N]$. Si P est un pavé et $f : P \to \mathbb{R} : x \mapsto f(x)$ une fonction, on note

$$\int_P f(x)\,\mathrm{d}x \quad \text{ou simplement} \quad \int_P f$$

le volume signé compris entre le graphe de f et l'espace « horizontal » des x (voir figure VI.8). Le signe du volume est positif ou négatif selon que ce volume se situe au-dessus ou en-dessous de l'espace des x. On peut donc écrire que

$$\begin{aligned}\int_P f(x)\,\mathrm{d}x = \text{volume}\big\{(x,y) \in \mathbb{R}^N \times \mathbb{R} : 0 \leqslant y \leqslant f(x)\big\} \\ - \text{volume}\big\{(x,y) \in \mathbb{R}^N \times \mathbb{R} : f(x) \leqslant y \leqslant 0\big\}\end{aligned}$$

Une définition précise de l'intégrale suit les mêmes lignes qu'à une dimension. On définit un *découpage* d'un pavé P comme un ensemble de pavés $\{P_1, \dots, P_n\}$ tel que $\bigcup_{i=1}^n P_i = P$ et $\operatorname{int} P_i \cap \operatorname{int} P_j = \varnothing$ pour tout $i \neq j$ (voir figure VI.9). Un *découpage pointé* d'un pavé P est un ensemble de couples $\{(\xi_1, P_1), \dots, (\xi_n, P_n)\}$

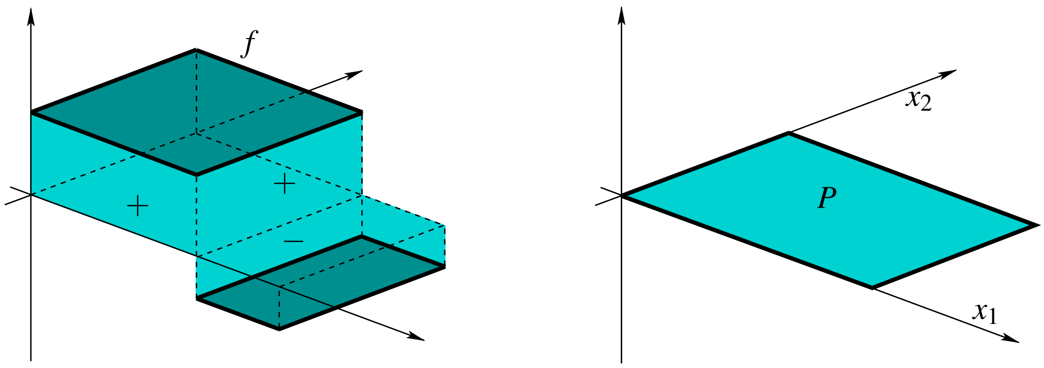


FIGURE VI.8 – Intégrale d'une fonction à deux variables

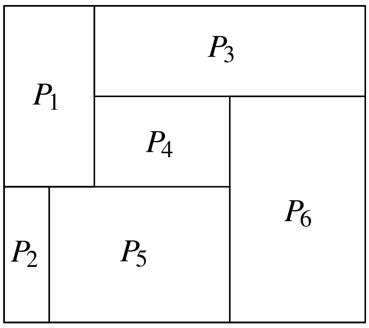


FIGURE VI.9 – Découpage

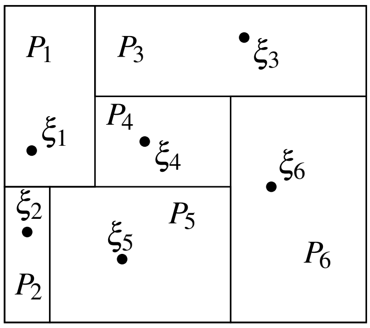


FIGURE VI.10 – Découpage pointé

tel que $\{P_1,\dots,P_n\}$ est un découpage de P et $\xi_i \in P_i$ pour tout i (voir figure VI.10). On appelle le *diamètre* d'un ensemble A au sens de la norme $\|\cdot\|$ la quantité

$$\operatorname{diam}_{\|\cdot\|} A := \sup\big\{ \|x_1 - x_2\| : x_1, x_2 \in A \big\}.$$

Le *volume* d'un pavé $P = \prod_{i=1}^{N} [a_i, b_i]$ est la quantité notée $\operatorname{vol}(P)$ définie par

$$\operatorname{vol}(P) = \prod_{i=1}^{N} |b_i - a_i|.$$

Définition VI.2 (Intégrale de Riemann)**.** Soit P un pavé de $\mathbb{R}^N$ et $f : P \to \mathbb{R}$ une fonction. On dit que f est *intégrable* sur P s'il existe un $c \in \mathbb{R}$ tel que, pour toute suite de découpages pointés $\mathbf{P}^{(k)} = \big\{(\xi_1^{(k)}, P_1^{(k)}), \dots, (\xi_{n(k)}^{(k)}, P_{n(k)}^{(k)})\big\}$, $k \in \mathbb{N}$, telle que

$$\max\big\{\operatorname{diam}_{\|\cdot\|} P_i^{(k)} : 1 \leqslant i \leqslant n(k)\big\} \xrightarrow[k\to\infty]{} 0,$$

on a

$$I_{\mathbf{P}^{(k)}}(f) := \sum_{i=1}^{n(k)} f(\xi_i^{(k)}) \operatorname{vol}(P_i^{(k)}) \xrightarrow[k\to\infty]{} c.$$

On étend la définition aux ensembles bornés $A \subseteq \mathbb{R}^N$ comme suit : $f : A \to \mathbb{R}$ est dite *intégrable* si la fonction

$$\tilde{f} : P \to \mathbb{R} : x \mapsto \begin{cases} f(x) & \text{si } x \in A \\ 0 & \text{sinon} \end{cases}$$

est intégrable où P est un pavé tel que $A \subseteq P$. Dans ca cas, l'*intégrale* de f sur A, notée $\int_A f$, est définie comme

$$\int_A f := \int_P \tilde{f}.$$

Ces définitions sont indépendantes du P considéré.

Les propriétés de l'intégrale à plusieurs variables sont très similaires à celles énoncées pour une variable. Les commentaires étant essentiellement les mêmes qu'alors, nous allons simplement les énoncer.

(i) L'ensemble des fonctions intégrables sur un ensemble $A \subseteq \mathbb{R}^N$ est un espace vectoriel et l'intégrale est une application linéaire de cet espace vers l'ensemble des réels $\mathbb{R}$.

(ii) Si $A \subseteq \mathbb{R}^N$ et $f : A \to \mathbb{R}$ est une fonction (mesurable) telle qu'il existe une fonction intégrable $g : A \to \mathbb{R}$ satisfaisant

$$\forall x \in A, \quad |f(x)| \leqslant g(x)$$

alors f est (Lebesgue) intégrable.

(iii) Si $f : A \to \mathbb{R}$ est une fonction définie sur un ensemble $A \subseteq \mathbb{R}^N$, alors

$$f \text{ est (Lebesgue) intégrable} \Leftrightarrow |f| \text{ est (Lebesgue) intégrable}$$

(iv) Si $f : A \to \mathbb{R}$ est une fonction continue définie sur un ensemble (mesurable) A, alors elle est intégrable.

(v) Si $f : A \to \mathbb{R}$ est une fonction intégrable et $B \subseteq A$ (est un ensemble mesurable) alors $f : B \to \mathbb{R}$ est intégrable.

(vi) Si $f : A \to \mathbb{R}$ est une fonction intégrable et $A_1, \dots, A_k$ sont des ensembles (mesurables) tels que $A = \bigcup_{i=1}^k A_i$ et $A_i \cap A_j = \varnothing$ pour $i \neq j$, alors

$$\int_A f = \sum_{i=1}^k \int_{A_i} f.$$

(vii) Si $f, g : A \to \mathbb{R}$ sont deux fonctions intégrables et $f \leqslant g$, alors

$$\int_A f \leqslant \int_A g.$$

VI.3 Calcul d'intégrales à une variable

L'outil principal de calcul de l'intégrale à une dimension est le théorème suivant :

Théorème VI.3 (Théorème fondamental de l'Analyse)**.** *Soit $f : [a,b] \to \mathbb{R}$ une fonction continue. Alors la fonction $[a,b] \to \mathbb{R} : x \mapsto \int_a^x f$ est continue sur $[a,b]$, dérivable sur $]a,b[$ et, pour tout $x_0 \in\,]a,b[$,*

$$\partial_x \Big(\int_a^x f \Big)(x_0) = f(x_0).$$

Ce théorème montre que l'intégrale et la dérivée ne sont pas deux concepts séparés — comme on aurait pu le croire à priori — mais qu'en quelque sorte ils sont « inverses » l'un de l'autre.

Démonstration. Soit $x_0 \in]a,b[$. Nous allons montrer que la fonction $[a,b] \to \mathbb{R} : x \mapsto \int_a^x f$ est dérivable en x_0 ce qui impliquera qu'elle est continue en x_0. La continuité en a et en b est laissée au lecteur.

On doit prouver que

$$\frac{1}{h}\Big(\int_a^{x_0+h} f - \int_a^{x_0} f\Big) = \frac{1}{h}\int_{x_0}^{x_0+h} f \xrightarrow[h\to 0]{} f(x_0). \tag{VI.4}$$

Soit $\varepsilon > 0$. Puisque f est continue, il existe un $\delta > 0$ tel que

$$\forall x \in [x_0-\delta, x_0+\delta], \quad f(x_0)-\varepsilon \leqslant f(x) \leqslant f(x_0)+\varepsilon. \tag{VI.5}$$

Montrons que ce δ convient dans la définition de convergence pour (VI.4). Soit $|h| \leqslant \delta$. Puisque $[x_0, x_0+h] \subseteq [x_0-\delta, x_0+\delta]$, on obtient en intégrant les inégalités (VI.5) que

$$\begin{aligned} h\cdot\big(f(x_0)-\varepsilon\big) &= \int_{x_0}^{x_0+h} f(x_0)-\varepsilon \,\mathrm{d}x \\ &\leqslant \int_{x_0}^{x_0+h} f(x)\,\mathrm{d}x \leqslant \int_{x_0}^{x_0+h} f(x_0)+\varepsilon\,\mathrm{d}x = h\cdot\big(f(x_0)+\varepsilon\big). \end{aligned}$$

Dès lors

$$\Big|\frac{1}{h}\int_{x_0}^{x_0+h} f(x)\,\mathrm{d}x - f(x_0)\Big| \leqslant \varepsilon$$

ce qui termine la preuve. □

Corollaire VI.4. *Soit $f : [a,b] \to \mathbb{R}$ une fonction continue et $F : [a,b] \to \mathbb{R}$ une fonction continue sur $[a,b]$ et dérivable sur $]a,b[$ telle que $\forall x \in]a,b[$, $\partial F(x) = f(x)$. Alors*

$$\int_a^b f(x)\,\mathrm{d}x = F(b) - F(a).$$

Une telle fonction F est appelée une *primitive* de f sur $[a,b]$.

Démonstration. Posons $G(x) := \int_a^x f$. Le théorème VI.3 affirme que $G : [a,b] \to \mathbb{R}$ est continue sur $[a,b]$, dérivable sur $]a,b[$ et $\partial G = f$. Par conséquent, $F - G$ est continue sur $[a,b]$, dérivable sur $]a,b[$ et $\partial(F-G) = \partial F - \partial G = f - f = 0$. Ceci implique qu'il existe une constante $c \in \mathbb{R}$ telle que $f(x) - G(x) = c$ pout tout $x \in [a,b]$. Dès lors, $F(b) - F(a) = G(b) - G(a)$. Il suffit alors de remarquer que $G(b) = \int_a^b f$ et $F(a) = \int_a^a f = 0$ pour conclure. □

Ce corollaire nous dit que si on trouve une primitive F d'une fonction f, il est facile de calculer $\int_a^b f$.

VI.4 Fubini

Théorème VI.5 (Fubini)**.** *Soient $A_1 \subseteq \mathbb{R}^{N_1}$ et $A_2 \subseteq \mathbb{R}^{N_2}$. Si $f : A_1 \times A_2 \to \mathbb{R} : (x,y) \mapsto f(x,y)$ est une fonction intégrable sur $A_1 \times A_2$, on peut calculer son intégrale en itérant des intégrales plus simples :*

$$\begin{aligned}\int_{A_1\times A_2} f(x,y)\,\mathrm{d}(x,y) &= \int_{A_2}\Big(\int_{A_1} f(x,y)\,\mathrm{d}x\Big)\mathrm{d}y\\ &= \int_{A_1}\Big(\int_{A_2} f(x,y)\,\mathrm{d}y\Big)\mathrm{d}x\end{aligned}$$

Notez que $\int_{A_2} f(x,y)\,\mathrm{d}x$ ne dépend plus de x mais seulement de y, c'est donc une fonction de y, $y \mapsto \int_{A_2} f(x,y)\,\mathrm{d}x$ qu'on peut dès lors intégrer sur A_2. Le théorème ci-dessus dit que ces intégrales successives existent (et qu'on peut les faire dans l'ordre qu'on veut) pour autant que f soit intégrable sur $A_1 \times A_2$. L'inverse n'est pas vrai. C'est-à-dire qu'il ne suffit pas de pouvoir calculer par exemple $\int_{A_2}(\int_{A_1} f(x,y)\,\mathrm{d}y)\mathrm{d}x$ pour que toutes les autres intégrales existent. Choisissons $f :]-1,1[\times]0,1[\to\mathbb{R} : (x,y)\mapsto x/y$. Quel que soit $y\in]0,1[$, $\int_{]-1,1[} x/y\,\mathrm{d}x = 0$. Donc, $\int_0^1(\int_{-1}^1 x/y\,\mathrm{d}x)\mathrm{d}y = 0$. Cependant, si on cherche à faire ces intégrales successives en commencant par y, on trouve que $\int_0^1 x/y\,\mathrm{d}y$ n'existe pas ! Le théorème suivant répond à cette préoccupation.

Théorème VI.6 (Tonelli)**.** *Soient $A_1 \subseteq \mathbb{R}^{N_1}$ et $A_2 \subseteq \mathbb{R}^{N_2}$. Si $f : A_1 \times A_2 \to \mathbb{R}$ est une fonction telle que*

- *pour tout $y \in A_2$, $\int_{A_1} |f(x,y)|\,\mathrm{d}x$ existe ;*
- *la fonction $A_2 \to \mathbb{R} : y \mapsto \int_{A_1} |f(x,y)|\,\mathrm{d}x$ est intégrable sur A_2 ;*

alors f est intégrable sur $A_1 \times A_2$ (et on peut donc appliquer Fubini).

Bien sûr, l'énoncé où les intégrales sont faites d'abord par rapport à y et ensuite par rapport à x est aussi vrai.

Bien que ce ne soit pas apparent au vu de l'énoncé, le théorème de Fubini a une application plus vaste que celui où f est définie sur un « rectangle » $A_1 \times A_2$. La remarque fondamentale est que, si $f : A \to \mathbb{R}$ est intégrable et que $A \subseteq A_1 \times A_2$, alors

$$\int_A f(x,y)\,\mathrm{d}(x,y) = \int_{A_1\times A_2} \tilde{f}(x,y)\,\mathrm{d}(x,y)$$

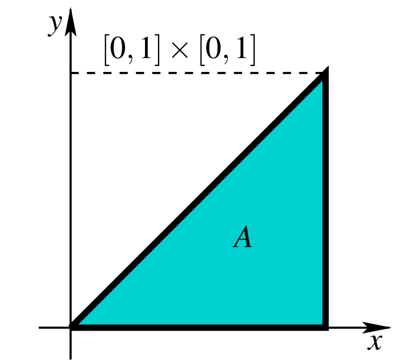
où $\tilde{f} = f$ sur A et $\tilde{f} = 0$ sur $(A_1 \times A_2)\setminus A$. Prenons un exemple pour mieux illustrer cette idée. Supposons qu'on veuille intégrer une fonction $f : A \to \mathbb{R}$ où $A \subseteq \mathbb{R}^2$ est le triangle (plein) de sommets $(0,0)$, $(1,0)$ et $(1,1)$. Clairement $A \subseteq [0,1] \times [0,1]$. On peut écrire :

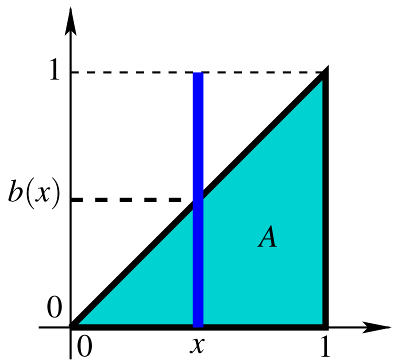
$$\int_A f = \int_{[0,1]\times[0,1]} \tilde{f} \qquad \text{où } \tilde{f} = \begin{cases} f(x,y) & \text{si } (x,y) \in A, \\ 0 & \text{sinon.} \end{cases}$$

Nous pouvons appliquer le théorème de Fubini à la fonction $\tilde{f}$ ce qui donne

$$\int_A f = \int_0^1 \Big(\int_0^1 \tilde{f}(x,y)\, \mathrm{d}y \Big) \mathrm{d}x.$$

(On aurait aussi pu choisir d'intégrer dans l'ordre inverse. Faites le !)





Évidemment, nous voudrions exprimer le membre de droite en fonction de f uniquement, $\tilde{f}$ n'étant à nos yeux qu'une fonction auxiliaire qui permet d'utiliser le théorème de Fubini. Pour cela, examinons de plus près l'intégrale intérieure $\int_0^1 \tilde{f}(x,y)\, \mathrm{d}y$. Calculer cette intégrale signifie fixer $x \in [0,1]$ et faire varier y de 0 à 1. On le voit sur le dessin, lorsque y varie, le point (x,y) appartient à A si $0 \leqslant y \leqslant b(x)$ auquel cas $\tilde{f} = f$, et $(x,y) \notin A$ si $b(x) < y \leqslant 1$ auquel cas $\tilde{f} = 0$. Autrement dit, on peut réécrire la définition de $\tilde{f}$ comme

$$\tilde{f} = \begin{cases} f(x,y) & \text{si } 0 \leqslant y \leqslant b(x), \\ 0 & \text{si } b(x) < y \leqslant 1. \end{cases}$$

Dès lors,

$$\int_0^1 \tilde{f}(x,y)\, \mathrm{d}y = \int_0^{b(x)} \tilde{f}(x,y)\, \mathrm{d}y + \int_{b(x)}^1 \tilde{f}(x,y)\, \mathrm{d}y = \int_0^{b(x)} f(x,y)\, \mathrm{d}y$$

Reste à déterminer $b(x)$. On voit sur le dessin que le point $(x, b(x))$ est à l'intersection de la droite verticale d'abcisse x et de la diagonale principale. Dès lors, $b(x) = x$. En rassemblant les résultats précédents, on trouve

$$\int_A f(x,y)\,\mathrm{d}(x,y) = \int_0^1 \Big(\int_0^x f(x,y)\,\mathrm{d}y\Big)\mathrm{d}x.$$

Ceci marche pour n'importe quelle fonction f intégrable sur A car c'est la géométrie du domaine qui détermine les bornes d'intégration, non la forme de la fonction.

VI.5 Changement de variables

Lorsque le domaine possède une géométrie « compliquée » ou simplement de nombreuses symétries, mieux vaut d'abord essayer de « déformer » ce domaine en quelque chose de plus simple avant d'y appliquer le théorème de Fubini. Cette « déformation » prend ici la forme d'un changement de variable. C'est l'objet du théorème qui suit.

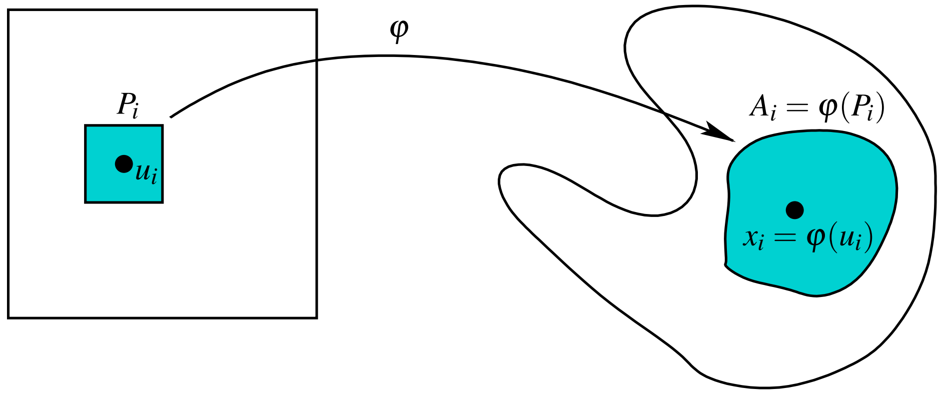
Théorème VI.7 (Intégration par changement de variables)**.** *Soient A et B deux ouverts de $\mathbb{R}^N$, $\varphi : B \to A$ un $\mathscr{C}^1$-difféomorphisme et $f : A \to \mathbb{R}$. On a que f est intégrable sur A ssi $f \circ \varphi\,|\det \partial\varphi|$ est intégrable sur B auquel cas*

$$\int_A f(x)\,\mathrm{d}x = \int_B f\big(\varphi(u)\big)\,|\det \partial\varphi(u)|\,\mathrm{d}u.$$

Rappelons que dire que φ est un $\mathscr{C}^1$-difféomorphisme signifie que φ est de classe $\mathscr{C}^1$ sur B, $\varphi : B \to A$ est bijective, et $\varphi^{-1} : A \to B$ est aussi de classe $\mathscr{C}^1$. Pour montrer qu'une fonction $\varphi : B \to A$ est un $\mathscr{C}^1$-difféomorphisme, il suffit de voir que $\varphi \in \mathscr{C}^1$, φ est bijective et $\det\big(\partial\varphi(u)\big) \neq 0$ pour tout $u \in B$.

On peut expliquer intuitivement cette formule par le diagramme suivant.

$$B \xrightarrow{\quad\varphi\quad} A \xrightarrow{\quad f\quad} \mathbb{R}$$



On veut intégrer f sur tous les $x \in A$. Ces x sont paramétrés par u. Lorsque x parcourt A, u parcourt B. Donc, puisque l'intégrale en x porte sur A, celle en u se fait sur B. Un découpage pointé $\{(u_i, P_i) : 1 \leqslant i \leqslant n\}$ de B engendre naturellement un recouvrement de A par des $A_i = \varphi(P_i)$ (tels que $\forall i \neq j$, $\operatorname{int} A_i \cap \operatorname{int} A_j = \varnothing$) avec des points $x_i = \varphi(u_i) \in A_i$. Puisqu'on va s'intéresser aux aires des A_i et P_i, il faut comprendre la relation qu'il y a entre les deux. Mais puisque chaque P_i est très petit, il est « concentré » autour du point u_i. Or une bonne approximation de φ près de u_i est le développement de Taylor d'ordre 1, c'est-à-dire :

$$A_i = \varphi(P_i) \approx \varphi(u_i) + \langle \partial\varphi(u_i), P_i - u_i \rangle.$$

Puisqu'additionner un vecteur constant à un ensemble consiste juste à le translater de ce vecteur, l'aire de l'ensemble ne change pas. On peut donc affirmer que

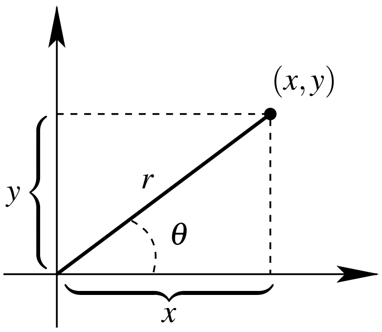
$$\operatorname{vol}(A_i) \approx \operatorname{vol}\langle \partial\varphi(u_i), P_i - u_i \rangle \quad \text{et} \quad \operatorname{vol}(P_i - u_i) = \operatorname{vol}(P_i).$$

Ainsi on est réduit à comprendre comment l'aire d'un ensemble se transforme par passage à travers une application linéaire. La valeur absolue du déterminant de l'application linéaire donne le facteur de multiplication. Donc,

$$\operatorname{vol}\langle \partial\varphi(u_i), P_i - u_i \rangle = |\det \partial\varphi(u_i)| \operatorname{vol}(P_i - u_i).$$

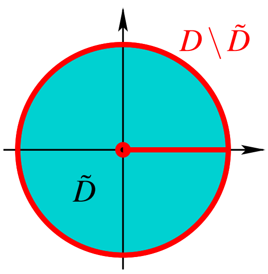
En mettant ensemble les idées précédentes, on trouve

$$\begin{aligned} \int_A f(x)\,\mathrm{d}x &= \lim \sum f(x_i) \operatorname{vol}(A_i) \\ &= \lim \sum f\big(\varphi(u_i)\big) \, |\det \partial\varphi(u_i)| \operatorname{vol}(P_i) = \int_B f\big(\varphi(u)\big) \, |\det \partial\varphi(u)| \,\mathrm{d}u \end{aligned}$$



Voyons maintenant l'utilisation de ce théorème sur un exemple très simple. Supposons que nous voulions calculer l'aire du disque D de $\mathbb{R}^2$ de rayon 1 et de centre $(0,0)$. En d'autres mots, nous voudrions calculer $\int_D 1$. Il y a de nombreuses manières d'arriver à la solution mais ici supposons que nous voulions exploiter la symétrie de rotation de D. Les coordonnées naturelles associées à une telle symétrie sont les coordonnées polaires et nous aimerions reformuler notre intégrale dans ces coordonnées. Parler de coordonnées polaires signifie choisir de repérer un point $(x,y) \in \mathbb{R}^2$ par sa distance r à l'origine et son angle θ avec l'axe des x. Formellement, cela donne lieu au difféomorphisme

$$\varphi : [0,1] \times [0,2\pi[\to D : (r,\theta) \mapsto (r\cos\theta, r\sin\theta).$$



À strictement parler, φ n'est *pas* un difféomorphisme puisque $\varphi\big(\{0\} \times [0,2\pi[\big) = \{0\}$. De plus il faudrait que φ soit défini sur un ouvert. Cela nous amène à restreindre φ à $]0,1[\times]0,2\pi[$. On se convaincra facilement que

$$\varphi :]0,1[\times]0,2\pi[\to \tilde{D},$$

où $\tilde{D} := \{(x,y) : x^2 + y^2 < 1 \text{ et } y \neq 0 \text{ si } x \geqslant 0\}$, est un difféomorphisme (voir le calcul de $\det \partial\varphi$ ci-dessous). Heureusement, comme $D \setminus \tilde{D}$ est d'aire nulle, $\int_D = \int_{\tilde{D}}$. La matrice Jacobienne de φ est aisée à calculer :

$$\frac{\partial\varphi}{\partial(r,\theta)}(r,\theta) = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}.$$

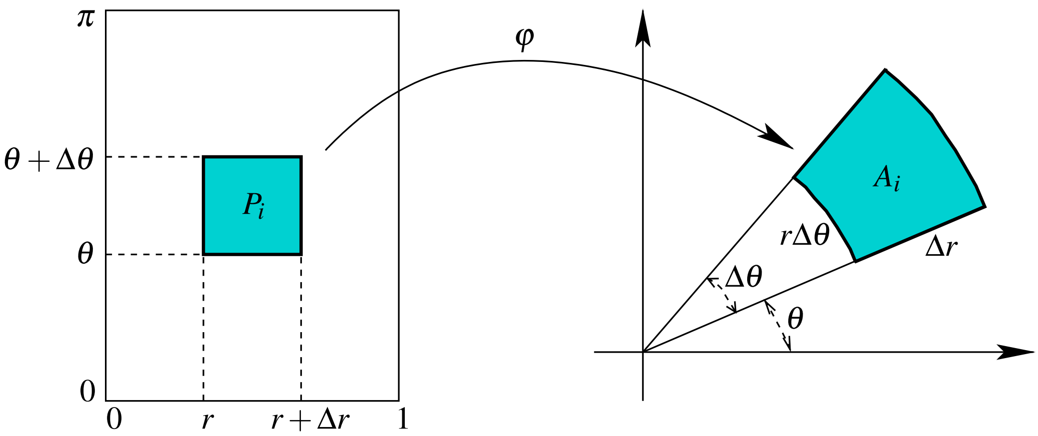
Dès lors, $\det \partial\varphi(r,\theta) = \det \frac{\partial\varphi}{\partial(r,\theta)}(r,\theta) = r\cos^2\theta + r\sin^2\theta = r > 0$ pour tout $(r,\theta) \in]0,1[\times]0,2\pi[$. Finalement, le théorème de changement de variables dit :

$$\int_D 1 \,\mathrm{d}(x,y) = \int_{]0,1[\times]0,2\pi[} 1 \cdot r \,\mathrm{d}(r,\theta).$$

Puisque l'intégrale de droite porte sur un rectangle, on peut lui appliquer le théorème de Fubini, ce qui donne

$$\int_D 1 \,\mathrm{d}(x,y) = \int_{]0,1[} \Big(\int_{]0,2\pi[} 1 \cdot r \,\mathrm{d}\theta \Big) \mathrm{d}r = \int_0^1 \Big(r \int_0^{2\pi} \mathrm{d}\theta \Big) \mathrm{d}r = \pi$$

Finissons par une représentation graphique de ce changement de variable qui explique pourquoi le déterminant de $\partial\varphi(r,\theta)$ est r.



Si on considère un petit élément d'aire carré P_i au point (r,θ), disons de longueur Δr et de hauteur $\Delta\theta$, son image $A_i = \varphi(P_i)$ est un arc d'anneau. Si on approxime A_i par un rectangle de cotés Δr et $r\Delta\theta$, on voit que son aire est plus ou moins égale à $r\Delta\theta\Delta r$, c'est-à-dire $r\,\mathrm{aire}(P_i)$. Ce facteur r est exactement celui donné par la différentielle de φ au point (r,θ).

VI.6 Exercices

Exercice VI.1. Calculez l'aire de l'ellipse (pleine) dont l'équation est $(x/a)^2 + (y/b)^2 \leqslant 1$.

Exercice VI.2. En utilisant un changement de variables en coordonnées sphériques, calculez le volume d'une sphère de rayon $R > 0$. (Expliquez en détail comment vous contournez le fait que ce changement de variables n'est pas un $\mathscr{C}^1$-difféomorphisme.)

Exercice VI.3. Soit une fonction $f : [a,b] \to \mathbb{R}^{\geqslant 0} : x \mapsto f(x)$ une fonction continue. Montrez que le volume de l'ensemble A délimité par rotation de f autour de l'axe des x, i.e., de

$$A = \left\{(x,y,z) \in \mathbb{R}^3 : x \in [a,b] \text{ et } \sqrt{y^2+z^2} \leqslant f(x)\right\},$$

est donné par $\pi \int_a^b f(x)^2 \,\mathrm{d}x$.

Exercice VI.4. Calculez le volume du tore dont le rayon du trou est $R > 0$ et le rayon de la partie pleine du tore est $\rho > 0$.

Notations

Ensembles

$\mathbb{N}$	ensemble des naturels : 0, 1, 2, 3,...
$\mathbb{Z}$	ensemble des entiers : ..., -2, -1, 0, 1, 2,...
$\mathbb{Q}$	ensemble des nombres rationnels, c'est-à-dire des fractions p/q où $p, q \in \mathbb{Z}$ avec $q \neq 0$.
$\mathbb{R}$	ensemble des nombres réels ; ceux-ci comprennent les rationnels mais aussi toutes les limites des suites rationelles de Cauchy (voir sections I et I.5).

Fonctions

$f\vert_A$	restriction de la fonction f à l'ensemble A. Si $f : X \circ\!\!\rightarrow Y$, sa restriction à $A \subseteq X$ est la fonction $f\vert_A : A \circ\!\!\rightarrow Y : x \mapsto f(x)$ avec $\operatorname{Dom}(f\vert_A) = A \cap \operatorname{Dom} f$.
$\mathbb{1}_X$	l'identité sur un ensemble X définie comme $\mathbb{1}_X : X \to X : x \mapsto x$.
pr_i	la projections sur la i^{e} composante : $\mathrm{pr}_i(x_1, \dots, x_N) = x_i$.
$\lceil \cdot \rceil$	$\lceil x \rceil$ est le plus petit entier plus grand ou égal à $x \in \mathbb{R}$. (Son existence dépend de l'axiome d'Archimède, voir page 32).

Alphabet grec

A	α	alpha	H	η	êta	N	ν	nu	T	τ	tau
B	β	beta	Θ	θ	theta	Ξ	ξ	xi	Y	υ	upsilon
Γ	γ	gamma	I	ι	iota	O	o	omicron	Φ	φ	phi
Δ	δ	delta	K	κ	kappa	Π	π	pi	X	χ	chi
E	ε	epsilon	Λ	λ	lambda	P	ρ	rho	Ψ	ψ	psi
Z	ζ	zeta	M	μ	mu	Σ	σ	sigma	Ω	ω	omega

Bibliographie

[1] E. LANDAU *Foundations of analysis*, Chelsea Publishing company New York.

[2] J. MAWHIN, *Analyse. Fondements, techniques, évolution*, De Boeck & Larcier (1997).

[3] E.W. SWOKOWSKI, *Analyse*, 5e édition (1993), De Boeck Université.

Index