

Topologie des espaces vectoriels normés

Romain JOLY

Dernière mise à jour : novembre 2022

Utilise tcolorbox de Thomas F. Sturm

Table des matières

Avant-propos

Chapitre 1 : Les nombres	1
1 Une rapide histoire des nombres	1
2 Un aperçu de la construction des nombres	2
3 Cardinalité	5
4 Compléments	9
Chapitre 2 : Rappels sur la topologie de la droite réelle	13
1 Les structures de \mathbb{R}	13
2 Les propriétés fondamentales de \mathbb{R}	19
3 Suites réelles	21
4 Fonctions réelles	29
5 Quelques applications	38
Chapitre 3 : Topologie des espaces vectoriels normés	47
1 Espaces vectoriels normés	47
2 Topologie des espaces vectoriels normés : définitions de base	61
3 Suites d'un espace vectoriel normé	71
4 Équivalence de normes	76
5 Compléments	80
Chapitre 4 : Fonctions sur les espaces vectoriels normés	87
1 Définitions générales	87
2 Limites	89
3 Continuité	92
4 Applications linéaires	99
5 Compléments	104
Chapitre 5 : Complétude	107
1 Suites de Cauchy et espaces complets	107
2 Convergence normale des séries et applications	109
3 Le point fixe des applications contractantes	113
Chapitre 6 : Compacité	119
1 Les compacts	119
2 Fonctions et compacité	126

Chapitre 7 : Compléments	129
1 Connexité par arc	129
2 Homéomorphismes	132

Avant-propos

La *topologie*, de *topos* et *logos*, est littéralement « l'étude d'un lieu ». Dans le sens qui nous intéressera, il s'agit de comprendre la géométrie d'un espace. Plus précisément, on souhaite le munir d'une notion de distance et comprendre comment définir une notion de limite de suite, de fonction continue... et quelles sont les propriétés découlant de ces notions.

La topologie est née progressivement il y a 100-150 ans. Au XIX^{ème} siècle, certains domaines de l'analyse deviennent tellement pointus qu'ils donnent des résultats apparemment paradoxaux. Il s'en suit un besoin de clarifier des preuves et notions qui restaient souvent basées sur l'intuition. Des mathématiciens comme Cauchy décident de remettre au propre toute l'analyse pour la fonder sur des bases solides. En même temps, le nouveau formalisme rigoureux permet de comprendre qu'on peut généraliser bien des notions. On peut par exemple définir des distances entre des fonctions, considérer des fonctions continues sur des espaces de fonctions...

Les objectifs de ce cours sont :

1. avoir un aperçu historique de cette refonte de l'analyse,
2. rappeler les bases de l'analyse sur \mathbb{R} ,
3. les généraliser à des espaces plus complexes,
4. posséder les bases de l'analyse pour le CAPES ou la grande majorité des métiers des mathématiques.

Il s'agit d'un cours de troisième année de licence de mathématiques (L3) qui cible le parcours B. Il existe des versions bien plus générales de la topologie, étudiant des espaces topologiques abstraits, sans notion de distance. Nous nous restreindrons ici aux espaces vectoriels normés et, de fait, nos exemples seront quasi-exclusivement dans \mathbb{R}^d ou $\mathcal{C}^0(I, \mathbb{R})$ muni des normes L^p .

Chapitre 1 : Les nombres

Ce chapitre est une divagation autour des ensembles de nombres. Ce sera l'occasion de discuter de leur construction formelle et de s'intéresser à la notion de cardinalité. Bien que certains points seront utiles pour la suite du cours, ce chapitre doit surtout être considéré pour la culture mathématique. Nous ne prétendons pas ici être complet sur les sujets abordés ni avoir toute la rigueur exigible.

1 Une rapide histoire des nombres

L'idée de nombre s'est construite avec les civilisations et l'écriture, c'est-à-dire très récemment à l'échelle de l'histoire de l'humanité.

- \mathbb{N}^* et \mathbb{Q}_+^* : les entiers naturels et les fractions positives sont utilisés depuis le début des temps historiques (premiers écrits). Toutes les cultures du monde ont des noms pour les premiers nombres entiers et des façons de les représenter. Les écritures des nombres ont été très diverses et plus ou moins pratiques. Il faut noter qu'il existe encore des peuples de chasseurs-cueilleurs qui n'ont pas de mot (et donc de représentation mentale) pour les nombres au-dessus de quelques unités (les Pirahãs comptent « un, deux, beaucoup »). Les grands nombres et le calcul sont donc des affaires de civilisations.
- $(\mathbb{R} \setminus \mathbb{Q})_+$: on sait que $\sqrt{2}$ est irrationnel depuis les pythagoriciens (vers 500 avant J.C). Les grecs de l'Antiquité avaient la vision des nombres comme des longueurs et donc se plaçant sur une ligne droite. Dans ce sens, on peut dire qu'ils voyaient les nombres positifs comme un continuum et donc incluaient les irrationnels.
- \mathbb{Z}^* et \mathbb{Q}_-^* : les nombres négatifs sont utilisés en Chine et en Inde deux siècles avant notre ère. Cela consiste à noter avec des couleurs différentes les dettes et de connaître les règles d'opération sur les signes. Ils ne seront utilisés que bien plus tard chez les arabes et en occident.
- L'écriture décimale et le chiffre zéro apparaissent en Inde vers le Vème siècle avant d'être adoptés par la civilisation arabe. D'ailleurs, ceux que nous appelons « chiffres arabes » étaient qualifiés d'« indiens » par al-Khwārizmī (IXème siècle, Perse). Ce dernier rédige un manuel d'utilisation de l'écriture décimale qui sera traduit et participera à l'essor de cette écriture en occident pendant la renaissance. Ceci explique que « chiffre » se dit « algarismo » en portugais

(une autre partie de l'œuvre d'al-Khwārizmī fera que son nom donnera aussi le mot « algorithmes » et le titre d'un de ces livres notre mot « algèbre »).

- 0 : le nombre zéro n'apparaît que vers 500 en Inde. Il ne faut pas le confondre avec le chiffre zéro qui ne sert qu'à indiquer une position vide dans une écriture positionnelle (parfois un espace blanc ou un dessin ont joué le même rôle que le chiffre zéro).
- \mathbb{C} : pendant la renaissance italienne, Tartaglia, Cardan et Ferrari développent la résolution des équations de degré 3 et 4. Leurs formules peuvent conduire à des calculs du type $\sqrt{-1} - \sqrt{-1}$. A priori, cela n'a pas de sens, mais si on admet $\sqrt{-1}$ comme un nombre « imaginaire », alors le calcul donne 0 et on obtient la solution qu'on sait exister. Il s'agit donc au début d'une astuce pour faire aboutir une méthode de calcul mais pas de nombres en tant que tels. C'est Raphaël Bombelli (1526-1572, Italie) puis Leonhard Euler (1707-1783, Suisse) qui feront de ces nombres de vrais nombres, bien que « complexes ».

2 Un aperçu de la construction des nombres

2.1 Entiers et fractions

Lors de la refondation des mathématiques, les mathématiciens ont cherché à recréer toutes les théories en partant du minimum de bagages : quelques axiomes fondant la théorie des ensembles. Comment créer les nombres entiers avec le moins de choses possibles ? Une façon très simple est due à John von Neumann (1903-1957, Hongrie et USA), un des pères de l'informatique. Il suffit de poser 0 comme l'ensemble vide \emptyset , puis 1 comme le singleton $\{0\} = \{\emptyset\}$, 2 est l'ensemble $\{0,1\} = \{\emptyset, \{\emptyset\}\}$ etc. On construit donc par récurrence l'entier n comme l'ensemble contenant tous les nombres précédents. Cela fournit naturellement l'opération « +1 ». Il faut ensuite expliquer comment retrouver les opérations standards. Par exemple, l'addition $n + 2$ est définie par $(n + 1) + 1$ et ainsi de suite : l'addition générale $n + m$ est définie comme étant $(n + (m - 1)) + 1$. Retrouver toutes les propriétés des entiers naturels devient juste un jeu de patience et de rigueur. Coder les nombres négatifs est simple puisqu'il suffit de rajouter un *bit* devant le nombre entier.

Une fraction est juste un couple d'entiers. Mais si on définit \mathbb{Q} comme $\mathbb{Z} \times \mathbb{N}^*$, on a deux couples différents (1,2) et (2,4) alors qu'il s'agit de la même fraction 1/2. Donc on va définir \mathbb{Q} comme étant $\mathbb{Z} \times \mathbb{N}^*$ quotienté par la relation d'équivalence $(p,q) \equiv (p',q')$ si $pq' = p'q$. Notez bien que le test d'égalité reste dans les entiers puisqu'on ne sait pas ce qu'est une fraction (c'est tout le sel du jeu). On peut alors définir la somme de deux fractions par $(p,q) + (p',q') = (pq' + p'q, qq')$ et continuer avec la multiplication, l'inverse etc.

2.2 L'écriture décimale

L'écriture décimale est une écriture positionnelle en base 10. Un nombre décimal est un nombre que l'on peut écrire exactement avec un nombre fini de chiffres.

Définition 1.1

Un nombre **décimal** est un nombre rationnel du type $p/10^k$ avec p et k dans \mathbb{Z} .

Proposition 1.2

Le nombre r est décimal si et seulement s'il existe un couple $n > m$ dans \mathbb{Z} et une suite de chiffres $a_n, \dots, a_m \in \{0, 1, 2, \dots, 9\}$ telle qu'on peut l'écrire sous la forme

$$a_n a_{n-1} \dots a_1 a_0, a_{-1} \dots a_m := \sum_{i=n}^m a_i 10^i$$

avec $a_n \neq 0$ et $a_m \neq 0$. Par ailleurs, si elle existe, cette écriture est unique.

On voit que des fractions comme $1/3$ ou $1/7$ ne sont pas des décimaux. On peut alors imaginer les coder et même créer tous les nombres réels en poursuivant l'écriture décimale avec un nombre infini de chiffres après la virgule. On a bien alors

$$0,333333\dots = \sum_{k=1}^{+\infty} 3 \cdot 10^{-k} = 3 \frac{10^{-1}}{1 - 10^{-1}} = \frac{3}{9} = \frac{1}{3}.$$

Formellement, on est en train de construire les réels comme limite des décimaux. C'est certainement la façon dont on les envisage dans la vie courante. Mais on rencontre un problème :

$$0,99999\dots = \sum_{k=1}^{+\infty} 9 \cdot 10^{-k} = 9 \frac{10^{-1}}{1 - 10^{-1}} = 1$$

ce qui veut dire qu'il existe deux écritures pour le même nombre 1. On introduit donc la définition suivante.

Définition 1.3

Un développement décimal infini $a_n a_{n-1} \dots a_1 a_0, a_{-1} a_{-2} \dots$ est dit **impropre** s'il existe i_0 tel que $a_i = 9$ pour tout $i \leq i_0$.

Pour ne plus avoir de doublon dans la représentation décimale des réels, on n'autorise que les développements décimaux propres. Si on a réglé le principal problème, il reste encore définir les règles d'addition et multiplication sur ces nombres décimaux pour retrouver les structures des réels. Or ceci est très pénible, en particulier à cause des écritures impropres : $0,444\dots + 0,555\dots$ n'est pas égal à $0,999\dots$! C'est pour cela qu'une construction par les nombres décimaux n'est pas jugée pratique. Par ailleurs, elle dépend fortement de la base choisie, ce que n'aiment pas trop les mathématiciens : un nombre existe indépendamment de la base dans laquelle on l'écrit.

2.3 Autres constructions des réels

Imaginons que nous avons construits les nombres rationnels. Nous voulons représenter tous les nombres réels et les opérations liées d'une façon qui n'utilise que des nombres rationnels. Cela permettra de s'assurer que les réels ne sont pas une invention absurde qui pourrait aboutir à une contradiction logique, mais bien une structure découlant des axiomes de base des mathématiques. Il a été proposé plusieurs approches pour ce faire, dont les deux principales sont les suivantes. Il faut évidemment avoir en tête que ces constructions ont été faites a posteriori, alors que les nombres réels étaient utilisés depuis longtemps.

Les coupures de Dedekind

L'idée de Dedekind est de représenter un nombre par une coupure (A, B) de \mathbb{Q} , i.e. deux ensembles complémentaires non vides A et B tels que pour tout $a \in A$ et $b \in B$, $a \leq b$. Il faut vraiment voir la coupure comme un trait coupant \mathbb{Q} en deux parties : A à gauche et B à droite. Ces coupures peuvent déjà représenter les rationnels en posant qu'un rationnel $q \in \mathbb{Q}$ est représenté par

$$A_q = \{r \in \mathbb{Q}, r < q\} \text{ et } B_q = \{r \in \mathbb{Q}, q \leq r\} .$$

On a le choix de mettre q dans A ou B . Il faut une convention et c'est celle que l'ensemble A ne doit pas avoir de plus grand élément. On a donc $q = \min B_q$ pour les rationnels. Mais on peut aussi regarder la coupure

$$A = \{r \in \mathbb{Q}, r^2 < 2 \text{ ou } r \leq 0\} \text{ et } B = \{r \in \mathbb{Q}, 2 < r^2 \text{ et } 0 \leq r\}$$

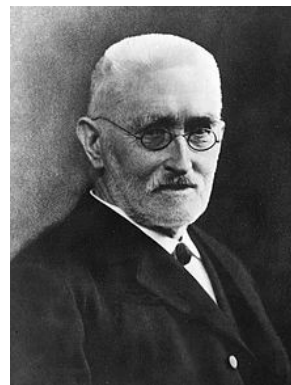
qui est une coupure admise, entièrement décrite par des rationnels, mais qui n'est pas de la forme (A_q, B_q) avec $q \in \mathbb{Q}$ puisque B n'a pas de minimum dans \mathbb{Q} . C'est donc un nouveau nombre, qui va correspondre à $\sqrt{2}$. Pour vraiment construire la théorie, il faut retrouver toutes les propriétés de \mathbb{R} , bien définir ce qu'est la somme de deux coupures, leur produit etc. Par exemple, $(A, B) \leq (A', B')$ peut être défini par $A \subset A'$. De même, on peut définir l'addition de coupures par $(A, B) + (A', B') = (A + A', B + B')$. On montre ainsi qu'on obtient une bonne représentation de ce qu'on appelle « nombre réel ».

Les suites de Cauchy

Louis Augustin Cauchy a eu lui l'idée d'introduire une notion de suite qui assurerait la convergence sans pour autant introduire la limite elle-même. Une suite $(u_n) \subset \mathbb{Q}$ est dite « de Cauchy » si

$$\forall \varepsilon \in \mathbb{Q}_+^* , \exists n_0 \in \mathbb{N} , \forall n, m \geq n_0 , |u_n - u_m| \leq \varepsilon .$$

Notons que cette définition n'utilise que des nombres rationnels (la définition plus standard des suites numériques de Cauchy sera rappelée dans le chapitre suivant). Une suite convergente dans \mathbb{Q} est de Cauchy mais certaines suites sont de Cauchy



Richard Dedekind
1831-1916
Allemagne

sans converger dans \mathbb{Q} car leur potentielle limite n'y est pas. On définit alors les réels comme toutes les limites possibles de suite de Cauchy de rationnels. Comme plusieurs suites donnent la même limite, on quotiente l'ensemble des suites de Cauchy par l'équivalence $(u_n) \equiv (v_n)$ si et seulement si $|u_n - v_n| \rightarrow 0$. L'ensemble des classes d'équivalence des suites de Cauchy de rationnels est \mathbb{R} suivant cette construction. La définition de l'addition et de la multiplication est simple car il suffit d'additionner et multiplier les suites (en vérifiant que c'est compatible avec les classes d'équivalence). Le procédé a aussi l'avantage d'être utilisable dans beaucoup d'autres situations mathématiques où il faut « compléter » un espace où certaines suites de Cauchy ne convergent pas. En contrepartie, certaines propriétés de \mathbb{R} sont plus difficiles à obtenir selon cette construction.



Louis Augustin Cauchy
1789-1857
France

3 Cardinalité

Le cardinal $|A|$ d'un ensemble A correspond à la taille de cet ensemble. On pourrait se limiter à distinguer simplement les ensembles de taille finie et ceux de taille infinie. Mais en fait, il existe plusieurs degré d'infini. On formalise cela avec les définitions suivantes. Dans toute la suite, si $a \leq b$ sont deux entiers, on note

$$\llbracket a, b \rrbracket := \{n \in \mathbb{Z}, a \leq n \leq b\}.$$

Définition 1.4 Cardinal d'un ensemble

Soient A et B deux ensembles. On dit que A et B sont **équipotents** ou **ont même cardinal** s'il existe une bijection entre A et B . On note alors $|A| = |B|$.

On dit que A est de cardinal plus petit que B s'il existe une injection de A dans B et on note alors $|A| \leq |B|$.

On dit que A est de cardinal **fini** s'il existe $n \in \mathbb{N}$ tel que $|A| = \llbracket 1, n \rrbracket$, c'est-à-dire s'il existe une bijection entre A et $\{1, 2, \dots, n\}$. On note plus simplement $|A| = n$ et on dit que le cardinal de A est n . Si A n'est pas de cardinal fini, on dit qu'il est **infini**.

Si A est de cardinal fini ou si $|A| = |\mathbb{N}|$ on dit que A est **dénombrable**. Sinon A est dit **indénombrable**.

Un ensemble dénombrable est un ensemble A dont on peut lister les éléments un par un sous la forme $A = \{a_0, a_1, a_2, \dots\}$. Ce n'est qu'au XIX^{ème} siècle que l'on a compris qu'il existe des ensembles indénombrables, c'est-à-dire de taille encore plus grande que \mathbb{N} . On verra ci-dessous que c'est le cas de \mathbb{R} . On peut construire toute une théorie de la cardinalité. Un des théorèmes importants est le suivant, que nous admettrons ici.

Théorème 1.5 Cantor-Schröder-Bernstein

Soient A et B deux ensembles, si $|A| \leq |B|$ et si $|B| \leq |A|$, alors $|A| = |B|$. Autrement dit, s'il existe une injection de A vers B et une de B vers A , alors on peut construire une bijection entre A et B .

Pour donner un peu l'esprit de ces preuves, faisons celle d'un résultat plus facile.

Proposition 1.6

Soient A et B deux ensembles non vides. On a $|A| \leq |B|$ si et seulement s'il existe une surjection de B vers A .

Démonstration : Supposons que $|A| \leq |B|$. Par définition, il existe une injection $f : a \in A \mapsto f(a) \in B$. Comme A est non vide, on peut choisir $a_* \in A$. On construit notre surjection $g : B \rightarrow A$ comme suit. Si b est dans l'image de f , on pose $g(b) = a$ avec $a \in A$ tel que $f(a) = b$. Si b n'est pas dans l'image de f , on pose $g(b) = a_*$.

Pour la réciproque, considérons une surjection $g : b \in B \mapsto g(b) \in A$. Pour tout $a \in A$, il existe au moins un $b_* \in B$ tel que $g(b_*) = a$. On pose simplement $f(a) = b_*$. Il est clair que f est une injection car si $f(a) = b_* = b'_* = f(a')$, $g(b_*) = g(b'_*)$ et donc $a = a'$. On notera au passage que pour chaque a , nous avons dû choisir un b_* dans l'ensemble $g^{-1}(a)$. Ça paraît évident, mais en fait cette opération n'est pas anodine et ne peut pas forcément être faite pour tout ensemble de points A quelle que soit sa taille. Elle demande l'utilisation d'un postulat appelé « axiome du choix ». Nous ne rentrerons pas dans les détails ici et de toute façon, nous n'utiliserons pas la proposition 1.6 dans un contexte si général que l'axiome du choix sera indispensable. \square

Nous allons chercher quels ensembles de nombres sont dénombrables. Évidemment, \mathbb{N} est dénombrable par définition. Avant de faire des propositions plus abstraites, nous allons faire à la main les cas de \mathbb{Z} et \mathbb{Q} .

Proposition 1.7

L'ensemble \mathbb{Z} est dénombrable.

Démonstration : On liste les entiers relatifs sous la forme $z_0 = 0$, $z_1 = 1$, $z_2 = -1$, $z_3 = 2$, $z_4 = -2$ et ainsi de suite. \square

Proposition 1.8

L'ensemble \mathbb{Q} est dénombrable.

Démonstration : On pose $r_0 = 0$, puis $r_1 = 1$ et $r_2 = -1$, puis $r_3 = 2$, $r_4 = 1/2$, $r_5 = -1/2$ et $r_6 = -2$, puis $r_7 = 3$, $r_8 = 3/2$ etc. À l'étape n , on liste les fractions du type p/q avec $p \in \llbracket -n, n \rrbracket$ et $q \in \llbracket 1, n \rrbracket$, en sautant celles qu'on a déjà listées. Notons qu'on avance progressivement car chaque étape ne contient qu'un nombre fini de nouvelles fractions. \square

En fait, ces propositions peuvent se déduire des résultats plus généraux suivants, puisque $\mathbb{Z} = \mathbb{N} \cup (-\mathbb{N})$ et \mathbb{Q} est une partie de $\mathbb{Z} \times \mathbb{N}^*$ si r est identifié à sa forme irréductible p/q .

Proposition 1.9

Un ensemble est dénombrable si et seulement s'il s'injecte dans \mathbb{N} . En conséquence, un sous-ensemble d'un ensemble dénombrable est dénombrable.

Démonstration : laissée au lecteur. \square

Proposition 1.10

Une union dénombrable d'ensembles dénombrables est dénombrable.

Démonstration : Soient A_i , $i \in J$ un ensemble d'ensembles avec J dénombrable. Par définition, quitte à utiliser une bijection pour changer les indices, on peut supposer que J est fini ou $J = \mathbb{N}$. Par la suite, on va faire le cas $J = \mathbb{N}$, le cas J fini étant plus simple et laissé au lecteur. On sait qu'on peut lister les éléments de chaque ensemble A_i sous la forme $A_i = \{a_i^0, a_i^1, a_i^2, \dots\}$. On va dénombrer les éléments de $\cup_{i \in \mathbb{N}} A_i$ suivant la liste

$$a_0^0 \quad a_1^0 \quad a_0^1 \quad a_2^0 \quad a_1^1 \quad a_0^2 \dots$$

où à l'étape n , on liste les éléments a_i^j tels que $i + j \leq n$. On obtient ainsi une surjection de \mathbb{N} dans $\cup_{i \in \mathbb{N}} A_i$ (car les A_i peuvent s'intersecter) et la proposition 1.6 conclut. \square

Proposition 1.11

Un produit fini d'ensembles dénombrables est dénombrable.

Démonstration : Soit A_i , $i = 0, \dots, p$ des ensembles dénombrables que l'on liste sous la forme $A_i = \{a_i^0, a_i^1, a_i^2, \dots\}$. On liste les éléments de $A_0 \times \dots \times A_p$ en listant à l'étape n tous les p -uplets $(a_1^{j_1}, a_2^{j_2}, \dots, a_p^{j_p})$ tels que $j_1 + j_2 + \dots + j_p \leq n$. Chaque étape ne contient qu'un nombre fini de p -uplets et chaque p -uplet sera bien compté lors d'une étape. \square

Une des plus jolies preuves sur la cardinalité correspond au paradoxe du « bar-
bier qui ne se rase pas lui-même ». On la doit à Cantor. Elle permet de montrer
que l'on peut construire des ensembles toujours plus grands et qu'il existe ainsi
une infinité de degrés dans la cardinalité infinie. Nous nous limiterons à la nuance
dénombrable/indénombrable dans ce cours.

Théorème 1.12 Cantor

Soit A un ensemble et $\mathcal{P}(A)$ l'ensemble des parties de A . Alors $|A| < |\mathcal{P}(A)|$.

Démonstration : Comme on a l'injection $f : a \in A \mapsto \{a\} \in \mathcal{P}(A)$, on a
 $|A| \leq |\mathcal{P}(A)|$. Supposons qu'il existe une surjection $g : A \rightarrow \mathcal{P}(A)$. On regarde
l'ensemble

$$E = \{a \in A, a \notin g(a)\} .$$

Il s'agit d'un sous-ensemble de A , donc il existe $e \in A$ tel que $g(e) = E$. Si on
a $e \in E$, alors par définition de E , $e \notin g(e) = E$, ce qui est absurde. Mais si
on avait $e \notin E = g(e)$, alors par définition de E , on devrait avoir $e \in E$, ce qui
est encore absurde. Donc l'existence de la surjection est absurde et d'après la
proposition 6, le cardinal de $\mathcal{P}(A)$ est strictement plus grand que celui de A .
□

Le théorème de Cantor montre que \mathbb{R} n'est pas
dénombrable en associant au développement en base
2 d'un réel $x \in]0,1[$ l'ensemble des entiers n tels que le
 n -ième chiffre de x est 1. Mais il existe une preuve bien
plus frappante, encore due à Cantor.



*Georg Cantor
(1845-1918)
Allemagne*

Proposition 1.13

L'ensemble $]0,1[$ est indénombrable.

Démonstration : On utilise l'argument diagonal de Cantor, qui est un rai-
sonnement par l'absurde. Imaginons avoir listé tous les réels de $]0,1[$ dans un
tableau de taille infinie :

$$\begin{aligned} x_1 &= 0, a_1^1 a_1^2 a_1^3 a_1^4 \dots \\ x_2 &= 0, a_2^1 a_2^2 a_2^3 a_2^4 \dots \\ x_3 &= 0, a_3^1 a_3^2 a_3^3 a_3^4 \dots \\ x_4 &= 0, a_4^1 a_4^2 a_4^3 a_4^4 \dots \\ &\vdots \end{aligned}$$

On construit un réel $y \in]0,1[$ en prenant comme i -ème chiffre un chiffre
différent de a_i^i et de 9 (et au moins un chiffre non nul pour éviter d'avoir
 $y = 0$). On obtient un réel avec une écriture décimale propre et y n'est pas

dans le tableau puisque s'il y était à la n -ième ligne, alors son n -ième chiffre serait a_n^n , ce qui est exclu par construction. Donc notre liste ne peut pas être exhaustive. \square

Proposition 1.14

Tout intervalle non trivial de \mathbb{R} est équipotent à \mathbb{R} . En conséquence, l'ensemble \mathbb{R} ainsi que tout intervalle non trivial de \mathbb{R} est indénombrable.

Démonstration : La fonction arctangente fournit une bijection entre \mathbb{R} et $] -\pi/2, \pi/2[$. L'exponentielle fournit une bijection entre \mathbb{R} et $]0, +\infty[$. Deux intervalles du type $]a, b[$ sont mis en bijection par de simples homothéties et translations, de même pour les intervalles d'autres types comme $[a, b]$, $]a, b]$, $[a, +\infty[$ etc. On peut obtenir des bijections explicites entre des intervalles de type différents en utilisant des arguments comme celui de l'hôtel de Hilbert (voir les compléments de ce chapitre). Si on ne cherche pas une forme explicite, on peut simplement utiliser le théorème 1.5 : tous les intervalles sont contenus dans \mathbb{R} et tous les intervalles non triviaux contiennent un intervalle du type $]a, b[$ qui est équipotent à \mathbb{R} . \square

Une conséquence assez troublante de ce résultat est la suivante. Comme écrire ou manipuler un nombre prend un certain laps de temps, aussi bien pour les humains que leurs ordinateurs, et qu'il n'y a qu'un nombre fini d'humains, alors on n'utilisera pendant toute l'histoire de l'humanité qu'une partie négligeable des nombres réels. De même, on ne pourra jamais trouver un système d'écriture permettant de décrire tous les réels en un nombre fini de caractères.

4 Compléments

4.1 Rationnels et irrationnels

Les nombres rationnels sont dénombrables alors que les réels sont indénombrables. Cela montre non seulement qu'il existe des irrationnels, mais surtout qu'il y en a « beaucoup plus » que de rationnels. De façon plus concrète, on connaît explicitement de nombreux irrationnels. Le premier connu serait $\sqrt{2}$. La preuve suivante suit la méthode d'Euclide (vers -300, Grèce).

Proposition 1.15 (irrationalité de $\sqrt{2}$)

Le nombre $\sqrt{2}$ ne peut s'écrire sous la forme a/b avec a et b entiers. En particulier, la diagonale d'un carré n'est pas commensurable à son côté.

Démonstration : Commençons par rappeler que le carré d'un nombre pair est pair car $(2p)^2 = 4p^2 = 2(2p^2)$ et que le carré d'un nombre impair est impair

car $(2p + 1)^2 = 4p^2 + 4p + 1 = 2(2p^2 + 2p) + 1$.

Supposons qu'il existe deux entiers positifs a et b tels que $\sqrt{2} = a/b$. On a alors $2b^2 = a^2$, et donc a^2 est pair. Comme le carré d'un impair est impair, ceci implique que a est pair et s'écrit $a = 2a'$ avec a' entier. Mais alors $2b^2 = a^2 = 4a'^2$ et donc $b^2 = 2a'^2$ est pair. On en déduit que b est pair et s'écrit $b = 2b'$ avec b' entier. On a donc aussi que $\sqrt{2} = a'/b'$.

Mais on pourrait de nouveau appliquer l'argument et diviser par deux chacun des nombres de la fraction et ainsi de suite une infinité de fois, ce qui est absurde. Donc notre hypothèse de départ est fautive et $\sqrt{2}$ ne peut s'écrire sous forme d'une fraction. \square

Cette démonstration s'adapte aux autres racines carrées qui ne sont pas entières. Ainsi, on peut montrer que $\sqrt{5}$ et donc le nombre d'or $\phi = (1 + \sqrt{5})/2$ sont irrationnels. On sait aussi que e est irrationnel (1737, Leonhard Euler) ainsi que π (1768, Jean-Henri Lambert).

Il existe un lien surprenant entre les rationnels et l'écriture décimale.

Proposition 1.16

Un nombre $x \in \mathbb{R}$ est rationnel si et seulement si son développement décimal est périodique à partir d'un certain rang.

Démonstration : Supposons que x a un développement décimal périodique à partir d'un certain rang, c'est-à-dire qu'il s'écrit sous la forme

$$x = a_n \dots a_0, a_{-1} \dots a_m b_1 \dots b_p b_1 \dots b_p b_1 \dots b_p \dots$$

On pose $A = a_n \dots a_m$ et $B = b_1 \dots b_p$, qui sont deux entiers. Par définition de l'écriture décimale, on a donc

$$x = A.10^m + B \sum_{k \geq 1} 10^{m-kp}.$$

Il ne reste plus qu'à voir que

$$\sum_{k \geq 1} 10^{m-kp} = \frac{10^{m-p}}{1 - 10^{-p}}$$

est rationnel.

Réciproquement, soit $x = p/q$ un rationnel. Le développement décimal de x se trouve en faisant la division de p par q selon la méthode classique. À chaque étape, on a un reste entre 0 et $p - 1$. Comme il y a un nombre fini de restes possibles, on tombe forcément à un moment sur un reste déjà vu. Dans ce cas, l'algorithme de division boucle et on retombe sur une suite de décimale qui vont se répéter. \square

Ceci permet de créer de nombreux irrationnels. Par exemple 0,1010010001000010...

est irrationnel ainsi que $0,12345678910111213\dots$

4.2 L'hôtel de Hilbert

L'hôtel de Hilbert est une illustration des paradoxes sur l'infini dont la paternité est attribuée à David Hilbert.

Imaginons un hôtel possédant une infinité (dénombrable) de chambres numérotées $1, 2, 3, \dots$ et imaginons que toutes les chambres sont occupées. Un nouveau client se présente, on pourrait penser qu'il n'aura pas de chambre pour lui. Mais l'hôtelier demande à chaque client de passer de la chambre n à la chambre $n + 1$. La première chambre devient libre et le nouveau client peut s'y installer. Mathématiquement, on a

montré que $\{-1\} \cup \mathbb{N}$ est équipotent à \mathbb{N} .

Imaginons maintenant qu'un car contenant une infinité de clients arrive à l'hôtel. Toutes les chambres sont prises, mais l'hôtelier demande à chaque client de passer de la chambre n à la chambre $2n$. Toutes les chambres impaires deviennent libres et les nouveaux clients peuvent s'y installer. En termes mathématiques, on a montré que \mathbb{Z} est équipotent à \mathbb{N} .

On peut utiliser la même idée pour obtenir des bijections explicites plus complexes.



David Hilbert
(1862-1943)
Allemagne

Exemple :

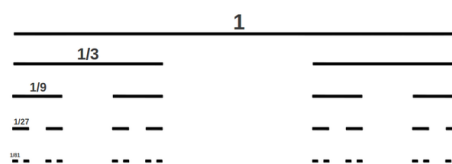
On veut obtenir une bijection explicite entre $]0,1]$ et $]0,1[$. On va utiliser le premier argument de l'hôtel avec la suite $1/n$. On pose $f(x) = 1/(n + 1)$ si $x = 1/n$ pour $n \in \mathbb{N}^*$ et $f(x) = x$ sinon. On a poussé d'un cran les « clients » des places $1/n$ pour caser le 1.

Notons que cette bijection n'est pas continue. On montrera dans ce cours qu'il n'est pas possible de construire une bijection f continue entre $]0,1]$ et $]0,1[$ pour des raisons topologiques, c'est-à-dire par l'essence même de ce qu'on apprendra dans ce cours.

4.3 L'ensemble de Cantor

L'ensemble de Cantor est un des fractals les plus simples. De ce fait, c'est un bon candidat pour obtenir des contre-exemples à des propriétés que l'on penserait vraies pour les ensembles de réels si on se limiterait à l'intuition que nous donnent les intervalles.

Il se construit ainsi : on part de $[0,1]$ et on supprime le tiers central pour ne garder que $[0,1/3] \cup [2/3,1]$. Puis pour chaque intervalle restant, on recoupe en trois et on élimine le tiers central, puis on répète ainsi l'opération à l'infini. On peut montrer que l'ensemble obtenu est sous la forme



Les premières étapes de construction de l'ensemble de Cantor

$$\mathcal{C} = \left\{ x \in [0,1] \mid \exists (\epsilon_n)_{n \geq 1} \in \{0,1\}^{\mathbb{N}}, x = \sum_{n \geq 1} \frac{2\epsilon_n}{3^n} \right\}.$$

Cette description correspond à dire qu'on ne garde que les $x \in [0,1]$ tels que x s'écrive en base 3 qu'avec les chiffres 0 et 2, en autorisant les écritures impropres. En effet, retirer le tiers central de $[0,1]$ correspond à ne pas prendre les nombres qui s'écrivent $0,1\dots$ en base deux. On s'autorise juste à prendre $0,1 = 0,022222\dots$. Le découpage de l'étape n élimine les nombres qui utilisent le chiffre 1 en position n .

Pour tout n , l'ensemble de Cantor est contenu dans l'ensemble de l'étape n qui est composé de 2^n intervalles de longueurs $1/3^n$. Il est donc de « mesure » plus petite que $(2/3)^n$ pour tout n . On dit que l'ensemble de Cantor est *de mesure nulle* (un sens plus rigoureux sera vu au second semestre). L'ensemble des rationnels est aussi de mesure nulle, mais ce qui est particulier à l'ensemble de Cantor, c'est qu'il est de mesure nulle tout en étant indénombrable. C'est ainsi un exemple d'ensemble indénombrable qui ne contient aucun segment $]a,b[$.

Pour montrer que l'ensemble de Cantor est indénombrable, on peut considérer la surjection suivante

$$f : x = \sum_{n \geq 1} \frac{2\epsilon_n}{3^n} \in \mathcal{C} \longmapsto f(x) = \sum_{n \geq 1} \frac{\epsilon_n}{2^n} \in [0,1].$$

Autrement dit, on associe à l'écriture en base 3 avec des 0 et des 2, l'écriture en base 2 avec des 0 et des 1. On obtient une surjection qui n'est pas une bijection à cause des écritures impropres. Mais cela suffit à dire que le cardinal de \mathcal{C} est plus grand que celui de $[0,1]$ et donc que \mathcal{C} est indénombrable.

Chapitre 2 : Rappels sur la topologie de la droite réelle

La plus grande partie du contenu de ce chapitre a normalement déjà été vue pendant les années précédentes. Nous n'allons pas redonner toutes les preuves ni même toutes les définitions (par exemple pour les différentes limites des fonctions réelles). Notre but est :

- de faire des rappels pour se rafraîchir la mémoire. En particulier, le lecteur pourra en profiter pour se replonger dans ses cours précédents et revoir les points dont il n'est plus très sûr.
- d'énoncer quelques résultats fondamentaux des réels qui nous serviront pour des démonstrations des chapitres suivants.
- d'observer comment les différentes notions topologiques sont définies dans \mathbb{R} pour pouvoir les généraliser à des espaces plus complexes dans la suite de ce cours.

1 Les structures de \mathbb{R}

1.1 Structure algébrique

Les opérations standards sur les réels sont l'addition et la multiplication ainsi que leurs inverses. Mathématiquement, on peut voir dans \mathbb{R} plusieurs structures.

L'ensemble \mathbb{R} muni de l'addition standard $+$ est un **groupe commutatif** :

- la somme de deux réels est un réel,
- le zéro est un élément neutre pour l'addition,
- pour tout $x \in \mathbb{R}$ il existe un inverse $-x$ pour l'addition,
- la somme est associative car $(a + b) + c = a + (b + c)$ et commutative dans le sens où $a + b = b + a$.

Si on y ajoute la multiplication standard \times , il s'agit d'un **corps commutatif** :

- \mathbb{R} muni de $+$ est un groupe commutatif ayant 0 comme élément neutre,
- $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$ muni de la multiplication est un groupe commutatif d'élément neutre 1,
- la multiplication est distributive par rapport à l'addition car $(a + b) \times c = a \times c + b \times c$.

Mais on peut le voir comme un cas très particulier d'**espace vectoriel réel** qui a juste une dimension :

- la somme de deux réels est réelle
- si $\lambda \in \mathbb{R}$ est un scalaire et $x \in \mathbb{R}$ un vecteur, alors $\lambda x \in \mathbb{R}$ est un vecteur.

Cette dernière structure peut paraître très artificielle mais nous allons étudier dans ce cours la topologie des espaces vectoriels réels. Il est donc intéressant de voir que \mathbb{R} en est un cas particulier dont nous allons nous inspirer pour le cas général.

1.2 La notion de distance

Pour généraliser des concepts comme la continuité des fonctions à d'autres ensembles que les réels, il nous faut avoir une vision plus géométrique de ces définitions.

La **distance** entre deux réels x et y est $|x - y|$.

Un petit intervalle $]x - \varepsilon, x + \varepsilon[$ avec $\varepsilon > 0$ est un voisinage du point x . Il s'agit de tous les réels à distance au plus ε de x . C'est donc aussi la boule ouverte de centre x et de rayon ε . Plus généralement, on appelle **voisinage** de $x \in \mathbb{R}$ tout ensemble qui contient une boule $]x - \varepsilon, x + \varepsilon[$ avec $\varepsilon > 0$.

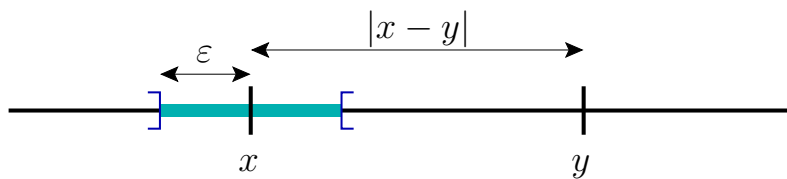


FIGURE 2.1 – la distance $|x - y|$ entre deux nombres réels correspond bien à la notion usuelle. Un voisinage de x contient un petit intervalle autour de x .

Il sera important de traduire toute phrase avec quantificateurs sous une forme plus géométrique et intuitive. Ainsi la convergence $u_n \rightarrow \ell$ qui s'écrit

$$\forall \varepsilon > 0, \exists n_0 \in \mathbb{N}, \forall n \geq n_0, |u_n - \ell| < \varepsilon$$

peut se lire « *Quel que soit le petit voisinage de ℓ considéré, la suite (u_n) finit par entrer dans ce voisinage et y rester.* ». Vu comme cela, on comprend qu'on peut généraliser la notion de limite pour des suites dans des espaces plus généraux une fois qu'on a défini un concept de *voisinage* ou de *boule*.

Comme on l'a vu, la valeur absolue est une fonction primordiale pour faire de l'analyse sur \mathbb{R} . Voici quelques rappels concernant la valeur absolue :

- si $x \geq 0$, $|x| = x$ et si $x \leq 0$, $|x| = -x$. En particulier, $|x| = |-x|$.
- $|x| \leq a$ est équivalent à $-a \leq x \leq a$. En particulier, les boules sur \mathbb{R} sont des intervalles car $\{y \in \mathbb{R}, |x - y| < \varepsilon\} =]x - \varepsilon, x + \varepsilon[$.
- $|x| \geq a$ est équivalent à $(x \leq -a$ ou $x \geq a)$.

- $|x \times y| = |x| \times |y|$

Une des propriétés fondamentales de la valeur absolue est la suivante.

Proposition 2.1

L'inégalité triangulaire s'énonce

$$|a + b| \leq |a| + |b|.$$

Il y a égalité si et seulement a et b ont même signe.

Pour $b = -c$, on obtient une majoration pour une différence $|a - c| \leq |a| + |c|$. C'est une inégalité triangulaire sur les distances si on l'écrit sous la forme

$$|x - y| = |x - z + z - y| \leq |x - z| + |y - z|.$$

On trouve bien qu'aller directement de x à y est toujours plus court que de passer par z . Il est aussi important de connaître une deuxième version de l'inégalité triangulaire.

Proposition 2.2

Pour tous réels a et b

$$||a| - |b|| \leq |a + b| \quad \text{et} \quad ||a| - |b|| \leq |a - b|.$$

Démonstration : On utilise la première version de l'inégalité triangulaire

$$|a| = |a + b - b| \leq |a + b| + |b|$$

et donc $|a + b| \geq |a| - |b|$. Mais les rôles de a et b sont symétriques, donc on trouve aussi $|a + b| \geq |b| - |a|$ (quitte à refaire l'argument en changeant les rôles de a et b). Pour la deuxième inégalité, il suffit de changer b en $-b$. \square

1.3 La structure d'ordre et les bornes supérieures

Nous avons l'habitude de munir \mathbb{R} d'une structure d'ordre. Il est notable que cette structure est compatible avec les opérations naturelles sur \mathbb{R} . Par exemple, si $a > a'$ et $b > b'$, on a $a + b > a' + b'$. On a aussi les règles connues concernant la multiplication (attention aux changements de sens pour les négatifs). Ce n'est pas anodin puisqu'une telle structure d'ordre n'est pas toujours possible. Par exemple, on ne peut munir \mathbb{C} d'une structure d'ordre qui se comporte raisonnablement bien par rapport aux opérations complexes (mais on peut munir \mathbb{C} d'une structure d'ordre quand même, par exemple $a + ib > a' + ib'$ si $a > a'$ ou si $a = a'$ et $b > b'$).

On introduit les notions suivantes.

Définition 2.3

Soit $A \subset \mathbb{R}$ un sous-ensemble des nombres réels.
On dit que $M \in \mathbb{R}$ est **un majorant** de A si

$$\forall a \in A, \quad a \leq M.$$

On dit que A est **majoré** s'il existe un réel M qui majore A .

On dit que $M \in \mathbb{R}$ est **la borne supérieure** de A , notée $\sup(A)$, si c'est le plus petit majorant de A , c'est-à-dire que c'est un majorant et que tout nombre plus petit n'est plus un majorant.

Si $M \in \mathbb{R}$ est la borne supérieure de A et que M appartient à A , alors on dit que M est **le maximum** de A et on le note $\max(A)$.

Exemples :

- Le segment $]0,1[$ est majoré par 2. Sa borne supérieure est 1 car d'une part tout $x \in]0,1[$ est plus petit que 1 et, d'autre part, tout nombre $1 - \varepsilon$ avec $\varepsilon > 0$ n'est pas majorant car $x = \max(1 - \varepsilon/2, 1/2)$ est dans $]0,1[$ et est plus grand que $1 - \varepsilon$. Par contre 1 n'appartient pas à $]0,1[$ donc 1 n'est pas le maximum de $]0,1[$ et écrire $\max(]0,1[)$ n'a pas de sens.
- $\{x \in \mathbb{Q}, x \leq 2\}$ a 2 pour maximum.
- $\{x \in \mathbb{Q}, x^2 \leq 2\}$ a $\sqrt{2}$ pour borne supérieure mais n'a pas de maximum.

Si $A = \emptyset$ est vide, alors A est majoré par tous les réels car une proposition concernant tous les éléments de \emptyset est trivialement vérifiée (il n'y a aucun élément à considérer!). L'ensemble des majorants de l'ensemble vide est donc \mathbb{R} tout entier et il n'y a pas de plus petit majorant. Donc $A = \emptyset$ n'a pas de borne supérieure. De même, un ensemble non majoré n'a aucun majorant et donc pas de borne supérieure. Mais pour simplifier les notations, il peut être agréable de quand même faire un abus de notation utilisant les infinis.

Définition 2.4

Si $A = \emptyset$ est vide, on pose $\sup(A) = -\infty$.

Si $A \subset \mathbb{R}$ n'est pas majorée, on pose $\sup(A) = +\infty$.



Attention : il ne s'agit que d'une notation. C'est commode pour énoncer des propositions sans avoir à faire plusieurs cas, mais il faut se méfier si on veut utiliser ces définitions comme des nombres concrets puisque des calculs comme $\infty - \infty$ n'ont pas de sens.

On procède de même pour la borne inférieure.

Définition 2.5

Soit $A \subset \mathbb{R}$ un sous-ensemble des nombres réels.
On dit que $m \in \mathbb{R}$ est **un minorant** de A si

$$\forall a \in A, \quad m \leq a .$$

On dit que A est **minoré** s'il existe un réel m qui minore A .

On dit que $m \in \mathbb{R}$ est **la borne inférieure** de A , notée $\inf(A)$, si c'est le plus grand minorant de A , c'est-à-dire que c'est un minorant et que tout nombre plus grand n'est plus un minorant.

Si $m \in \mathbb{R}$ est la borne inférieure de A et que m appartient à A , alors on dit que m est **le minimum** de A et on le note $\min(A)$.

De nouveau, pour simplifier les notations, on peut écrire

Définition 2.6

Si $A = \emptyset$ est vide, on pose $\inf(A) = +\infty$.

Si $A \subset \mathbb{R}$ n'est pas minorée, on pose $\inf(A) = -\infty$.

Et naturellement

Définition 2.7

Un ensemble à la fois majoré et minoré est dit **borné**.

Exemples :

- Remarquons que, par définition, l'ensemble vide est borné.
- Un intervalle du type $]a,b]$ est borné. Il admet un maximum qui est b et une borne inférieure qui est a mais n'a pas de minimum.
- Un intervalle du type $[a, +\infty[$ est minoré mais pas majoré, il n'est donc pas borné et n'admet pas de borne supérieure (ni de maximum). Il admet a comme minimum (et donc aussi comme borne inférieure).
- Supposons que $A \subset \mathbb{R}$ est un ensemble pour lequel on a trouvé une borne $M > 0$ telle que pour tout $a \in A$, $|a| \leq M$. Alors A est borné et contenu dans $[-M, M]$.
- L'ensemble $\{x \in \mathbb{R}, x^2 \leq 1\}$ est borné (c'est le segment $[-1, 1]$). Mais l'ensemble $\{x \in \mathbb{R}, 1 \leq x^2\}$ n'est ni majoré, ni minoré (il est égal à $] - \infty, 1] \cup [1, + \infty[$). Attention donc à bien décrypter ce qu'est l'ensemble avant de décider s'il est majoré ou minoré (ne pas se fier juste aux sens des inégalités le définissant).

De ces définitions découlent des propriétés classiques. Nous n'allons pas toutes les énoncer et plusieurs seront vues en TD. Voici quelques exemples dont l'intérêt réside

aussi dans la preuve. Dans tous les cas, on admet l'existence des bornes supérieures qui sera énoncée dans le théorème 2.12 plus loin (et qui sera de toute façon admise).

Proposition 2.8

Soit A un ensemble non vide majoré de réel. L'ensemble des majorants de A est exactement $[\sup A, +\infty[$.

Démonstration : Soit $M \in \mathbb{R}$. Si $M < \sup A$, alors M n'est pas un majorant de A puisque $\sup A$ est le plus petit d'entre eux. Si $M \geq \sup A$, alors comme $\sup A$ majore A , pour tout $a \in A$, $a \leq \sup A \leq M$. Donc M majore A . \square

Proposition 2.9

Si A et B sont deux ensembles non vides majorés de réels avec $A \subset B$, alors $\sup A \leq \sup B$.

Démonstration : Par définition, $\sup B$ est un majorant de B . Soit $a \in A$, comme $A \subset B$, a est aussi dans B et donc $a \leq \sup B$. Comme a est quelconque, on vient de montrer que $\sup B$ est un majorant de A . Comme $\sup A$ est le plus petit d'entre eux, on a $\sup A \leq \sup B$. \square

Proposition 2.10

Si A et B sont deux ensembles non vides majorés de réels, alors $\sup(A \cup B) = \max(\sup A, \sup B)$.

Démonstration : Soit $M = \max(\sup A, \sup B)$. Pour tout $x \in A \cup B$, soit $x \in A$ et $x \leq \sup A \leq M$ (car $\sup A$ majore A), soit $x \in B$ et $x \leq \sup B \leq M$. Dans les deux cas, $x \leq M$ et donc M majore $A \cup B$.

Soit $m < M$. Supposons que $M = \sup A$ (sinon $M = \sup B$ et le raisonnement est symétrique). Comme $m < \sup A$, m n'est pas majorant de A (car par définition, $\sup A$ est le plus petit d'entre eux). Donc il existe $a \in A$ tel que $a > m$. Comme a est aussi dans $A \cup B$, m ne majore pas $A \cup B$ et M est bien le plus petit des majorants. \square

Proposition 2.11

Soit $A \subset \mathbb{R}$ un ensemble non vide et majoré. Les propriétés suivantes sont équivalentes :

- (i) le nombre M est la borne supérieure de A ,
- (ii) le nombre M majore A et il existe une suite (a_n) d'éléments de A qui converge vers M .

Démonstration : Si (ii) est vérifiée, on sait que M majore A donc il reste à montrer que c'est le plus petit majorant. Soit $x < M$ et soit $\varepsilon = M - x > 0$. Comme (a_n) tend vers M , il existe un rang n_0 à partir duquel $a_n > M - \varepsilon = x$. Donc $a_{n_0} > x$ et x ne peut pas majorer A . Donc M est bien la borne supérieure de A .

Supposons maintenant que (i) est vraie, i.e. que M est la borne supérieure de A . Par définition, M majore A . Nous allons construire la suite (a_n) comme suit. Pour tout $n \in \mathbb{N}$, on considère $M - 2^{-n}$. Comme M est le plus petit majorant, $M - 2^{-n}$ ne majore pas A et il existe au moins un élément a_n de A tel que $a_n > M - 2^{-n}$. Par ailleurs, comme M majore A , on a aussi $a_n \leq M$. Donc par théorème des gendarmes, on a bien $a_n \rightarrow M$. \square

2 Les propriétés fondamentales de \mathbb{R}

Dans cette partie, nous allons voir plusieurs propriétés importantes des réels. Il s'agit de propriétés fondamentales dans le sens où elles proviennent plus ou moins directement de la construction même de \mathbb{R} . Pour les démontrer, il faudrait donc avoir défini concrètement les réels selon une des constructions possibles et utiliser les axiomes de la construction. Comme nous n'avons pas fait le travail de construction précisément, nous ne pourrions donc pas démontrer ces propriétés que nous admettrons.

• Propriété de la borne supérieure

Une des propriétés fondamentales des réels est l'existence d'une borne supérieure.

Théorème 2.12

Tout ensemble $A \subset \mathbb{R}$ non vide et majoré admet une borne supérieure.
 Tout ensemble $A \subset \mathbb{R}$ non vide et minoré admet une borne inférieure.

Exemple :

L'ensemble

$$A = \{x \in \mathbb{R}, x^5 - 3x - 1 \leq 0\}$$

est majoré et contient $x = 0$. Il admet donc une borne supérieure, mais qu'on ne peut pas écrire à l'aide des fonctions usuelles. C'est donc le théorème 2.12 qui donne l'existence de cette borne supérieure même si on ne peut pas écrire ce qu'elle vaut exactement.

Même si elle peut paraître naturelle, il s'agit d'une propriété *topologique* fondamentale de \mathbb{R} . Prenons l'ensemble $A = \{r \in \mathbb{Q}, r^2 < 2\} \subset \mathbb{Q}$ qui est un ensemble de rationnels dont la définition n'a utilisé que des rationnels. Clairement, A est non vide car $0 \in A$ et A est majoré car si $x \in A$ alors $x \leq 2$ et pourtant A n'admet pas de borne supérieure dans \mathbb{Q} . Mais si on étend notre vue à tous les réels, c'est-à-dire qu'on regarde A comme sous-ensemble de \mathbb{R} , alors A admet une borne supérieure qu'on note $\sqrt{2}$. Donc le théorème 2.12 n'est plus vrai si on remplace \mathbb{R} par \mathbb{Q} et

c'est la prise en compte de tous les réels qui permet d'obtenir toutes les bornes supérieures.

Il s'agit d'une propriété par nature assez élémentaire selon la construction de Dedekind, mais plus délicate à obtenir par la construction selon les suites de Cauchy.

• **Complétude de \mathbb{R}**

Rappelons la définition des suites de Cauchy.

Définition 2.13

Une suite de réels $(u_n)_{n \in \mathbb{N}} \subset \mathbb{R}$ est dite **de Cauchy** si

$$\forall \varepsilon > 0, \exists n_0 \in \mathbb{N}, \forall p \geq q \geq n_0, |u_p - u_q| < \varepsilon.$$

On dit que \mathbb{R} est complet car il vérifie la propriété suivante.

Théorème 2.14

Toute suite de Cauchy de \mathbb{R} admet une limite dans \mathbb{R} .

Ce résultat est important car il permet de montrer qu'une suite converge sans même savoir quelle pourrait en être la limite. Il est à la base de nombreux résultats. Notons encore que tous les ensembles n'ont pas cette propriété puisque \mathbb{Q} possède des suites de Cauchy qui ne convergent pas dans \mathbb{Q} (par exemple si on considère une suite d'approximations rationnelles de $\sqrt{2}$).

Les théorèmes 2.12 et 2.14 énoncent des résultats en fait très semblables puisqu'ils reviennent à dire que, intuitivement, \mathbb{R} « n'a pas de trous ».

• **Densité des rationnels**

Nous introduirons plus tard une notion de densité plus générale. Dans \mathbb{R} , cette notion s'écrit comme suit.

Définition 2.15

Un sous-ensemble $A \subset \mathbb{R}$ est dit **dense** s'il vérifie une des caractérisations équivalentes suivantes :

i) chaque voisinage d'un point x de \mathbb{R} contient un point de A

$$\forall x \in \mathbb{R}, \forall \varepsilon > 0, \exists a \in A, a \in]x - \varepsilon, x + \varepsilon[.$$

ii) A rencontre tout intervalle ouvert de \mathbb{R}

$$\forall (x, y) \in \mathbb{R}^2, x < y \implies (\exists a \in A, x < a < y).$$

iii) tout réel peut être approché par une suite de A

$$\forall x \in \mathbb{R}, \exists (a_n)_{n \in \mathbb{N}} \subset A, a_n \xrightarrow[n \rightarrow +\infty]{} x.$$

La construction des réels nous donne assez facilement que \mathbb{Q} est dense dans \mathbb{R} .

Il est un peu moins évident de voir que c'est aussi le cas des irrationnels.

Théorème 2.16

Les ensembles \mathbb{Q} et $\mathbb{R} \setminus \mathbb{Q}$ sont tous les deux denses dans \mathbb{R} .

Cela implique par exemple qu'entre deux réels, il y a une infinité de rationnels et aussi une infinité d'irrationnels.

• Corps ordonné archimédien

La définition suivante peut paraître triviale, mais il s'agit d'une propriété importante des réels. Les grecs l'avaient déjà identifiée puisqu'on trouve dans le livre V des *Éléments* d'Euclide la définition « *des grandeurs sont dites avoir une raison entre elles lorsque ces grandeurs, étant multipliées, peuvent se surpasser mutuellement* ».

Définition 2.17

Un corps ordonné K est **archimédien** si pour tout x et y dans K avec $y > x > 0$, il existe un entier n tel que $n.x > y$.

On peut utiliser le résultat suivant comme une définition des réels.

Théorème 2.18

L'ensemble \mathbb{R} muni de ses structures habituelles est l'unique corps ordonné archimédien vérifiant la propriété de la borne supérieure.

Par « unique corps », il faut comprendre qu'un autre corps qui aurait les mêmes propriétés peut être mis en bijection avec \mathbb{R} avec une bijection transportant exactement toutes les structures en question (l'image de la somme est la somme des images, l'ordre des images est le même que celui des antécédents etc.).

Exemple :

On considère la suite définie par $u_n = 1/n$ pour $n \in \mathbb{N}^*$. Soit $\varepsilon > 0$. Comme \mathbb{R} est archimédien, il existe $n_0 \in \mathbb{N}^*$ tel que $n_0\varepsilon > 1$. Mais alors $0 < 1/n_0 < \varepsilon$ et donc $|u_{n_0}| < \varepsilon$. Les règles liant relation d'ordre et opérations sur \mathbb{R} , qui viennent du fait que \mathbb{R} est un corps ordonné, montrent que pour $n \geq n_0$, $0 < u_n \leq u_{n_0}$. Donc pour tout $n \geq n_0$, $|u_n| < \varepsilon$. On vient de montrer que la suite $u_n = 1/n$ tend vers 0 grâce aux propriétés fondamentales de \mathbb{R} .

3 Suites réelles

Une suite réelle $(u_n)_{n \in \mathbb{N}} \subset \mathbb{R}$ est une fonction de \mathbb{N} dans \mathbb{R} . Il est plus agréable de la voir comme une liste ordonnée u_0, u_1, u_2, \dots de réels, l'indice $n \in \mathbb{N}$ étant un temps discret qui s'écoule. Le sujet principal d'étude des suites est de comprendre leur comportement quand $n \rightarrow +\infty$. Rappelons la définition classique.

Définition 2.19

Une suite réelle $(u_n)_{n \in \mathbb{N}} \subset \mathbb{R}$ admet le réel $\ell \in \mathbb{R}$ comme **limite** si

$$\forall \varepsilon > 0, \exists n_0 \in \mathbb{N}, \forall n \geq n_0, |u_n - \ell| < \varepsilon.$$

On note $\lim_{n \rightarrow +\infty} u_n = \ell$ ou $u_n \xrightarrow[n \rightarrow +\infty]{} \ell$.

On dit aussi que (u_n) **converge** ou **tend vers** ℓ .

On dit (u_n) est **convergente**, ou bien converge, s'il existe $\ell \in \mathbb{R}$ tel que (u_n) tend vers ℓ . Dans le cas contraire, on dit que (u_n) est **divergente**.

On retrouve ci-dessus une transcription dans le formalisme des quantificateurs de Cauchy de ce que Gottfried Wilhelm Leibniz (1646-1716, Allemagne) tentait d'écrire avec des phrases comme

« à mesure qu'on la considère de plus en plus loin, l'erreur sera moindre que toute grandeur donnée. »

On peut aussi définir une notion de convergence vers $\pm\infty$.

Définition 2.20

Soit $(u_n)_{n \in \mathbb{N}}$ une suite réelle. On dit que (u_n) a pour **limite** $+\infty$ si

$$\forall M \in \mathbb{R}, \exists n_0 \in \mathbb{N}, \forall n \geq n_0, u_n \geq M.$$

Symétriquement, on dit que (u_n) a pour **limite** $-\infty$ si

$$\forall M \in \mathbb{R}, \exists n_0 \in \mathbb{N}, \forall n \geq n_0, u_n \leq M.$$

Attention, si on peut parler de limite infinie, on considère bien qu'une suite qui tend vers $\pm\infty$ est divergente.

Nous n'allons pas reprendre toutes les propriétés des limites réels ni leurs démonstrations. Nous allons reprendre seulement les propriétés suivantes, d'une part car elles nous seront utiles pour la suite et d'autre part pour bien montrer qu'elles proviennent des propriétés fondamentales de \mathbb{R} .

Théorème 2.21

Toute suite croissante majorée converge (vers une limite réelle finie). Symétriquement, toute suite décroissante minorée converge (vers une limite réelle finie).

Démonstration : Si (u_n) est croissante et majorée, alors l'ensemble $\{u_n, n \in \mathbb{N}\}$ est un ensemble majoré non vide de \mathbb{R} . Il admet donc une borne supérieure $\ell \in \mathbb{R}$. Par construction, comme ℓ est un majorant, $u_n \leq \ell$ pour tout $n \in \mathbb{N}$. Comme ℓ est le plus petit majorant, pour tout $\varepsilon > 0$, $\ell - \varepsilon$ n'est plus un majorant et il existe $n_0 \in \mathbb{N}$ tel que $u_{n_0} > \ell - \varepsilon$. Comme la suite est croissante,

on a que pour tout $n \geq n_0$, $u_n \geq u_{n_0} > \ell - \varepsilon$. On a donc que pour tout $n \geq n_0$, $u_n \in]\ell - \varepsilon, \ell]$ et donc que $|u_n - \ell| < \varepsilon$.

Le cas symétrique se démontre de la même façon. \square

Nous avons utilisé la propriété de la borne supérieure. Cette proposition n'est pas vraie dans \mathbb{Q} car la suite des approximations décimales de π (écriture décimale tronquée à n chiffres) est une suite de rationnels croissante et majorée mais qui ne converge pas dans \mathbb{Q} .

Corollaire 2.22

Si $(u_n)_{n \in \mathbb{N}}$ est une suite croissante, alors on a l'alternative suivante :

- i) soit (u_n) est majorée et converge vers une limite finie,
- ii) soit (u_n) n'est pas majorée et diverge vers $+\infty$.

Démonstration : Si (u_n) est majorée, alors le théorème 2.21 conclut. Supposons donc que (u_n) n'est pas majorée : pour tout $M \in \mathbb{R}$, il existe forcément $n_0 \in \mathbb{N}$ tel que $u_{n_0} > M$ puisque M ne peut pas être un majorant. Mais comme la suite est croissante, on a alors pour tout $n \geq n_0$, $u_n \geq u_{n_0} > M$. Par définition, cela veut dire que (u_n) tend vers $+\infty$. \square

Un autre résultat important utilisant une des propriétés fondamentales de \mathbb{R} est celui des suites adjacentes.

Théorème 2.23 (suites adjacentes)

Soient (u_n) et (v_n) deux suites que l'on suppose **adjacentes** c'est-à-dire que :

- i) la suite (u_n) est croissante,
- ii) la suite (v_n) est décroissante,
- iii) on a $|u_n - v_n| \rightarrow 0$ quand n tend vers $+\infty$.

Alors (u_n) et (v_n) convergent toutes les deux vers une même limite $\ell \in \mathbb{R}$ et on a l'encadrement

$$\forall n \in \mathbb{N}, u_n \leq \ell \leq v_n.$$

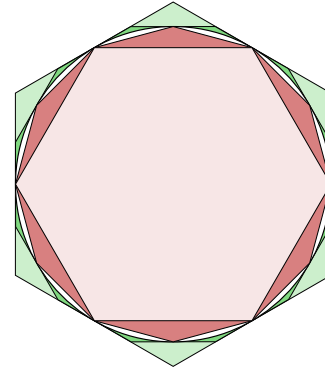
Démonstration : D'après la définition de la convergence, il existe un rang $n_0 \in \mathbb{N}$ tel que, pour tout $n \geq n_0$, $|u_n - v_n| \leq 1$. En particulier, à partir de ce rang, $v_n \geq u_n - 1$. D'après la monotonie des suites, si $n \geq n_0$, on a $v_n \geq u_n - 1 \geq u_{n-1} - 1 \geq \dots \geq u_0 - 1$. Pour $n \leq n_0$, on a aussi $v_n \geq v_{n_0} \geq u_{n_0} - 1 \geq u_0 - 1$. La suite (v_n) est donc décroissante et minorée, elle converge vers une limite $\ell \in \mathbb{R}$ d'après le théorème 2.21. Par ailleurs, la preuve de ce théorème montre bien que ℓ est la borne inférieure de (v_n) et donc que $\ell \leq v_n$ pour tout n .

On peut faire le raisonnement symétrique pour (u_n) et montrer qu'elle converge vers une limite $\ell' \in \mathbb{R}$. Il reste à voir qu'il s'agit en fait de la même limite. Mais

comme les deux suites convergent, leur différence $u_n - v_n$ tend vers $\ell' - \ell$ et donc $|u_n - v_n|$ tend vers $|\ell' - \ell|$. Par hypothèse, ce nombre est égal à 0 donc $\ell = \ell'$. \square

Exemples :

- Un exemple très visuel est donné par la méthode de quadrature du cercle qu'a utilisée Archimède au III^{ème} siècle avant J.C. pour donner l'encadrement $3 + 10/71 < \pi < 3 + 1/7$. On note u_n la surface du polygone à 6×2^n côtés inscrit dans le cercle et v_n la surface du polygone à 6×2^n côtés circonscrit au cercle. On part donc d'un hexagone puis on double le nombre de côtés comme sur la figure jointe. Il est clair que (u_n) est croissante et (v_n) décroissante et on peut se



convaincre que $u_n - v_n \rightarrow 0$. Le calcul de u_n et v_n peut se faire par itération de formules trigos. On obtient deux suites adjacentes qui donnent des approximations de plus en plus proches de π . Archimède avait été jusqu'à $n = 5$, c'est-à-dire pour des polygones à 96 côtés, obtenant les trois premiers chiffres significatifs de π . Notons que cette construction de suites adjacentes nous fournit l'existence du nombre qu'on appelle π .

- On considère les suites (u_n) et (v_n) données par

$$u_n = \sum_{k=0}^n \frac{1}{k!} = 1 + 1 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} + \dots + \frac{1}{n!} \quad \text{et} \quad v_n = u_n + \frac{1}{n!}.$$

Comme u_n est une somme de plus en plus grande de termes positifs, la suite (u_n) est clairement croissante. On peut montrer que (v_n) est décroissante (calcul laissé au lecteur ou aux TDs). Comme $|u_n - v_n| = 1/n! \rightarrow 0$, il s'agit de deux suites adjacentes. La puissance du théorème 2.23 est de montrer qu'elles convergent, même si on ne sait rien de cette limite. Il se trouve que leur limite est un nombre important en mathématique, on va donc lui donner un nom : c'est le nombre e . On peut ainsi s'autoriser à écrire

$$e = \sum_{k=0}^{\infty} \frac{1}{k!} = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \frac{1}{5!} + \dots$$

Le résultat suivant s'appuie sur les mêmes arguments mais avec un point de vue plus géométrique.

Théorème 2.24 (théorème des segments emboîtés)

Soient $I_n = [a_n, b_n]$ une suite de segments fermés bornés qui sont emboîtés dans le sens où $I_{n+1} \subset I_n$. Alors l'intersection $\bigcap_{n \in \mathbb{N}} I_n$ est non vide.

Démonstration : Comme les segments sont emboîtés, la suite (a_n) est croissante et majorée par b_0 et la suite (b_n) est décroissante et minorée par a_0 . Toutes les deux convergent donc vers des limites a_∞ et b_∞ . On vérifie facilement que $\bigcap_{n \in \mathbb{N}} I_n = [a_\infty, b_\infty]$. \square

Exemples :

- On procède à une recherche par dichotomie. On part du segment $I_0 = [0,1]$ que l'on coupe en deux et on choisit l'une des moitiés $I_1 = [0,1/2]$ ou $I_2 = [1/2,1]$. Puis I_1 est de nouveau coupé en deux et on choisit pour I_2 l'une des moitiés etc. Le théorème 2.24 nous assure qu'il restera au moins un point à l'intersection de tous ces choix (en fait juste un seul car la taille des segments tend vers 0).
- Revenons à l'ensemble de Cantor avec la vision géométrique. On part du segment $I_0 = [0,1]$ puis on coupe l'intervalle en deux en retirant le tiers central. On choisit pour I_1 l'un des tiers restants et on le redécoupe puis on choisit pour I_2 un des tiers restants etc. De nouveau, on est assuré qu'il reste un point à l'intersection de tous ces choix. Non seulement cela montre que l'ensemble de Cantor est non vide, mais cela montre qu'on peut exactement le mettre en bijection avec les suites infinies de choix du type $DGGGDDGGDGG\dots$. Par l'argument diagonal de Cantor, on montre que l'ensemble de ces suites est indénombrable.
- On fait attention que, même si la propriété peut paraître évidente, elle repose sur des propriétés topologiques. Par exemple, elle n'est plus vraie sur \mathbb{Q} puisque $\bigcap_{n \in \mathbb{N}} [a_n, b_n] = \emptyset$ si (a_n) et (b_n) sont des approximations décimales de $\sqrt{2}$ par en-dessous et par au-dessus à 10^{-n} près. Si on reste sur \mathbb{R} , il faut faire attention à la topologie des intervalles. Ainsi $\bigcap_{n \in \mathbb{N}}]0, 2^{-n}] = \emptyset$ et $\bigcap_{n \in \mathbb{N}} [n, +\infty[= \emptyset$.

Une suite convergente est forcément bornée, mais il existe des suites bornées qui ne convergent pas comme $(-1)^n$. La propriété fondamentale suivante correspond à ce que nous appellerons plus tard la *compacité*. Elle permet d'obtenir quand même une convergence, quitte à ne considérer qu'une partie de la suite. Son nom vient des mathématiciens Bernard Bolzano (1781-1848, Hongrie) et Karl Weierstrass (1815-1897, Allemagne).

Théorème 2.25 (théorème de Bolzano-Weierstrass)

Soit $(u_n)_{n \in \mathbb{N}} \subset \mathbb{R}$ une suite réelle bornée. Alors il existe une fonction $\varphi : \mathbb{N} \rightarrow \mathbb{N}$ strictement monotone telle que la suite extraite $(u_{\varphi(n)})_{n \in \mathbb{N}}$ converge vers une limite $\ell \in \mathbb{R}$.

Démonstration : Notre raisonnement consiste à appliquer une méthode de dichotomie. Comme (u_n) est bornée, il existe un segment $[a_0, b_0]$ tel que $u_n \in [a_0, b_0]$ pour tout $n \in \mathbb{N}$. On prend $\varphi(0) = 0$. Soit $m = (a_0 + b_0)/2$ le milieu de $[a_0, b_0]$. Comme il y a une infinité de termes de la suite dans $[a_0, b_0]$, il faut bien

qu'il y en ait une infinité dans au moins une des moitié $[a_0, m]$ et/ou $[m, b_0]$ (principe des tiroirs). On note $[a_1, b_1]$ une des moitiés qui convient et on pose $\varphi(1)$ comme le premier $n > \varphi(0) = 0$ tel que u_n soit dans $[a_1, b_1]$. De nouveau, on prend le milieu $m = (a_1 + b_1)/2$. Comme il y a une infinité de termes de la suite dans $[a_1, b_1]$, il faut bien qu'il y en ait une infinité dans au moins une des moitié $[a_0, m]$ et/ou $[m, b_0]$, que l'on note $[a_2, b_2]$. On pose $\varphi(2)$ comme le premier $n > \varphi(1)$ tel que u_n soit dans $[a_2, b_2]$. De proche en proche, on peut extraire ainsi une sous-suite. À partir du rang n_0 , tous les termes de la sous-suite se retrouvent dans un intervalle de taille $2^{-n_0}|b_0 - a_0|$ puisqu'on divise l'intervalle considéré par deux à chaque étape. Cela montre que cette sous-suite est une suite de Cauchy. Le théorème 2.14 nous dit qu'elle converge donc. \square

Quand on parle de compacité, on pense plutôt à prendre une suite de points dans un ensemble. La version topologique du théorème précédent s'énoncera plutôt comme suit.

Corollaire 2.26 (compacité simplifiée dans \mathbb{R})

Soient $a < b$ deux réels. Soit $(x_n) \subset [a, b]$ une suite de points de l'intervalle fermé borné $[a, b]$, alors on peut en extraire une sous-suite $(x_{\varphi(n)})$ qui converge vers une limite $\ell \in [a, b]$.

Démonstration : Comme $a \leq x_n \leq b$ pour tout n , la suite est bornée et on peut utiliser le théorème précédent pour en extraire une sous-suite $(x_{\varphi(n)})$ qui converge vers une limite $\ell \in \mathbb{R}$. Comme $x_{\varphi(n)} \geq a$ pour tout n , les théorèmes de comparaison nous disent que $\ell \geq a$. De même, on obtient que $\ell \leq b$, ce qui conclut. \square

Les limites possibles des suites extraites forment ce qu'on appelle les *valeurs d'adhérence* de la suite.

Définition 2.27

Soit $(u_n)_{n \in \mathbb{N}}$ une suite réelle. On dit que $\ell \in \mathbb{R}$ est une **valeur d'adhérence** de (u_n) s'il existe une sous-suite $(u_{\varphi(n)})$ de (u_n) qui converge vers ℓ .

Exemples :

- La suite $(-1)^n$ a deux valeurs d'adhérence qui sont ± 1 .
- Les valeurs d'adhérence de la suite $u_n = \cos n$ forment exactement le segment $[-1, 1]$, mais ce n'est pas facile à montrer (il faut utiliser que $\mathbb{Z} + \pi\mathbb{Z}$ est dense dans \mathbb{R}).
- La suite définie par $u_n = 0$ si n pair et $u_n = n$ si n impair est une suite non bornée mais qui a 0 pour valeur d'adhérence.
- Le théorème 2.25 dit donc que toute suite bornée a au moins une valeur d'adhérence.

Le concept de valeur d'adhérence se généralise à des espaces autres que \mathbb{R} (voir chapitres suivants). Il existe par contre un concept généralisant celui de limites qui est fortement lié à l'ordre sur \mathbb{R} .

Définition 2.28

Soit $(u_n)_{n \in \mathbb{N}}$ une suite réelle. Si (u_n) est majorée, on appelle **limite supérieure** ou plus simplement « limsup » la limite

$$\limsup_{n \rightarrow +\infty} u_n := \lim_{n \rightarrow +\infty} \left(\sup_{k \geq n} u_k \right).$$

Si (u_n) est minorée, on appelle **limite inférieure** ou plus simplement « liminf » la limite

$$\liminf_{n \rightarrow +\infty} u_n := \lim_{n \rightarrow +\infty} \left(\inf_{k \geq n} u_k \right).$$

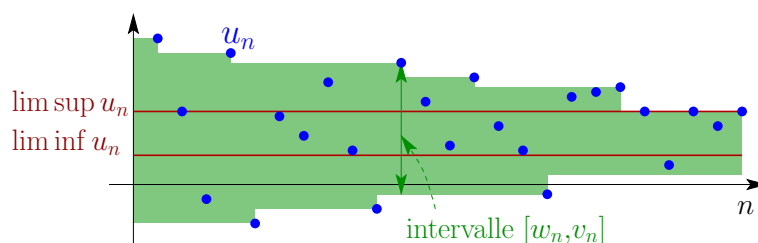


FIGURE 2.2 – L'exemple d'une suite qui ne converge pas mais qui possède une limite supérieure et une limite inférieure. En vert, on voit l'intervalle dans lequel se trouve tous les u_k pour $k \geq n$, qui se réduit petit à petit pour converger vers l'intervalle $[\liminf u_n, \limsup u_n]$.

La définition ci-dessus semble assurer que les limites existent bien. C'est en effet le cas comme le dit la proposition suivante.

Proposition 2.29

Soit $(u_n)_{n \in \mathbb{N}}$ une suite réelle. Si (u_n) est majorée par $M \in \mathbb{R}$, alors $\limsup u_n$ existe et appartient à $[-\infty, M]$. Si (u_n) est minorée par $m \in \mathbb{R}$, alors $\liminf u_n$ existe et appartient à $[m, +\infty]$.

La suite (u_n) est bornée si et seulement si les limsup et liminf existent et sont finies. On a alors toujours $\liminf u_n \leq \limsup u_n$, avec égalité si et seulement si (u_n) converge.

Démonstration : Si (u_n) est majorée par M , l'ensemble $\{u_k, k \geq n\}$ est majoré par M et non vide. On peut donc poser $v_n = \sup_{k \geq n} u_k$ et $v_n \leq M$ par définition de la borne supérieure. Comme $\{u_k, k \geq n+1\} \subset \{u_k, k \geq n\}$, on a $v_{n+1} \leq v_n$. La suite (v_n) étant décroissante, soit elle tend vers $-\infty$, soit elle converge vers une limite finie, forcément inférieure à M puisque $v_n \leq M$ pour tout n .

Le cas de la liminf se traite pareil en posant $w_n = \inf_{k \geq n} u_k$, qui est une suite croissante.

Par définition, on a $w_n \leq u_n \leq v_n$. Si les limsup et liminf existent et sont finies, (w_n) et (v_n) sont deux suites convergentes dans \mathbb{R} . En particulier, elles sont bornées et donc (u_n) est aussi bornée. Réciproquement, si $(u_n) \subset [m, M]$, alors on a $m \leq w_n \leq u_n \leq v_n \leq M$. La suite (v_n) est décroissante minorée et (w_n) est croissante majorée. Donc ces deux suites convergent vers des limites finies. De $w_n \leq u_n \leq v_n$ on obtient que $\liminf u_n \leq \limsup u_n$ et que s'il y a égalité, alors (u_n) converge vers la même limite (théorème des gendarmes).

Il reste à montrer que si (u_n) converge vers $\ell \in \mathbb{R}$, alors $\liminf u_n = \limsup u_n = \ell$. Soit $\varepsilon > 0$, il existe $n_0 \in \mathbb{N}$ tel que si $k \geq n_0$, alors $\ell - \varepsilon \leq u_k \leq \ell + \varepsilon$. Donc pour tout $n \geq n_0$, on a $v_n \geq u_n \geq \ell - \varepsilon$ (la borne sup est un majorant) et aussi $v_n \leq \ell + \varepsilon$ (car $\ell + \varepsilon$ est aussi un majorant et la borne sup est le plus petit de tous). D'où $|\ell - v_n| \leq \varepsilon$. On en déduit que (v_n) tend vers ℓ et donc $\limsup u_n = \ell$. Le cas de la limite inférieure est symétrique. \square

Il est possible que $\limsup u_n = -\infty$, ce qui correspond à $u_n \rightarrow -\infty$. La définition n'est pas alors abusive car elle correspond à une suite réelle (v_n) qui tend vers $-\infty$. Mais dans ce cas, la suite (u_n) n'est pas minorée et on a $\inf_{k \geq n} u_k = -\infty$. La suite (w_n) serait alors toujours égale à $-\infty$ et ce n'est plus une suite réelle. On utilisera alors l'abus de notation $\liminf u_n = -\infty$ même si cela ne correspond pas à la définition de la limite inférieure. De même, on posera $\limsup u_n = +\infty$ pour une suite non majorée.

Exemples :

- Si $u_n = (-1)^n$, on a $\liminf u_n = -1$ et $\limsup u_n = 1$.
- On pose $u_n = (-1)^n + 1/n$. On a $v_{2n-1} = v_{2n} = (-1)^n + 1/n$ et $w_n = -1$. Donc $\liminf u_n = -1$ et $\limsup u_n = 1$.
- On regarde de nouveau la suite définie par $u_n = 0$ si n pair et $u_n = n$ si n impair. On a $\liminf u_n = 0$ et $\limsup u_n = +\infty$.

Proposition 2.30

Soit $(u_n)_{n \in \mathbb{N}}$ une suite bornée. Alors $\limsup u_n$ et $\liminf u_n$ sont respectivement les plus grandes et plus petites valeurs d'adhérence de (u_n) .

Démonstration : Soit ℓ une valeur d'adhérence de (u_n) , c'est-à-dire la limite d'une sous-suite $(u_{\varphi(n)})$. On a $w_{\varphi(n)} \leq u_{\varphi(n)} \leq v_{\varphi(n)}$ et à la limite on trouve $\liminf u_n \leq \ell \leq \limsup u_n$. Il reste à montrer que les limsup et liminf sont des valeurs d'adhérence.

Soit $\ell = \limsup u_n$, qui est la limite de $v_n = \sup_{k \geq n} u_k$. Comme v_n tend vers ℓ en décroissant, à partir d'un rang n_1 , on a $\ell \leq v_n \leq \ell + 1$. Par définition, on a $u_k \leq v_{n_1}$ pour tout $k \geq n_1$. Comme $v_{n_1} - 1$ ne majore pas $\sup_{k \geq n_1} u_k$,

il existe $k_1 \geq n_1$ tel que $v_{n_1} - 1 \leq u_{k_1}$. On pose $\varphi(1) = k_1$ et on note que $\ell - 1 \leq u_{\varphi(1)} \leq \ell + 1$.

On prend maintenant un rang n_2 à partir duquel $\ell \leq v_n \leq \ell + 1/2$. Quitte à le prendre plus grand encore, on peut supposer que $n_2 > \varphi(1)$. Par définition, on a $u_k \leq v_{n_2}$ pour tout $k \geq n_2$. Comme $v_{n_2} - 1/2$ ne majore pas $\sup_{k \geq n_2} u_k$, il existe $k_2 \geq n_2$ tel que $v_{n_2} - 1/2 \leq u_{k_2}$. On pose $\varphi(2) = k_2$ et on note que $\ell - 1/2 \leq u_{\varphi(2)} \leq \ell + 1/2$ et aussi que $\varphi(2) > \varphi(1)$.

On continue ainsi de suite et on construit une suite extraite $(u_{\varphi(n)})$ telle que $\ell - 1/n \leq u_{\varphi(n)} \leq \ell + 1/n$. Ceci montre que ℓ est une valeur d'adhérence de (u_n) .

La preuve pour la liminf est évidemment symétrique. \square

4 Fonctions réelles

Nous allons ici revoir certaines définitions et propriétés concernant les fonctions réelles, c'est-à-dire les fonctions du type

$$f : \mathcal{D}_f \subset \mathbb{R} \longrightarrow \mathbb{R} .$$

Nous allons le faire avec le même esprit que la partie sur les suites : nous n'allons pas tout reprendre mais nous concentrer sur certains aspects qui seront utiles pour les prochains chapitres, soit comme propriété fondamentale, soit comme modèle pour une généralisation. Notons qu'on ne parlera que très peu de dérivation dans ce cours. En effet, la topologie est surtout liée aux concepts de limites et de continuité. L'aspect dérivation et intégration relève plutôt de ce qu'on appelle la *géométrie différentielle* et sera vue au second semestre. Cela n'empêchera évidemment pas d'utiliser les concepts d'intégrale ou de dérivée, mais ils ne seront pas l'objet principal de notre étude. En particulier, ils n'apparaîtront pas dans cette partie.

La définition de la limite dans le cas des fonctions suit exactement les mêmes principes que dans le cas des suites. Il y a beaucoup de cas différents donc plutôt que de les apprendre tous par cœur, l'important est de comprendre comment ils sont construits :

- les voisinages d'un point $x \in \mathbb{R}$ sont du type $]x - \varepsilon, x + \varepsilon[$,
- les voisinages de $+\infty$ sont du type $]M, +\infty[$,
- les voisinages de $-\infty$ sont du type $] - \infty, M[$.

et on peut retenir le principe général :

La définition de

« $f(x)$ tend vers $\ell \in [-\infty, +\infty]$ quand x tend vers $\ell' \in [-\infty, +\infty]$ »
s'écrit

« pour tout voisinage \mathcal{V} de ℓ , il existe un voisinage \mathcal{V}' de ℓ' tel que $f(\mathcal{V}') \subset \mathcal{V}$ »
ce qui peut se comprendre comme

« si x est suffisamment proche de ℓ' , alors $f(x)$ reste aussi proche de ℓ que voulu ».

Nous n'allons pas écrire les 9 cas différents de limites (voire plus si on compte aussi les limites à droite et à gauche). De toute façon, ce concept de limite sera revu dans un cadre bien plus général ensuite. Voyons juste quelques exemples.

Définition 2.31

Soit $f : \mathcal{D}_f \rightarrow \mathbb{R}$ une fonction réelle définie autour de x_0 dans le sens où pour tout $\delta > 0$, $]x_0 - \delta, x_0 + \delta[\cap \mathcal{D}_f$ est non vide. Alors,

- on dit que f tend vers $\ell \in \mathbb{R}$ quand x tend vers x_0 si

$$\forall \varepsilon > 0, \exists \delta > 0, \forall x \in]x_0 - \delta, x_0 + \delta[\cap \mathcal{D}_f, |f(x) - \ell| < \varepsilon.$$

- On dit que f tend vers $+\infty$ quand x tend vers x_0 si

$$\forall M \in \mathbb{R}, \exists \delta > 0, \forall x \in]x_0 - \delta, x_0 + \delta[\cap \mathcal{D}_f, f(x) \geq M.$$

Définition 2.32

Soit $f : \mathcal{D}_f \rightarrow \mathbb{R}$ une fonction réelle définie près de $+\infty$, c'est-à-dire que pour tout $x_0 > 0$, $[x_0, +\infty[\cap \mathcal{D}_f$ est non vide. On dit que f tend vers $-\infty$ quand x tend vers $+\infty$ si

$$\forall M \in \mathbb{R}, \exists x_0 \in \mathbb{R}, \forall x \in [x_0, +\infty[\cap \mathcal{D}_f, f(x) \leq M.$$

Ces définitions sont un peu lourdes pour pouvoir inclure le cas où f n'est pas définie partout, ni même dans un voisinage de $+\infty$. Par exemple, on notera que si f est définie de \mathbb{N} dans \mathbb{R} , alors la notion de limite de f en $+\infty$ retombe bien sur la notion de limite pour la suite $(f(n))_{n \in \mathbb{N}}$. C'est naturel puisqu'une suite n'est rien d'autre qu'une fonction définie sur \mathbb{N} .

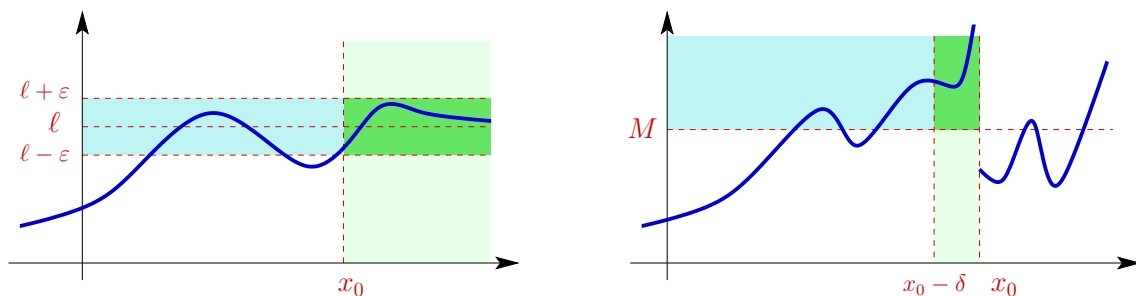


FIGURE 2.3 – Des illustrations d'autres exemples de limites. À gauche un exemple où $f(x) \rightarrow \ell$ quand $x \rightarrow +\infty$: pour tout $\varepsilon > 0$, il existe un point x_0 à partir duquel, pour $x \geq x_0$, $f(x)$ est toujours dans l'intervalle $]l - \varepsilon, l + \varepsilon[$. À droite un exemple où $f(x) \rightarrow +\infty$ quand $x \rightarrow x_0^-$: pour tout $M > 0$, il existe un intervalle $]x_0 - \delta, x_0[$ tel que pour tout $x \in]x_0 - \delta, x_0[$, $f(x)$ est toujours plus grand que M .

À partir de ces définitions, on retrouvera toutes les propriétés des limites qu'on connaît déjà. Comme on fera le cas général plus loin, nous laissons au lecteur la démonstration dans le cas de \mathbb{R} . Voyons un exemple d'énoncé pour s'entraîner. Notons qu'il s'agit d'une propriété fortement liée à \mathbb{R} dans le sens où la preuve utilise la propriété de la borne supérieure.

Proposition 2.33

Soit $f : [0, +\infty[\rightarrow \mathbb{R}$ une fonction décroissante et minorée. Alors $f(x)$ converge vers une limite finie $\ell \in \mathbb{R}$ quand x tend vers $+\infty$.

Démonstration : Comme f est minorée, son image $\{f(x), x \geq 0\}$ est non vide et minorée et admet donc une borne inférieure $\ell := \inf_{x \geq 0} f(x)$. Comme ℓ est un minorant de l'image de f , pour tout $x \geq 0$, $f(x) \geq \ell$. Soit $\varepsilon > 0$. Comme $\ell + \varepsilon$ n'est plus un minorant de l'image de f , il existe $x_0 \geq 0$ tel que $f(x_0) < \ell + \varepsilon$. Mais comme f est décroissante, on a $f(x) < \ell + \varepsilon$ pour tout $x \geq x_0$. Au total, on a $\ell \leq f(x) < \ell + \varepsilon$ pour tout $x \geq x_0$, ce qui montre que f tend vers ℓ quand x tend vers $+\infty$. \square

Exemple :

On considère les étirements $y(t)$ d'un ressort de raideur k qui est soumis à un frottement d'intensité γ . La longueur $y(t)$ est solution de l'équation différentielle $my''(t) + \gamma y'(t) = -ky(t)$. Si on considère l'énergie $E(t) = \frac{1}{2}(m|y'(t)|^2 + k|y(t)|^2)$, on a $E'(t) = my'(t)y''(t) + ky(t)y'(t) = -\gamma|y'(t)|^2 \leq 0$. On trouve que $E(t)$ est décroissante et elle est clairement positive, donc $E(t)$ admet une limite finie positive quand $t \rightarrow +\infty$. Ceci est la première étape pour montrer que l'énergie se dissipe jusqu'à ce que le ressort revienne à l'équilibre.

Pour les fonctions réelles, il y a plusieurs façons équivalentes de définir la continuité. On peut donc en choisir une comme définition de base et les autres comme caractérisations équivalentes.

Définition 2.34

Soit $f : \mathcal{D}_f \rightarrow \mathbb{R}$ une fonction réelle. On dit que f est **continue en** $x_* \in \mathcal{D}_f$ si on a

$$\forall \varepsilon > 0, \exists \delta > 0, \forall x \in \mathcal{D}_f, |x - x_*| < \delta \implies |f(x) - f(x_*)| < \varepsilon.$$

On dit que f est **continue sur un ensemble** $E \subset \mathcal{D}_f$ si f est continue en tout point de E .

On note $\mathcal{C}^0(E, F)$ l'ensemble des fonctions continues sur $E \subset \mathbb{R}$ à valeurs dans $F \subset \mathbb{R}$.

Exemples :

- Là où elles sont définies, les fonctions usuelles sont continues, sauf la partie entière. Donc toute fonction définie par une formule sera continue là où la formule fait sens (sauf dans le cas rare où la partie entière entre en jeu).
- Beaucoup de grandeurs physiques sont en général considérées comme continues, comme la température, la position, la vitesse... Si bien qu'on pourrait penser qu'il n'y a pas à s'embêter avec les fonctions discontinues. Mais dans beaucoup de modèles, il est intéressant de prendre des fonctions discontinues. Par exemple, si on modélise un circuit électronique dont on allume le courant avec un interrupteur au temps $t = 0$, il est plus simple de penser que l'intensité est la fonction I définie par $I(t) = 0$ si $t \leq 0$ et $I(t) = 1$ si $t > 0$ qui est discontinue en 0. En effet, même si la vraie intensité est possiblement continue à cause d'un passage de courant progressif quand l'interrupteur se ferme, cela est trop compliqué à modéliser et il est très probablement non pertinent de s'embêter avec cela. On préférera donc une fonction discontinue. De la même façon, si on regarde une bille qui fait un rebond parfait sur un mur, on supposera le choc élastique. Si la position varie continûment par rapport au temps, sa vitesse sera réfléchiée instantanément lors du rebond et ne sera pas continue. Là encore, si on regarde tout en détail, le changement n'est pas immédiat, mais alors la conservation du moment cinétique obligerait à prendre en compte les déformations du mur, ce qui est trop difficile à faire.

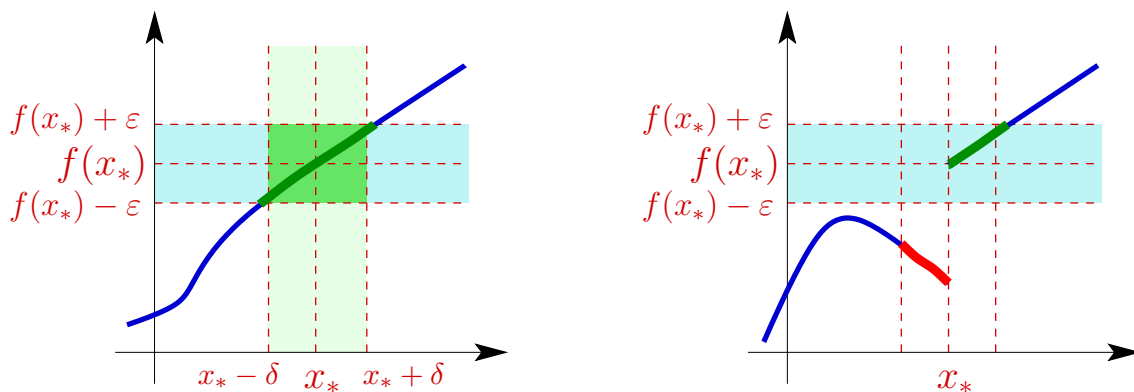


FIGURE 2.4 – À gauche, une fonction continue en x_* : pour tout écart $\varepsilon > 0$, on peut trouver un petit intervalle $]x_* - \delta, x_* + \delta[$ autour de x_* dont l'image reste à distance moins de ε de $f(x_*)$. À droite, la fonction n'est pas continue en x_* : quand $\varepsilon > 0$ est suffisamment petit, il y a toujours des points aussi proches que l'on veut de x_* dont l'image est plus loin que ε de $f(x_*)$.

On pourra utiliser à tout moment les caractérisations équivalentes suivantes.

Théorème 2.35 (critères équivalents pour la continuité)

Soit $f : \mathcal{D}_f \rightarrow \mathbb{R}$ une fonction réelle et $x_* \in \mathcal{D}_f$. Les propositions suivantes sont équivalentes

- i) f est continue en x_* , i.e. pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que si $x \in \mathcal{D}_f$ vérifie $|x - x_*| < \delta$ alors $|f(x) - f(x_*)| < \varepsilon$,
- ii) pour toute suite $(x_n) \subset \mathcal{D}_f$ qui tend vers x_* , $f(x_n)$ tend vers $f(x_*)$,
- iii) les limites à gauche et à droite de f en x_* existent, sont finies et égales à la valeur de f en x_* , c'est-à-dire, si ces limites ont un sens,

$$\lim_{x \rightarrow x_*^-} f(x) = \lim_{x \rightarrow x_*^+} f(x) = f(x_*).$$

Démonstration : Nous allons procéder par une boucle d'implications.

Commençons par supposer que i) est vérifiée. Soit une suite (x_n) tendant vers x_* et soit $\varepsilon > 0$. Par continuité, il existe $\delta > 0$ tel que si $x \in]x_* - \delta, x_* + \delta[$ alors $|f(x) - f(x_*)| < \varepsilon$. Comme (x_n) tend vers x_* , il existe un rang $N \in \mathbb{N}$ tel que $|x_n - x_*| < \delta$ pour $n \geq N$. Donc pour $n \geq N$, $|f(x_n) - f(x_*)| < \varepsilon$ et donc ii) est vérifiée.

Montrons que ii) implique iii) par contraposée. Imaginons que la limite à droite de f en x_* n'existe pas ou bien est différente de $f(x_*)$, c'est-à-dire que la phrase

$$\forall \varepsilon > 0, \exists \delta > 0, \forall x \in]x_*, x_* + \delta[\cap \mathcal{D}_f, |f(x) - f(x_*)| < \varepsilon$$

est fausse. On a donc qu'il existe $\varepsilon > 0$ tel que

$$\forall \delta > 0, \exists x \in]x_*, x_* + \delta[, |f(x) - f(x_*)| \geq \varepsilon \quad (2.1)$$

En appliquant (2.1) à $\delta = 1$, on trouve un point x_1 dans $]x_*, x_* + 1[$ tel que $|f(x_1) - f(x_*)| \geq \varepsilon$. Puis en appliquant (2.1) à $\delta = 1/2$, on trouve un point x_2 dans $]x_*, x_* + 1/2[$ tel que $|f(x_2) - f(x_*)| \geq \varepsilon$ et on recommence ainsi : pour $\delta = 1/n$, on trouve un point x_n dans $]x_*, x_* + 1/n[$ tel que $|f(x_n) - f(x_*)| \geq \varepsilon$. On a ainsi une suite (x_n) qui tend vers x_* et telle que $f(x_n)$ reste à distance plus grande que $\varepsilon > 0$ de $f(x_*)$. Ceci contredit ii). La démonstration est similaire si le problème vient de la limite à gauche.

Supposons finalement que iii) est vraie. Pour tout $\varepsilon > 0$, d'après les définitions des limites à gauche et à droite, il existe δ_+ et $\delta_- > 0$ tels que pour tout $x \in]x_* - \delta_-, x_*[$ et pour tout $x \in]x_*, x_* + \delta_+[$, $|f(x) - f(x_*)| < \varepsilon$. On pose $\delta = \min(\delta_-, \delta_+)$, on a donc $|f(x) - f(x_*)| < \varepsilon$ pour tout $x \in]x_* - \delta, x_* + \delta[$ (le cas $x = x_*$ s'incluant de façon triviale). \square

Le théorème des valeurs intermédiaires correspond à l'idée simple de la continuité comme « le tracé sans lever le crayon ». Dans ce sens, il peut paraître simpliste mais c'est un théorème fondamental qui est plus profond qu'il paraît.

Théorème 2.36 (Théorème des valeurs intermédiaires dit T.V.I.)

Soit $[a,b]$ un intervalle de \mathbb{R} et $f \in \mathcal{C}^0([a,b],\mathbb{R})$ une fonction continue sur $[a,b]$. Soit y une valeur strictement comprise entre les images de a et b , c'est-à-dire que soit $f(a) < y < f(b)$, soit $f(b) < y < f(a)$. Alors, il existe $x \in]a,b[$ tel que $y = f(x)$.

Démonstration : La fonction $g : x \mapsto f(x) - y$ est aussi continue sur $[a,b]$. Le problème revient alors à trouver un point $x \in [a,b]$ où g s'annule en sachant que $g(a)$ et $g(b)$ sont de signes opposés.

On pose $u_0 = a$ et $v_0 = b$. Soit $m_0 = (a + b)/2$ le milieu du segment. Comme $g(a)$ et $g(b)$ sont de signes opposés, on a soit que $g(m_0)$ est du même signe que $g(u_0) = g(a)$ et on pose alors $u_1 = m_0$ et $v_1 = b$, soit $g(m_0)$ est du même signe que $g(b)$ est on pose alors $u_1 = a$ et $v_1 = m_0$. Puis on reprend $m_1 = (u_1 + v_1)/2$ le milieu du nouveau segment. Si $g(m_1) = 0$, on a trouvé notre x tel que $g(x) = 0$ et on peut s'arrêter. Si $g(m_1)$ est du même signe que $g(u_1)$ et on pose alors $u_2 = m_1$ et $v_2 = v_1$, si $g(m_1)$ est du même signe que $g(v_1)$ on pose alors $u_2 = u_1$ et $v_2 = m_1$. . . On continue ainsi en coupant chaque segment en deux et en gardant le morceaux pour lequel les images des bords sont de signes opposés. Soit le processus s'arrête car on a trouvé un point x où g s'annule, soit il se poursuit infiniment. Mais dans ce dernier cas, cela nous construit deux suites (u_n) et (v_n) qui sont par construction adjacentes car (u_n) est croissante, (v_n) est décroissante et $|u_n - v_n|$ est la taille de l'intervalle à l'étape n qui vaut $2^{-n}(b - a)$. Donc (u_n) et (v_n) convergent vers un même limite x . Comme $a = u_0 \leq u_n \leq v_n \leq v_0 = b$, $x \in [a,b]$. Par ailleurs, $g(u_n)$ et $g(v_n)$ sont de signe opposés. Comme g est continue, $g(u_n)$ converge vers $g(x)$ et $g(x)$ est du même signe que $g(u_n)$ au sens large. Mais de même, $g(x)$ est du même signe que $g(v_n)$ au sens large. Le seul nombre que a les deux signes au sens large est $y = 0$. Donc $g(x) = y$ et on a trouvé le point cherché.

Il reste juste à remarquer que x n'est pas seulement dans $[a,b]$ mais en fait dans $]a,b[$. En effet, $g(a)$ et $g(b)$ sont supposés non nuls, donc x ne peut être ni a , ni b . □

Exemples :

- Soit $P : x \mapsto ax^3 + bx^2 + cx + d$ un polynôme de degré 3, c'est-à-dire que $a \neq 0$. C'est une fonction continue sur \mathbb{R} . Supposons $a > 0$. Quand x tend vers $+\infty$, $P(x)$ tend vers $+\infty$ donc pour x assez grand $P(x) > 0$: il existe $b \in \mathbb{R}$ tel que $P(b) > 0$. Quand x tend vers $-\infty$, $P(x)$ tend vers $-\infty$ donc il existe a assez négatif pour que $P(a) < 0$. Le théorème des valeurs intermédiaires nous dit qu'il existe $x \in]a,b[$ tel que $P(x) = 0$. Le cas $a < 0$ est symétrique. On obtient donc le résultat que tout polynôme réel de degré 3 admet au moins une racine réelle. Notons qu'il existe des polynômes de degré 2 sans racines réelles (comme $P(x) = x^2 + 1$).
- Soit $d(t)$ la distance d'un solide à un point de référence, il est naturel de

considérer $d(t)$ comme une fonction continue du temps. Si à $t = 0$ le solide était sur le point de référence et si à $t = T > 0$, il était à distance $d(T) = 100$ m, alors à un moment entre 0 et T , il a été à distance $d(t) = 10$ m.

- Un récipient contient une quantité de liquide que l'on vide progressivement dans un autre récipient qui était vide au départ. Il existe un moment où les deux récipients contiennent exactement le même volume de liquide. En effet, si $V(t)$ est la quantité de liquide dans le récipient d'origine, alors on a au départ $V(0) > 0$ et à la fin $V(T) = 0$. Comme $V(t)$ est naturellement une quantité physique continue, il existe un temps $t \in]0, T[$ tel que $V(t) = V(0)/2$.
- La fonction f définie sur \mathbb{R} par $f(x) = 0$ si $x < 0$ et $f(x) = 1$ si $x \geq 0$ n'est pas continue en $x = 0$. Les valeurs entre $f(-1) = 0$ et $f(1) = 1$ ne sont pas prises par la fonction. Celle-ci ne vérifie pas la propriété des valeurs intermédiaires.
- Une bille de vitesse $V > 0$ subit un choc élastique contre un mur et rebondit en repartant à vitesse $-V < 0$. Pourtant la bille n'a jamais été au repos car son énergie cinétique étant conservée, elle vaut toujours $\frac{1}{2}mV^2 \neq 0$. C'est parce que dans cette modélisation, la vitesse passe brutalement de V à $-V$: elle est discontinue et ne vérifie pas forcément le T.V.I.

Le deuxième résultat théorique important concernant la continuité est lié à ce qu'on appelle la compacité. Il permet non seulement de borner une fonction mais il garantit l'existence d'extrema. On lui associe parfois le nom de Karl Weierstrass (1815-1897, Allemagne).

Théorème 2.37 (théorème des valeurs extrêmes)

Soit $[a, b] \subset \mathbb{R}$ un intervalle fermé et borné et $f \in C^0([a, b], \mathbb{R})$ une fonction réelle continue sur $[a, b]$. Alors f est bornée et atteint ses bornes sur $[a, b]$. Autrement dit, il existe x_{\max} et x_{\min} tels que

$$f(x_{\max}) = \max_{x \in [a, b]} f(x) = \sup_{x \in [a, b]} f(x) \quad \text{et} \quad f(x_{\min}) = \min_{x \in [a, b]} f(x) = \inf_{x \in [a, b]} f(x).$$

Démonstration : On va montrer que f est majorée sur $[a, b]$ et atteint son maximum. Le cas du minimum est symétrique.

Supposons que f ne soit pas majorée sur $[a, b]$, alors par définition, pour tout $n \in \mathbb{N}$, il existe un point $x_n \in [a, b]$ tel que $f(x_n) \geq n$ et en particulier $f(x_n) \rightarrow +\infty$. D'après le théorème de Bolzano-Weierstrass (corollaire 2.26), on peut extraire de (x_n) une sous-suite $(x_{\varphi(n)})$ qui converge vers une limite $\ell \in [a, b]$. On a que $f(x_{\varphi(n)})$ tend vers $+\infty$ car c'est une sous-suite de $f(x_n)$ mais aussi que $f(x_{\varphi(n)})$ tend vers $f(\ell)$ par continuité de f . Comme $f(\ell)$ est un nombre fini, c'est contradictoire et donc f est majorée sur $[a, b]$.

On pose $M = \sup_{x \in [a, b]} f(x)$. Pour tout $n \in \mathbb{N}$, $M - 2^{-n}$ n'est plus un majorant et il existe donc $x_n \in [a, b]$ tel que $M - 2^{-n} < f(x_n) \leq M$. Donc $f(x_n)$ tend vers M . Mais comme précédemment, on peut extraire de (x_n) une sous-suite

$(x_{\varphi(n)})$ qui converge vers une limite $x_{\max} \in [a,b]$. Et par continuité $f(x_{\varphi(n)})$ tend vers $f(x_{\max})$. Donc $f(x_{\max}) = \sup_{x \in [a,b]} f(x)$ est le maximum cherché. \square



Comme on le voit dans les exemples ci-dessous, il est important que f soit continue mais aussi que $[a,b] \subset \mathbb{R}$ soit un intervalle fermé et borné, c'est-à-dire qu'il inclut ses bornes a et b qui sont des réels (finis). On fera plus tard la généralisation de ce résultat. Le point important est que $[a,b]$ est un *compact* qui vérifie le corollaire 2.26.

Exemples :

- Soit $d(t)$ la distance d'un solide à un point de référence, il est naturel de considérer $d(t)$ comme une fonction continue du temps. Pendant un intervalle de temps $[0,T]$, le solide s'est éloigné au maximum d'une distance D et il existe un temps $t_0 \in [0,T]$ où il était pile à distance D .
- La fonction $f : x \mapsto x$ est continue sur $[0, +\infty[$ mais n'est pas majorée dessus. On ne peut pas appliquer le théorème précédent car $[0, +\infty[$ n'est pas un intervalle borné.
- La fonction $f : x \mapsto 1/x$ est continue sur $]0,1]$ mais n'est pas majorée dessus. On ne peut pas appliquer le théorème précédent car $]0,1]$ n'est pas un intervalle fermé.
- La fonction f définie sur $[0,1]$ par $f(0) = 0$ et $f(x) = 1/x$ si $x \in]0,1]$ n'est pas majorée sur $[0,1]$. Même si $[0,1]$ est fermé et borné, on ne peut pas appliquer le théorème précédent car f n'est pas continue.

En rassemblant les deux énoncés de cette partie, on obtient ce qu'on pourra appeler dans pas très longtemps « l'image d'un connexe compact par une fonction continue est un connexe compact ».

Corollaire 2.38

L'image d'un intervalle $[a,b] \subset \mathbb{R}$ par une fonction réelle continue est un intervalle $[\alpha,\beta] \subset \mathbb{R}$.

Démonstration : Le théorème 2.37 nous dit que l'image par f continue d'un intervalle $[a,b]$ est bornée et que les bornes sont atteintes. Donc $\alpha = \min_{[a,b]} f$ et $\beta = \max_{[a,b]} f$ sont dans l'image de f , atteints aux points x_{\min} et x_{\max} respectivement. Par définition de ces extrema, l'image de f est incluse dans $[\alpha,\beta]$. Mais le théorème des valeurs intermédiaires appliqué sur l'intervalle $[x_{\min},x_{\max}]$ (ou $[x_{\max},x_{\min}]$) nous dit que toutes les valeurs de $[\alpha,\beta]$ sont atteintes. \square

Finissons ce chapitre par introduire l'uniforme continuité dans le cas de \mathbb{R} .

Définition 2.39

Une fonction $f : \mathcal{D}_f \rightarrow \mathbb{R}$ est dite **uniformément continue** si

$$\forall \varepsilon > 0, \exists \delta > 0, \forall x, y \in \mathcal{D}_f, |x - y| < \delta \implies |f(x) - f(y)| < \varepsilon.$$



Si on écrit la continuité de f sous la même forme, on écrirait

$$\forall \varepsilon > 0, \forall x \in \mathcal{D}_f, \exists \delta > 0, \forall y \in \mathcal{D}_f, |x - y| < \delta \implies |f(x) - f(y)| < \varepsilon.$$

On voit alors que l'uniforme continuité est strictement plus forte que la continuité : si la fonction est continue, δ est choisi après x et peut dépendre de x ; alors que si la fonction est uniformément continue, δ est choisi avant x et est donc le même pour tout x , i.e. il est uniforme en x .

Le résultat le plus important concernant l'uniforme continuité dans \mathbb{R} est le suivant. Son nom vient du mathématicien Eduard Heine (1821-1881, Allemagne). De nouveau, la notion importante cachée derrière est celle de *compacité* et c'est pour cela qu'il est important de considérer un intervalle $[a, b]$ fermé et borné.

Théorème 2.40 (théorème de Heine)

Soit $a < b$ et $f : [a, b] \rightarrow \mathbb{R}$ une fonction continue. Alors f est uniformément continue sur $[a, b]$.

Démonstration : Nous raisonnons par l'absurde. Si f n'est pas uniformément continue, c'est qu'il existe $\varepsilon > 0$ tel que pour tout $\delta = 1/n$, il existe x_n et y_n dans $[a, b]$ tels que $|x_n - y_n| < 1/n$ et $|f(x_n) - f(y_n)| \geq \varepsilon$. Par le théorème de Bolzano-Weierstrass (corollaire 2.26), on peut extraire une sous-suite $(x_{\varphi(n)})$ telle que $x_{\varphi(n)} \rightarrow x_* \in [a, b]$. Comme $|x_n - y_n| < 1/n$, on doit aussi avoir $y_{\varphi(n)} \rightarrow x_*$. Par continuité, il faut que $f(x_{\varphi(n)})$ et $f(y_{\varphi(n)})$ aient la même limite $f(x_*)$. C'est-à-dire qu'on devrait avoir $|f(x_{\varphi(n)}) - f(y_{\varphi(n)})| \rightarrow 0$, ce que contredirait l'inégalité $|f(x_n) - f(y_n)| \geq \varepsilon$. \square

Exemples :

- Le résultat précédent nous donne beaucoup d'exemples : toutes les fonctions usuelles sur un intervalle $[a, b]$ sont uniformément continue.
- La fonction $x \in [0, +\infty[\mapsto 2x$ est uniformément continue. En effet, pour tout $\varepsilon > 0$, on peut prendre $\delta = \varepsilon/2$.
- La fonction $x \in [0, +\infty[\mapsto x^2$ n'est uniformément continue. En effet, pour $\varepsilon = 1$ et pour tout $\delta > 0$, on peut prendre $x = \frac{2}{\delta}$ et $y = x + \frac{\delta}{2}$. On a alors

$$|x^2 - y^2| = \frac{\delta}{4}(4x + \delta) > \delta x = 2 > \varepsilon.$$

Ceci contredit la définition de l'uniforme continuité. Visuellement, on voit que la distance séparant deux images de points proches x et $x + \delta$ est de plus grande quand x va vers $+\infty$. C'est donc que l'uniforme continuité n'est pas possible.

- La fonction $x \in]0,1] \mapsto \frac{1}{x}$ n'est uniformément continue. En effet, pour $\varepsilon = 1/2$ et pour tout $\delta > 0$, on pose $\alpha = \min(\delta, 1/2)$ et on peut prendre $x = \alpha$ et $y = 2\alpha$. On a alors

$$\left| \frac{1}{x} - \frac{1}{y} \right| = \frac{1}{2\alpha} \geq 1 > \varepsilon.$$

Ceci contredit la définition de l'uniforme continuité. Visuellement, on voit que la distance séparant deux images de points proches x et $x + \delta$ est de plus grande quand x tend vers 0. C'est donc que l'uniforme continuité n'est pas possible.

L'uniforme continuité peut paraître une définition technique, mais elle est centrale pour beaucoup de preuves de propriétés qui sont fausses si la fonction est seulement continue.

Exemple :

Si $f :]0,1] \rightarrow \mathbb{R}$ est définie par $f(x) = \cos(1/x)$, il s'agit d'une fonction continue et bornée qui n'a pas de limite quand $x \rightarrow 0$. En effet, $f(1/2n\pi) \rightarrow 1$ et $f(1/(2n+1)\pi) \rightarrow -1 \neq 1$. Notons qu'elle n'est pas uniformément continue (ce qui pourra se démontrer par la contraposée de ce qui suit).

Mais si on prend $f :]0,1] \rightarrow \mathbb{R}$ quelconque qui est uniformément continue sur $]0,1]$ alors on va montrer que f admet une limite quand $x \rightarrow 0$. On pose $u_n = f(1/n)$ et on prétend que (u_n) est de Cauchy. En effet, pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que $|f(x) - f(y)| < \varepsilon$ si $|x - y| < \delta$. Donc si p et q sont assez grands, par exemple si $p, q \geq 1/\delta$, alors $|1/p - 1/q| < \delta$ et donc $|u_p - u_q| = |f(1/p) - f(1/q)| < \varepsilon$. Par le théorème 2.14, la suite (u_n) converge vers une limite $\ell \in \mathbb{R}$. Montrons maintenant que cette limite est bien celle de toute la fonction (nous n'avons que le cas $f(1/n)$ pour le moment). Soit $\varepsilon > 0$. On applique l'uniforme continuité : il existe $\delta > 0$ tel que si $|x - y| < \delta$ alors $|f(x) - f(y)| < \varepsilon/2$. On a aussi un rang n_0 tel que si $n \geq n_0$, $|\ell - f(1/n)| < \varepsilon/2$. Soit $x \in]0, \delta[$ et soit $n \geq \max(n_0, 1/\delta + 1)$. On a alors $1/n \in]0, \delta[$ et donc $|x - 1/n| < \delta$. On obtient au final

$$|\ell - f(x)| \leq |\ell - f(1/n)| + |f(1/n) - f(x)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

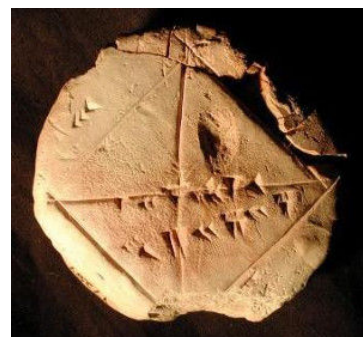
Ceci montre bien que $f(x) \rightarrow \ell$ quand $x \rightarrow 0$.

5 Quelques applications

5.1 La méthode de Héron

Il y a plus de 3.000 ans, les savants de l'Antiquité mésopotamienne et égyptienne savaient extraire des racines carrées. On trouve ainsi des tablettes d'argile avec la valeur de la diagonale d'un carré montrant qu'ils savaient calculer $\sqrt{2}$ à la précision du

millionième. Les scribes n'ont pas expliqué leur méthode, mais on pense qu'il s'agit du même algorithme que celui utilisé par les grecs plusieurs siècles plus tard. Il est appelé *méthode de Héron* ou *méthode babylonienne*. L'idée est géométriquement simple. Supposons que l'on veuille calculer \sqrt{a} , cela revient à trouver quel est le côté d'un carré de surface a . On part d'un rectangle de surface a , par exemple de côtés $u_0 = a$ et $v_0 = 1$ (en pratique, il est plus rapide de partir d'une meilleure approximation du bon carré, par exemple $u_0 = 4$ et $v_0 = 2,5$ pour calculer $\sqrt{10}$). Ce premier rectangle a la bonne surface mais n'est pas un carré. Pour le rendre « plus carré », on prend comme nouveau côté la moyenne des côtés en posant $u_1 = (u_0 + v_0)/2$ et donc $v_1 = a/u_1$. Notre nouveau rectangle est encore d'aire a mais il est proche du carré. On itère ainsi le procédé et on se rapproche de plus en plus de la racine carrée cherchée.



Un calcul de $\sqrt{2}$ datant de plus de 3.500 ans

Mathématiquement parlant, comme $v_n = a/u_n$, cela revient à partir de $u_0 = a$ et itérer la récurrence

$$u_{n+1} = \frac{1}{2} \left(u_n + \frac{a}{u_n} \right). \quad (2.2)$$

L'étude de la fonction $f : x \mapsto (x + a/x)/2$ montre qu'elle est décroissante sur $]0, \sqrt{a}]$ et croissante sur $[\sqrt{a}, +\infty[$ et que $f(\sqrt{a}) = \sqrt{a}$. En particulier, f envoie $[\sqrt{a}, +\infty[$ sur lui-même et pour tout n , $u_n \geq \sqrt{a}$. On a alors que $v_n = a/u_n \leq \sqrt{a}$ et donc que $u_{n+1} = (u_n + v_n)/2 \leq (u_n + u_n)/2 = u_n$. La suite (u_n) est donc décroissante et minorée par \sqrt{a} et elle converge vers un réel $\ell \geq \sqrt{a}$. En passant à la limite dans (2.2), on obtient que $\ell = (\ell + a/\ell)/2$ ce qui donne $\ell^2 = a$. Comme ℓ est positive, on a donc $\ell = \sqrt{a}$.

Pour obtenir une bonne approximation de \sqrt{a} , il suffit donc de partir d'une valeur raisonnable et d'itérer la formule (2.2) qui n'est composée que d'opérations élémentaires. On peut prouver que cette méthode est très efficace car elle converge quadratiquement : à chaque étape, on double le nombre de chiffres exacts de notre approximation.

5.2 La méthode de dichotomie

Si on regarde bien la preuve du théorème des valeurs intermédiaires, elle fournit une méthode simple, constructive et explicite pour chercher les zéros d'une fonction continue, c'est-à-dire pour trouver une solution à une équation $f(x) = 0$. Supposons que f soit continue sur un intervalle $[a, b]$ et que l'on sache que $f(a)$ et $f(b)$ sont de signes opposés. Alors on coupe l'intervalle $[a, b]$ en deux et on garde l'intervalle où f admet des valeurs de signes opposés au bord. Puis on recoupe cette intervalle en deux etc. On sait par le T.V.I. qu'il existe bien une solution de $f(x) = 0$ dans chacun des intervalles et plus on avance, plus ces intervalles sont petits et plus on obtient une bonne approximation de cette solution. Attention toutefois que cette méthode ne garantit pas de trouver toutes les solutions même si elle permet d'en trouver au moins une. L'algorithme est schématiquement :

Tant que $(a-b)/2 > \text{précision}$ faire boucle

 pose $m=(a+b)/2$

 si $f(a)f(m)<0$

 alors pose $b=m$

 sinon pose $a=m$

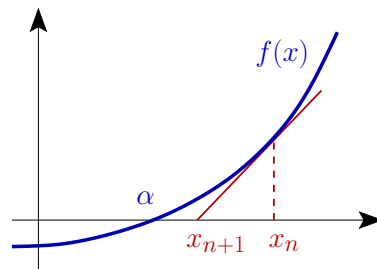
 fin si

fin boucle

Écrit "la solution vaut" $(a+b)/2$ "avec la précision \pm " précision

5.3 L'algorithme de Newton

La méthode de Newton est utilisée pour approcher les zéros de fonctions. Considérons une fonction f dont on veut trouver un zéro α , c'est-à-dire une solution de l'équation $f(\alpha) = 0$. L'idée géométrique est de partir d'une valeur assez proche x_0 , puis d'approcher la fonction f par sa tangente en x_0 et de prendre comme nouvelle position x_1 le zéro de cette tangente. Puis on recommence le procédé et on espère que la suite (x_n) tende bien vers α . Ce n'est pas toujours le cas et l'algorithme peut échouer, ne serait-ce que si la tangente que l'on regarde est horizontale et donc n'a pas de zéro. La méthode a été introduite par Isaac Newton en 1669 mais celui-ci ne l'a utilisée que pour les polynômes. Il a fallu un peu de temps avant que l'on comprenne la généralité de l'algorithme et qu'il prenne sa forme moderne.



Concrètement, si f est dérivable, sa tangente au point x_n est la droite d'équation $y = f'(x_n)(x - x_n) + f(x_n)$. Comme x_{n+1} est le zéro de cette tangente, on a

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}. \quad (2.3)$$

Supposons que sur l'intervalle $[\alpha, x_0]$, la fonction f soit croissante. Elle est donc positive et f' est positive aussi. On obtient alors que $x_{n+1} \leq x_n$. Supposons en outre que f soit convexe sur $[\alpha, x_0]$, elle est alors au-dessus de ses tangentes et $\alpha \leq x_{n+1}$ (voir la figure ci-dessus). Sous ces hypothèses, la suite (x_n) est donc décroissante et minorée par α et elle converge vers une limite $\ell \in [\alpha, x_0]$. En passant à la limite dans (2.3), on a que $\ell = \ell - f(\ell)/f'(\ell)$ et donc que $f(\ell) = 0$. Comme α est le seul zéro de f dans l'intervalle considéré, $\ell = \alpha$. Dans ce cas, la suite (x_n) partant de x_0 et définie par la récurrence (2.3) converge bien vers un zéro f . Dans d'autre cas, la convergence n'a pas forcément lieu mais on peut montrer qu'elle est très rapide si elle a lieu.

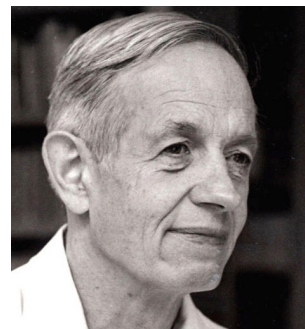
Une application simple de l'étude précédente est le cas du plus grand zéro des polynômes scindés. Prenons par exemple $f(x) = x^2 - a$ avec $a \in \mathbb{R}$. La fonction f est croissante pour $x \geq 0$ et convexe sur \mathbb{R} . Si on part d'un point $x_0 \geq \sqrt{a}$ et qu'on itère la méthode de Newton, la suite (x_n) va donc converger vers \sqrt{a} . Dans le cas $f(x) = x^2 - a$, (2.3) devient

$$x_{n+1} = x_n - \frac{x_n^2 - a}{2x_n} = x_n - \frac{x_n}{2} + \frac{a}{2x_n} = \frac{1}{2} \left(x_n + \frac{a}{x_n} \right).$$

On retrouve la méthode de Héron !

5.4 Un théorème de point fixe

Les points fixes d'une fonction f , c'est-à-dire les solutions de l'équation $f(x) = x$, jouent un rôle important dans beaucoup de théories et applications. Par exemple, leur interprétation en tant qu'équilibres d'une dynamique $u_{n+1} = f(u_n)$ se retrouve en théorie des jeux et donc en économie. En se limitant aux fonctions à une seule variable réelle, cela revient souvent à prendre des modèles simplistes, mais essayons d'en comprendre le principe général. Imaginons un agent économique qui peut produire une quantité x de biens et en tirer un profit $h(x,y)$ qui dépend de sa production x et de l'état du marché y . Il va tenter de maximiser ses gains et donc produire la quantité $x_* = g(y)$ telle que $h(x_*,y) = \max_x h(x,y)$. Mais l'état du marché y dépend aussi de la production x de l'agent (une surproduction peut baisser les prix etc.). Si l'agent produit une quantité x , que le marché est dans $y(x)$, que son gain est $h(x,y(x))$ mais que produire la quantité $x_* = g(y(x))$ lui apporterait un meilleur gain, il sera insatisfait et changera sa production en x_* . Mais cela va changer l'état $y(x)$ du marché en $y(x_*)$ et la production optimale ne sera pas forcément toujours x_* etc. L'idée est que si le marché est stable, c'est qu'on a fini dans un état tel que $x_* = g(y(x_*))$. On est amené à chercher un point fixe de la fonction $f : x_* \mapsto g(y(x_*))$. C'est un *équilibre de Nash* du nom du mathématicien John Nash, lauréat du prix dit Nobel d'économie et du prix Abel.



John Nash
1928-2015
États-Unis

Théorème 2.41

Soit $[a,b] \subset \mathbb{R}$ et $f \in \mathcal{C}^0([a,b],[a,b])$ une fonction continue envoyant $[a,b]$ sur lui-même. Alors f a au moins un point fixe, c'est-à-dire qu'il existe $x_* \in [a,b]$ tel que $f(x_*) = x_*$.

Démonstration : Si $f(a) = a$, c'est gagné, donc supposons que l'on a $f(a) \neq a$, ce qui implique $f(a) > a$ car $f(a) \in [a,b]$. De même, si $f(b) = b$, c'est gagné, donc supposons que l'on a $f(b) \neq b$, ce qui implique $f(b) < b$. On considère la fonction $g : x \mapsto f(x) - x$. La fonction g est continue sur $[a,b]$ comme somme de fonctions continues. Par ailleurs, les hypothèses ci-dessus nous donnent que $g(a) = f(a) - a > 0$ et $g(b) = f(b) - b < 0$. En appliquant le théorème des valeurs intermédiaires, on voit qu'il existe $x_* \in]a,b[$ tel que $g(x_*) = 0$. Mais cela veut dire que $f(x_*) - x_* = 0$ et donc x_* est un point fixe de f . \square

5.5 Le minimum d'un puits de potentiel

On considère un potentiel continu $V : \mathbb{R} \rightarrow \mathbb{R}$. On suppose que $\lim_{x \rightarrow \pm\infty} V(x) = +\infty$. Cela veut dire que ce potentiel est un *potentiel puits* qui a tendance à ramener notre système vers une zone bornée. Par les résultats ci-dessus, nous pouvons

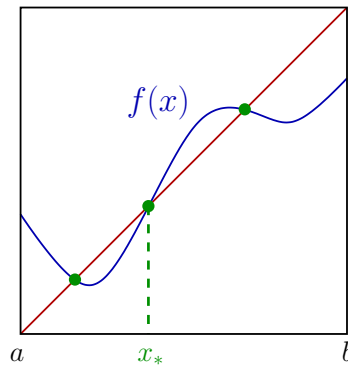


FIGURE 2.5 – Une illustration du théorème 3.41 : le graphe de la fonction continue f doit forcément intersecter la droite $y = x$ et donc f a forcément au moins un point fixe (ici elle en a trois).

montrer que V admet un minimum global, c'est-à-dire un point x_* où l'énergie potentielle est la plus petite possible et donc pour lequel, un système dans cet état x_* resterait stable proche de cette énergie minimale. En effet, soit $M = V(0) + 1$. Comme $\lim_{x \rightarrow -\infty} V(x) = +\infty$, il existe x_- tel que si $x \leq x_-$ alors $V(x) \geq M$. De même, comme $\lim_{x \rightarrow +\infty} V(x) = +\infty$, il existe x_+ tel que si $x \geq x_+$ alors $V(x) \geq M$. Donc pour tout x en dehors de $[x_-, x_+]$, $V(x)$ est minoré par $V(0) + 1$, et en particulier $0 \in [x_-, x_+]$ puisque $V(0) < V(0) + 1$. Par ailleurs, comme V est continue sur l'intervalle fermé borné $[x_-, x_+]$, V y est minoré et atteint son minimum en un point x_* . On a donc que pour tout $x \in [x_-, x_+]$, $V(x) \geq V(x_*)$. Mais comme $0 \in [x_-, x_+]$, pour tout $x \notin [x_-, x_+]$, $V(x) \geq V(0) + 1 > V(0) \geq V(x_*)$. Au total, on a bien que $V(x) \geq V(x_*)$ pour tout $x \in \mathbb{R}$ et V admet un minimum global en x_* .

5.6 Le théorème de Weierstrass

Cette dernière application concerne en fait un théorème important de l'analyse.

Théorème 2.42 (théorème de Weierstrass et polynômes de Bernstein)

Soient $a < b$. Pour toute fonction continue $f \in \mathcal{C}^0([a, b], \mathbb{R})$, il existe une suite de polynômes $(P_n) \subset \mathbb{R}[X]$ qui converge uniformément vers f sur $[a, b]$, c'est-à-dire que

$$\sup_{x \in [a, b]} |f(x) - P_n(x)| \xrightarrow{n \rightarrow +\infty} 0. \quad (2.4)$$

Plus précisément, si $[a, b] = [0, 1]$, la suite

$$P_n = \sum_{k=0}^n f\left(\frac{k}{n}\right) \binom{n}{k} X^k (1 - X)^{n-k} \quad (2.5)$$

satisfait la convergence uniforme demandée.



Karl Weierstrass
1815-1897
Allemagne



Sergueï Bernstein
1880-1968
Russie

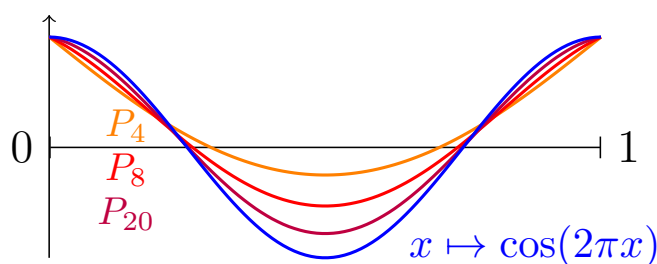
Pour comprendre l'importance du théorème de Weierstrass, il faut se rappeler que les polynômes et leur quotient sont les seules fonctions que l'on peut calculer de façon exacte. Des opérations comme la racine carré ou le log ne sont en fait implémentées que de façon approchée. Par ailleurs, il est facile de coder et stocker un polynôme car il s'agit juste d'une suite de coefficients. Le théorème de Weierstrass nous dit qu'avec ces polynômes, on peut faire toutes les formes possibles (au moins de façon approchée). Ainsi, on peut utiliser les polynômes pour coder tous les dessins sur un ordinateur et la police de caractères de ce poly est elle-même codée à l'aide de polynômes.

Il existe plusieurs suites (P_n) satisfaisant (2.4). La suite (2.5) est due à Bernstein. La forme (2.5) est valable pour le segment $[0,1]$ mais la transformation $x \mapsto (x - a)/(b - a)$ permet de passer de $[a,b]$ à $[0,1]$ et donc de transposer cette formule à tout segment $[a,b]$.

Les polynômes de Taylor nous fournissent déjà des approximations de fonctions par développement limité. Mais il ne s'agit que d'une approximation proche d'un point et qui peut rester très mauvaise ailleurs, même quand on va de plus en plus loin dans le développement. L'interpolation de Lagrange fournit une interpolation polynomiale qui passe par un certain nombre de points. Mais là encore, l'approximation peut rester mauvaise ailleurs, même si on augmente le nombre de points.

Exemples :

- Les polynômes de Bernstein (2.5) fournissent une suite explicite. On peut faire le test ci-dessous avec la fonction $x \in [0,1] \mapsto \sin(2\pi x)$. Les premières expressions sont simples, par exemples $P_4 = -4X^4 + 8X^3 - 4X + 1$ mais on obtient rapidement des polynômes peu pratiques mêmes si explicites. Par ailleurs, la convergence semble très lente.



Il se trouve que les polynômes de Bernstein ne sont pas les meilleurs pour faire des bonnes approximations, même s'ils suffisent à obtenir le théorème comme nous allons le voir dans la preuve. En pratique, on utilise d'autres approximations polynomiales plus adéquates.

- On rappelle que la limite uniforme d'une suite de fonctions continues est continue. Le théorème de Weierstrass échoue donc si f n'est pas continue.
- Il est important de considérer un intervalle *compact*, c'est-à-dire $[a,b]$ fermé et borné. Ainsi, on ne peut pas approcher uniformément toutes les fonctions par des polynômes sur \mathbb{R} . Par exemple, la fonction $x \in \mathbb{R} \mapsto \sin x$ est bornée mais pas constante. Elle n'est pas bien approchée par les constantes mais dès qu'on prend un polynôme non constant, celui-ci a des limites infinies en $\pm\infty$ et donc l'erreur dans (2.4) n'est même pas majorée sur \mathbb{R} et $\sup_x |f(x) - P_n(x)| = +\infty$. De même, si on considère $x \in [0, +\infty[\mapsto e^x$, la croissance en $+\infty$ de l'exponentielle est trop forte pour pouvoir être approchée par des polynômes. Enfin, le même problème arrive si on considère un intervalle borné mais non fermé. Ainsi $x \in]0,1] \mapsto 1/x$ n'est pas bornée alors que tout polynôme est borné sur $]0,1]$. Donc la distance entre $1/x$ et un polynôme ne pourra pas être petite uniformément sur tout $]0,1]$.

Bernstein s'est appuyé sur les probabilités pour trouver la formule (2.5) qui a fourni une nouvelle preuve du théorème de Weierstrass. Imaginons une expérience aléatoire qui a une chance $p \in [0,1]$ de réussir. La variable aléatoire X comptant le nombre de succès de n expériences indépendantes suit la loi binomiale de paramètre p définie par

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} .$$

Il est connu (cf n'importe quel cours de probabilités de niveau L2) que

$$E(X) = np \quad \text{et} \quad \text{Var}(X) = np(1-p) .$$

On en déduit les résultats suivants.

Lemme 2.43

Pour tout $x \in \mathbb{R}$ et $n \in \mathbb{N}$,

$$\sum_{k=0}^n \binom{n}{k} x^k (1-x)^{n-k} = 1 .$$

Démonstration : Il s'agit simplement de la formule du binôme $\sum_{k=0}^n \binom{n}{k} x^k (1-x)^{n-k} = (x+(1-x))^n = 1$. Mais on peut aussi la voir comme le fait que la loi binomiale est une loi de probabilité et donc $\sum_{k=0}^n P(X = k) = 1$ pour $p = x \in [0,1]$. Comme les deux membres de l'égalité sont des polynômes en x , s'ils sont égaux sur $[0,1]$, ils le sont sur tout \mathbb{R} . \square

Lemme 2.44

Pour tout $x \in \mathbb{R}$ et $n \in \mathbb{N}$,

$$\sum_{k=0}^n \left(x - \frac{k}{n}\right)^2 \binom{n}{k} x^k (1-x)^{n-k} = \frac{x(1-x)}{n}.$$

Démonstration : Comme ci-dessus, il suffit de montrer l'égalité pour tout $x = p \in [0,1]$. Notre connaissance de la variance de la loi binomiale nous dit que

$$\begin{aligned} np(1-p) &= \text{Var}(X) = \sum_{k=0}^n P(X=k)(k - E(X))^2 \\ &= \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} (k - np)^2 \end{aligned}$$

On obtient la formule cherchée en divisant par n^2 des deux côtés. \square

Nous allons utiliser ces lemmes dans une estimation astucieuse grâce à la continuité uniforme.

Démonstration du théorème 2.42 : Comme dit plus haut, la transformation $x \mapsto (x-a)/(b-a)$ permet de passer de $[a,b]$ à $[0,1]$ et elle transforme les polynômes en polynômes. Il suffit donc de montrer que la suite P_n donnée par (2.5) vérifie (2.4) dans le segment $[a,b] = [0,1]$.

Soit $f \in \mathcal{C}^0([0,1], \mathbb{R})$ et soit $\varepsilon > 0$, on doit montrer que pour n assez grand $\sup_x |f(x) - P_n(x)| < \varepsilon$. On commence par utiliser le lemme 2.43 pour écrire

$$f(x) - \sum_{k=0}^n f\left(\frac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k} = \sum_{k=0}^n \left(f(x) - f\left(\frac{k}{n}\right)\right) \binom{n}{k} x^k (1-x)^{n-k}.$$

Nous allons maintenant utiliser le théorème de Heine (théorème 2.40) : la fonction f est uniformément continue. Donc il existe $\delta > 0$ tel que pour tout x, y avec $|x - y| < \delta$, on a $|f(x) - f(y)| < \varepsilon/2$. Par ailleurs, le théorème 2.37 nous assure qu'il existe une borne $M > 0$ telle que $\sup_{x \in [0,1]} |f(x)| \leq M$.

Pour tout $n \in \mathbb{N}$ et $x \in [0,1]$, on introduit l'ensemble $K_{n,x,\varepsilon}$ des indices $k \in \llbracket 0,1 \rrbracket$ tels que $|x - k/n| < \delta$. On découpe alors la somme en deux en utilisant que $x \in [0,1]$

$$\begin{aligned} \left| f(x) - \sum_{k=0}^n f\left(\frac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k} \right| &\leq \sum_{k=0}^n \left| f(x) - f\left(\frac{k}{n}\right) \right| \binom{n}{k} x^k (1-x)^{n-k} \\ &\leq \sum_{k \in K_{n,x,\varepsilon}} \left| f(x) - f\left(\frac{k}{n}\right) \right| \binom{n}{k} x^k (1-x)^{n-k} \\ &\quad + \sum_{k \notin K_{n,x,\varepsilon}} \left| f(x) - f\left(\frac{k}{n}\right) \right| \binom{n}{k} x^k (1-x)^{n-k} \end{aligned}$$

Pour les indices dans $K_{n,x,\varepsilon}$, on utilise simplement le lemme 2.43 et la définition de cet ensemble : l'erreur entre $f(x)$ et $f(k/n)$ y est petite et donc (toujours en utilisant la positivité de x et $1-x$)

$$\sum_{k \in K_{n,x,\varepsilon}} \left| f(x) - f\left(\frac{k}{n}\right) \right| \binom{n}{k} x^k (1-x)^{n-k} \leq \frac{\varepsilon}{2} \sum_{k \in K_{n,x,\varepsilon}} \binom{n}{k} x^k (1-x)^{n-k} \leq \frac{\varepsilon}{2}$$

L'autre partie suit l'idée de l'inégalité de Bienaymé-Tchebychev. On se débarasse grossièrement de f par le théorème 2.37 qui nous assure qu'il existe une borne $M > 0$ telle que $\sup_{x \in [0,1]} |f(x)| \leq M$. Puis si k/n est à distance plus que δ de x , alors $|x - k/n|^2 > \delta^2$. Alors le lemme 2.44 permet d'obtenir que

$$\begin{aligned} \sum_{k \notin K_{n,x,\varepsilon}} \left| f(x) - f\left(\frac{k}{n}\right) \right| \binom{n}{k} x^k (1-x)^{n-k} &\leq 2M \sum_{k \notin K_{n,x,\varepsilon}} \binom{n}{k} x^k (1-x)^{n-k} \\ &\leq \frac{2M}{\delta^2} \sum_{k \notin K_{n,x,\varepsilon}} |x - k/n|^2 \binom{n}{k} x^k (1-x)^{n-k} \\ &\leq \frac{2M}{\delta^2} \sum_{k=0}^n |x - k/n|^2 \binom{n}{k} x^k (1-x)^{n-k} \\ &\leq \frac{2M}{n\delta^2} x(1-x) \leq \frac{M}{2n\delta^2} \end{aligned}$$

En regroupant les deux estimations précédentes, on obtient pour tout $\varepsilon > 0$, si $\delta > 0$ est le module de continuité uniforme défini plus haut et M une borne uniforme sur $|f|$, alors pour tout $x \in [0,1]$,

$$\left| f(x) - \sum_{k=0}^n f\left(\frac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k} \right| \leq \frac{\varepsilon}{2} + \frac{M}{2n\delta^2}.$$

On note que cette estimation est indépendante de x . Il reste maintenant à prendre n assez grand tel que $M/n\delta^2 < \varepsilon$ et on obtient

$$\sup_{x \in [0,1]} \left| f(x) - \sum_{k=0}^n f\left(\frac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k} \right| < \varepsilon$$

ce qui conclut. □

Chapitre 3 : Topologie des espaces vectoriels normés

Ce chapitre est le cœur de ce cours. Nous allons introduire la notion de norme et toute la topologie qui en découle.

1 Espaces vectoriels normés

1.1 Définitions

Nous voulons munir un ensemble d'une notion de distance. Nous connaissons la distance usuelle dans le plan ou l'espace, dite euclidienne. Si on veut regarder un espace différent, qu'est-ce que doit vérifier une bonne notion de distance ?

Définition 3.1

Soit E un ensemble. Une **distance** sur E est une application

$$d : E \times E \longrightarrow [0, +\infty[$$

vérifiant

- la distance sépare les points : $d(a,b) = 0$ si et seulement si $a = b$,
- la distance est symétrique : $d(a,b) = d(b,a)$,
- l'inégalité triangulaire est vérifiée : $d(a,c) \leq d(a,b) + d(b,c)$.

Exemples :

- la distance euclidienne $d(x,y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$ est bien une distance sur \mathbb{R}^2
- on munit \mathbb{R}^2 de la *distance SNCF* : la distance entre deux points A et B est soit la distance euclidienne classique $\|\overrightarrow{AB}\|$ si A et B sont alignés avec l'origine O , soit la somme $\|\overrightarrow{AO}\| + \|\overrightarrow{OB}\|$ sinon. Cette distance fournit une notion de proximité différente de la distance classique car deux points de \mathbb{R}^2 peuvent être proches au sens classiques mais loin pour la distance SNCF car il faut « passer par O ».
- On peut définir une distance entre deux mots comme le nombre minimal

d'opérations (remplacements, ajouts ou suppressions de lettres) pour passer d'un mot à l'autre. Ainsi $d(\text{CHAT}, \text{CHIOT}) = 2$. Cette distance peut être utile pour les corrections automatiques (saisies sur smartphone...).

Nous pouvons maintenant parler de distance pour de nombreux ensembles et les munir ainsi d'une géométrie dont on peut étudier les propriétés. Dans ce cours, nous allons être moins général. Nous allons nous concentrer sur les espaces vectoriels réels. Nous rappelons qu'un *espace vectoriel réel* est un ensemble E que l'on a muni d'une opération $+$ et de la multiplication par un scalaire de telle sorte que

- $(E, +)$ est un groupe abélien : il existe un élément neutre 0 , la somme est commutative et associative et tout élément a un opposé.
- la multiplication par un scalaire réel est distributive et associative : $\lambda(u + v) = \lambda.u + \lambda.v$, $(\lambda + \mu).u = \lambda.u + \mu.u$ et $(\lambda\mu).u = \lambda.(\mu.u)$. Par ailleurs, l'élément neutre de cette opération est 1 , i.e. $1.u = u$.

On peut munir un espace vectoriel de tout un tas de distances. Mais il est intéressant que la notion de distance soit cohérente avec les opérations sur E . Pour ce faire, la notion fondamentale est la notion de norme.

Définition 3.2

Soit $(E, +, \cdot)$ un espace vectoriel réel. Une **norme** sur E est une fonction

$$x \in E \longmapsto \|x\| \in [0, +\infty[$$

telle que les propriétés suivantes soient vérifiées

- séparation : pour tout $x \in E$, $\|x\| = 0$ si et seulement si $x = 0$.
- homogénéité : pour tout $x \in E$ et $\lambda \in \mathbb{R}$, $\|\lambda x\| = |\lambda| \|x\|$.
- inégalité triangulaire : pour tous $x, y \in E$, $\|x + y\| \leq \|x\| + \|y\|$.

On dit que $(E, \|\cdot\|)$ est un **espace vectoriel normé**.

La norme $\|x\|$ d'un point $x \in E$ peut être vue comme la distance de x à 0 ou comme la taille du vecteur de $\vec{0x}$. La notion de distance canonique induite par une norme est donnée par la proposition suivante. Ce qui est important est que cette distance respecte les propriétés géométriques fondamentales des espaces vectoriels normés comme l'invariance par translation ou dilatation.

Proposition 3.3 (distance induite par une norme)

Soit $(E, \|\cdot\|)$ un espace vectoriel normé. La fonction

$$d : (x, y) \in E^2 \longmapsto d(x, y) = \|x - y\|$$

est une distance sur E qui est invariante par translation et homogène pour les homothéties.

Démonstration : Par définition, $d(x,y) = \|x - y\|$ est à valeurs positives. La séparation découle immédiatement de celle de la norme car $d(x,y) = \|x - y\| = 0$ équivaut à $x - y = 0$. La distance est aussi symétrique car

$$d(y,x) = \|y - x\| = \|(-1).(x - y)\| = |-1|\|x - y\| = \|x - y\| = d(x,y) .$$

De même, l'inégalité triangulaire découle de celle de la norme :

$$d(x,z) = \|x - z\| = \|(x - y) + (y - z)\| \leq \|x - y\| + \|y - z\| = d(x,y) + d(y,z) .$$

L'invariance par translation se traduit juste par

$$d(x + c, y + c) = \|(x + c) - (y + c)\| = \|x - y\| = d(x,y)$$

et l'homogénéité par

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

□

Les normes vérifient aussi des variantes de l'inégalité triangulaire de façon similaire à celles de la valeur absolue. Par exemple, en changeant y et $-y$ dans l'inégalité classique, on a

$$\|x - y\| = \|x + (-y)\| \leq \|x\| + \|-y\| = \|x\| + \|y\| .$$

La version suivante est importante pour minorer une norme.

Proposition 3.4

Pour tout x et y dans un espace vectoriel normé,

$$|\|x\| - \|y\|| \leq \|x - y\| \quad \text{et} \quad |\|x\| - \|y\|| \leq \|x + y\|$$

Démonstration : On a $\|x\| = \|(x - y) + y\| \leq \|x - y\| + \|y\|$ et donc $\|x\| - \|y\| \leq \|x - y\|$. Si on change les rôles de x et y , on obtient $\|y\| - \|x\| \leq \|y - x\| = \|x - y\|$. On a donc la majoration de $\|x\| - \|y\|$ et de son opposé, quel que soit son signe et on obtient la première estimation. La deuxième se démontre en changeant y en $-y$. □

Nous verrons dans les paragraphes suivants les exemples fondamentaux de ce cours. Donnons ici juste deux exemples triviaux.

Exemples :

- La valeur absolue $\|x\| = |x|$ est une norme sur \mathbb{R} . Dire cela signifie qu'on regarde \mathbb{R} comme un espace vectoriel réel, à une dimension, c'est-à-dire une droite. Les termes d'une multiplication ont un statut particulier : le premier réel est un scalaire et le second un point de l'espace vectoriel. Ce constat est important en tant que cas particulier de ce qui sera vu dans \mathbb{R}^d .

- La distance SNCF sur \mathbb{R}^2 n'est pas invariante par les translations du plan (l'origine jouant un rôle particulier). Cette distance n'est donc pas issue d'une norme sur \mathbb{R}^2 et est moins intéressante si le plan est considéré comme un espace vectoriel. Elle ne sera pas étudiée dans ce cours.

Maintenant que nous avons une distance, nous pouvons parler de ce qui est « proche », de « voisinage » etc.

Définition 3.5

Soit $(E, \|\cdot\|)$ un espace vectoriel normé.

La **boule ouverte** de centre $x \in E$ et de rayon $r > 0$ est l'ensemble

$$B(x,r) = \{y \in E, \|x - y\| < r\} .$$

La **boule fermée** de centre $x \in E$ et de rayon $r \geq 0$ est l'ensemble

$$\overline{B}(x,r) = \{y \in E, \|x - y\| \leq r\} .$$

La norme respectant la géométrie des espaces vectoriels, les boules vérifient aussi les propriétés associées.

Proposition 3.6

On peut passer d'une boule à l'autre par une translation et une homothétie. En particulier,

$$\forall x \in E, \forall r > 0, B(x,r) = x + rB(0,1) .$$

Démonstration : Par homogénéité de la norme, $\|x\| < 1$ si et seulement si $\|rx\| < r$ et donc $rB(0,1) = B(0,r)$. Par ailleurs $\|(x+y) - x\| = \|y\|$ donc y est dans $B(0,r)$ si et seulement si $(x+y)$ est dans $B(x,r)$. De cette formule, on déduit comment passer d'une boule à l'autre

$$B(x',r') = \frac{r'}{r}B(x,r) + \left(x' - \frac{r'}{r}x\right) . \quad \square$$

1.2 Exemples dans \mathbb{R}^d

Dans un espace vectoriel de dimension $d \geq 1$, on se repère par rapport aux coordonnées dans une base de d vecteurs. On peut donc toujours se ramener à \mathbb{R}^d qui est le modèle de tout espace vectoriel de dimension finie. On peut munir \mathbb{R}^d d'une infinité de normes différentes. Dans ce paragraphe, nous allons voir les normes p qui sont les plus importantes et peuvent servir de modèles pour construire les autres normes.

Commençons par le cas de deux normes dont l'expression est simple.

Définition 3.7

Dans \mathbb{R}^d ($d \geq 1$), on appelle **norme infini** ou **norme du max** la fonction

$$x = (x_1, \dots, x_d) \in \mathbb{R}^d \longmapsto \|x\|_\infty = \max(|x_1|, |x_2|, \dots, |x_d|) .$$

Proposition 3.8

La norme infini est une norme sur \mathbb{R}^d .

Démonstration : Il est clair que $\|x\|_\infty$ est un réel positif et que $\|(0, 0, \dots)\|_\infty = 0$. Si $\|x\|_\infty = 0$, alors le maximum des $|x_i|$ est nul. Comme chaque quantité est positive, c'est que tous les x_i sont nuls et donc $x = 0_{\mathbb{R}^d}$. L'homogénéité est aussi assez simple :

$$\begin{aligned} \|\lambda x\|_\infty &= \max(|\lambda x_1|, \dots, |\lambda x_d|) = \max(|\lambda| |x_1|, \dots, |\lambda| |x_d|) \\ &= |\lambda| \max(|x_1|, \dots, |x_d|) = |\lambda| \|x\|_\infty . \end{aligned}$$

Enfin, l'inégalité triangulaire découle de celle sur \mathbb{R} . Soient x et y donnés, notons i_0 une coordonnée telle que $|x_i + y_i|$ soit le plus grand possible (cette coordonnée existe car il n'y en a qu'un nombre fini). On a

$$\begin{aligned} \|x + y\|_\infty &= \max(|x_1 + y_1|, \dots, |x_d + y_d|) = |x_{i_0} + y_{i_0}| \leq |x_{i_0}| + |y_{i_0}| \\ &\leq \max(|x_1|, \dots, |x_d|) + \max(|y_1|, \dots, |y_d|) = \|x\|_\infty + \|y\|_\infty . \quad \square \end{aligned}$$

Définition 3.9

Dans \mathbb{R}^d ($d \geq 1$), on appelle **norme 1** la fonction

$$x = (x_1, \dots, x_d) \in \mathbb{R}^d \longmapsto \|x\|_1 = |x_1| + |x_2| + \dots + |x_d| .$$

Proposition 3.10

La norme 1 est une norme sur \mathbb{R}^d .

Démonstration : Les propriétés sont aussi évidentes que pour la norme infini, sauf l'inégalité triangulaire qui ne se montre pas exactement pareil :

$$\begin{aligned} \|x + y\|_1 &= |x_1 + y_1| + \dots + |x_d + y_d| \leq |x_1| + |y_1| + \dots + |x_d| + |y_d| \\ &\leq |x_1| + \dots + |x_d| + |y_1| + \dots + |y_d| = \|x\|_1 + \|y\|_1 . \quad \square \end{aligned}$$

Ces deux normes induisent des notions de distance différentes. Imaginons une expérience ou un relevé statistique dont les résultats devraient former une liste de

cent valeurs idéales $x = (x_1, x_2, \dots, x_{100})$. Imaginons qu'il y ait des erreurs (de mesure, de fabrication...) et que la vraie liste soit $\tilde{x} = (x_1 + e_1, \dots)$, i.e. la liste des erreurs est $e = \tilde{x} - x = (e_1, \dots, e_{100})$. Si on mesure la taille de l'erreur en norme infini, on a $\|e\|_\infty$ qui est la plus grande de toutes les erreurs. Il suffit d'une seule erreur pour que cette distance $d_\infty(x, \tilde{x})$ soit grande. En outre, cette distance ne voyant que cette plus grande erreur, on ne voit pas s'il y a beaucoup de grandes erreurs dans la liste, ou juste une seule. Inversement la distance mesurée par la norme 1 est la somme de toutes les erreurs, donc la moyenne des erreurs à un facteur près. On voit mieux s'il y a beaucoup d'erreurs. Mais d'un autre côté, on ne sait pas s'il s'agit de beaucoup de petites erreurs acceptables ou bien d'une unique qui serait trop grande pour être tolérée.

e	$\ e\ _\infty$	$\ e\ _1$
$(1, 0, 0, \dots, 0)$	1	1
$(1, 1, 0, \dots, 0)$	1	2
$(1, 1, 1, \dots, 1)$	1	100
$(100, 0, 0, \dots, 0)$	100	100

Que dire de la norme euclidienne classique ? En fait, elle fait partie de la grande famille des normes p pour l'exposant $p = 2$. C'est aussi le cas de la norme 1 (pour $p = 1$). Le nom de la norme infini vient de ce qu'il s'agit de la limite $p \rightarrow +\infty$ dans la formule de la norme p .

Définition 3.11

Dans \mathbb{R}^d ($d \geq 1$), pour $p \geq 1$, on appelle **norme p** la fonction

$$x = (x_1, \dots, x_d) \in \mathbb{R}^d \longmapsto \|x\|_p = (|x_1|^p + |x_2|^p + \dots + |x_d|^p)^{1/p}.$$

Pour démontrer qu'il s'agit en général d'une norme, on peut commencer par constater que l'homogénéité est vérifiée car

$$\|\lambda x\|_p = (|\lambda x_1|^p + \dots + |\lambda x_d|^p)^{1/p} = (|\lambda|^p (|x_1|^p + \dots + |x_d|^p))^{1/p} = |\lambda| \|x\|_p.$$

Le point clef est l'inégalité triangulaire qui demande d'abord l'inégalité de Hölder.

Proposition 3.12 (inégalité de Hölder)

Soient x et y dans \mathbb{R}^d . Pour tous p et q dans $[1, +\infty]$ avec $1/p + 1/q = 1$, on a

$$\|xy\|_1 = \sum_{i=1}^d |x_i y_i| \leq \|x\|_p \|y\|_q.$$

Démonstration : Les cas $p = 1$ ou $p = +\infty$ sont particulier à cause de l'expression de la norme infini. Ces cas sont plus simples et comme de toute façon nous savons déjà que les normes 1 et infinie sont des normes, nous les laissons au lecteur. De même, les cas $x = 0$ ou $y = 0$ sont triviaux, donc on va supposer que toutes les normes sont non nulles.

On commence par utiliser que le logarithme est concave et donc que pour tout

$a, b > 0$, on a

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

L'exponentielle étant croissante, on peut l'appliquer de chaque côté de l'inégalité pour obtenir

$$\frac{1}{p}a + \frac{1}{q}b \geq a^{1/p}b^{1/q} .$$

Utilisons cette inégalité sur notre estimation. On va répartir un poids $\theta > 0$ à choisir plus tard en prenant $a = \theta|x_i|^p$ et $b = \theta^{-1}|y_i|^q$:

$$\sum_{i=1}^d |x_i y_i| = \sum_{i=1}^d |(\theta x_i)(\theta^{-1} y_i)| \leq \sum_{i=1}^d \left(\frac{1}{p} \theta^p |x_i|^p + \frac{1}{q} \theta^{-q} |y_i|^q \right) = \frac{\theta^p}{p} \|x\|_p^p + \frac{\theta^{-q}}{q} \|y\|_q^q$$

Nous cherchons maintenant à optimiser cette estimation : plus le majorant sera petit, meilleure l'estimation sera. Quand $\theta \rightarrow 0$ ou $\theta \rightarrow +\infty$, le majorant tend vers $+\infty$. Le minimum se trouve au milieu. En dérivant, on trouve

$$\frac{d}{d\theta} \left(\frac{\theta^p}{p} \|x\|_p^p + \frac{\theta^{-q}}{q} \|y\|_q^q \right) = \theta^{p-1} \|x\|_p^p - \theta^{-q-1} \|y\|_q^q .$$

Le minimum se trouve à l'endroit où la dérivée s'annule, qui correspond à $\theta^p \|x\|_p^p = \theta^{-q} \|y\|_q^q$. Comme $1/p + 1/q = 1$ et donc aussi $pq = p + q$, on obtient

$$\theta = \left(\frac{\|y\|_q^q}{\|x\|_p^p} \right)^{1/(p+q)} = \frac{\|y\|_q^{1/p}}{\|x\|_p^{1/q}} .$$

En utilisant que $p - p/q = p(1 - 1/q) = p/p = 1$ et l'expression symétrique, le majorant optimisé en θ devient

$$\begin{aligned} \frac{\theta^p}{p} \|x\|_p^p + \frac{\theta^{-q}}{q} \|y\|_q^q &= \frac{1}{p} \|x\|_p^{p-p/q} \|y\|_q + \frac{1}{q} \|x\|_p \|y\|_q^{q-q/p} \\ &= \left(\frac{1}{p} + \frac{1}{q} \right) \|x\|_p \|y\|_q = \|x\|_p \|y\|_q . \end{aligned}$$

□

Proposition 3.13 (inégalité de Minkowski)

Pour tout $p \geq 1$ et pour tous x et y dans \mathbb{R}^d , on a

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p .$$

Démonstration : Le cas $p = 1$ a déjà été fait plus haut. Nous supposons donc $p \in]1, +\infty[$. Soit $q \in]1, +\infty[$ l'exposant conjugué, c'est-à-dire tel que

$1/p + 1/q = 1$. On écrit

$$\|x + y\|_p^p = \sum_{i=1}^d |x_i + y_i| \cdot |x_i + y_i|^{p-1} \leq \sum_{i=1}^d |x_i| \cdot |x_i + y_i|^{p-1} + \sum_{i=1}^d |y_i| \cdot |x_i + y_i|^{p-1} .$$

L'inégalité de Hölder nous donne que

$$\sum_{i=1}^d |x_i| \cdot |x_i + y_i|^{p-1} \leq \|x\|_p \left(\sum_{i=1}^d |x_i + y_i|^{q(p-1)} \right)^{1/q} .$$

Mais $q = 1/(1 - 1/p) = p/(p - 1)$ et donc $q(p - 1) = p$ et

$$\sum_{i=1}^d |x_i| \cdot |x_i + y_i|^{p-1} \leq \|x\|_p \|x + y\|_p^{p/q} .$$

Si on applique la même idée à l'autre terme, on obtient que

$$\|x + y\|_p^p \leq (\|x\|_p + \|y\|_p) \|x + y\|_p^{p/q} .$$

Il reste à voir que $p - (p/q) = p(1 - 1/q) = p/p = 1$ pour simplifier de chaque côté les puissances de $\|x + y\|_p$. \square

La proposition précédente étant exactement l'inégalité triangulaire pour la norme $p \geq 1$, nous avons bien démontré ce qu'il fallait pour obtenir le résultat principal de ce paragraphe.

Proposition 3.14

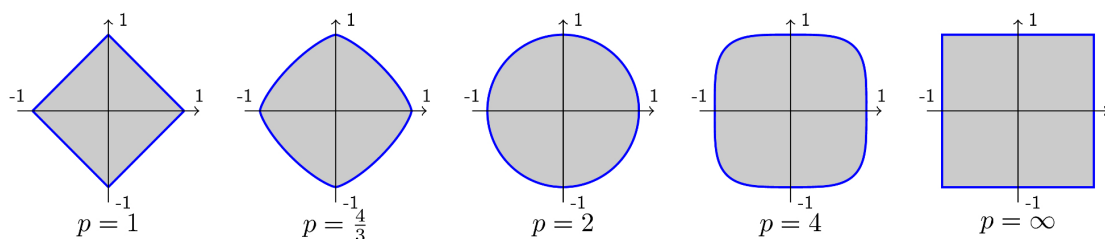
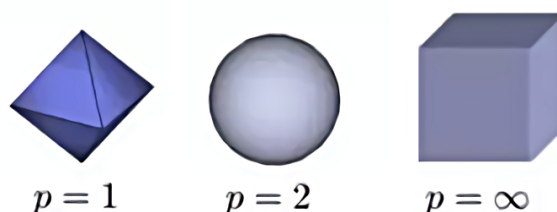
Pour tout $p \in [1, +\infty]$, la norme p est une norme sur \mathbb{R}^d .

Attention que la même expression $\|x\|_p$ avec $p < 1$ ne donne pas une norme car l'inégalité triangulaire devient fausse. Les normes 1 et infinie sont donc les deux extrêmes de cette famille de normes.

Il est intéressant de regarder les différentes formes des boules données par ces normes. Il s'agit d'un exercice de géométrie assez simple pour le cas 1 et infini et on obtient des polygones ou polyèdres dans les deux cas. Évidemment, le cas $p = 2$ donne la boule classique. Les cas intermédiaires sont des interpolations de ces trois cas.

1.3 Espaces de fonctions

Ce cours considère aussi des espaces de dimension infinie. Nous regarderons simplement quelques espaces de fonctions comme exemples types.


 FIGURE 3.1 – Les boules unités $B_p(0,1) = \{x \in \mathbb{R}^2, \|x\|_p < 1\}$ du plan.

 FIGURE 3.2 – Les boules unités $B_p(0,1) = \{x \in \mathbb{R}^2, \|x\|_p < 1\}$ de l'espace. La différence entre les cas $p = 1$ et $p = \infty$ devient plus claire.

Proposition 3.15

Soit $D \subset \mathbb{R}$ un domaine non vide et soit $\mathcal{B}(D, \mathbb{R})$ l'ensemble des fonctions $f : D \rightarrow \mathbb{R}$ bornées. Pour tout $f \in \mathcal{B}(D, \mathbb{R})$, on pose

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

L'espace $\mathcal{B}(D, \mathbb{R})$ muni de $\|\cdot\|_\infty$ est un espace vectoriel normé.

Démonstration : On vérifie facilement que $\mathcal{B}(D, \mathbb{R})$ est un espace vectoriel (la somme de fonctions bornées est bornée etc.). Le fait que $\|\cdot\|_\infty$ est bien une norme se vérifie comme pour la proposition 3.8. Il y a juste le cas de la somme où nous avons utilisé que le maximum était atteint dans \mathbb{R}^d . Ici, il faut revenir aux propriétés du sup. Nous prétendons que

$$\sup_{x \in D} |f(x) + g(x)| \leq \sup_{x \in D} |f(x)| + \sup_{x \in D} |g(x)| . \quad (3.1)$$

En effet, si $x \in D$, alors

$$|f(x) + g(x)| \leq |f(x)| + |g(x)| \leq \sup_{x' \in D} |f(x')| + \sup_{x' \in D} |g(x')| .$$

Le terme de droite est donc bien un majorant de $\{|f(x) + g(x)|\}$ et par définition de la borne sup, l'inégalité triangulaire (3.1) est vérifiée. \square

Plutôt que ce résultat général, nous utiliserons surtout les deux corollaires suivants.

Corollaire 3.16

Soient $a < b$ deux réels. L'espace $\mathcal{C}^0([a,b],\mathbb{R})$ des fonctions continues réelles sur $[a,b]$, muni de la norme infini

$$\|f\|_\infty = \sup_{x \in [a,b]} |f(x)|$$

est un espace vectoriel normé.

Démonstration : Il est facile de vérifier que $\mathcal{C}^0([a,b],\mathbb{R})$ est bien un espace vectoriel. Par ailleurs, si f est continue sur $[a,b]$ compact, alors f est bornée (cf le théorème 2.37). Donc $\mathcal{C}^0([a,b],\mathbb{R})$ est un sous-espace vectoriel de $\mathcal{B}([a,b],\mathbb{R})$. Il reste maintenant à remarquer que si $\|\cdot\|$ est une norme sur un espace E , alors c'est aussi une norme sur un sous-espace $F \subset E$ (c'est évident car toutes les conditions sont a fortiori vraies sur F qui est plus petit). La proposition 3.15 nous montre donc que $\|\cdot\|_\infty$ est bien une norme sur $\mathcal{C}^0([a,b],\mathbb{R})$. \square

Corollaire 3.17

Soit ℓ^∞ l'espace des suites bornées, i.e.

$$\ell^\infty = \{ u = (u_n) \in \mathbb{R}^\mathbb{N} \mid \exists M > 0, \forall n \in \mathbb{N}, |u_n| < M \} .$$

La norme infini définie par

$$\forall u = (u_n)_{n \in \mathbb{N}} \in \ell^\infty, \quad \|u\|_\infty = \sup_{n \in \mathbb{N}} |u_n| .$$

est bien une norme sur ℓ^∞ .

Démonstration : On applique la proposition 3.15 au cas $D = \mathbb{N}$. \square

Corollaire 3.18

Soit $\mathbb{R}[X]$ l'espace des polynômes. Si $P \in \mathbb{R}[X]$ est de degré d , il s'écrit $P = a_0 + a_1X + a_2X^2 + \dots + a_dX^d$ et on pose $\|P\| = \max_i |a_i|$. Ceci définit bien une norme sur $\mathbb{R}[X]$.

Démonstration : Si on représente un polynôme par la suite de ces coefficients, on peut associer à chaque $P \in \mathbb{R}[X]$ une suite (a_n) de ℓ^∞ qui a en plus la propriété d'être nulle à partir d'un certain rang (car P n'a qu'un nombre fini de monômes) :

$$\mathbb{R}[x] \longleftrightarrow \{ (a_n)_{n \in \mathbb{N}} \in \ell^\infty, \exists p \in \mathbb{N}, \forall i > p, a_i = 0 \} .$$

Il s'agit bien d'un sous-espace de ℓ^∞ et la norme proposée est exactement la norme $\|\cdot\|_\infty$ du corollaire 3.17. Ses propriétés sont a fortiori vraies si on les considère sur un espace plus petit et donc on a bien une norme sur l'espace des polynômes. \square

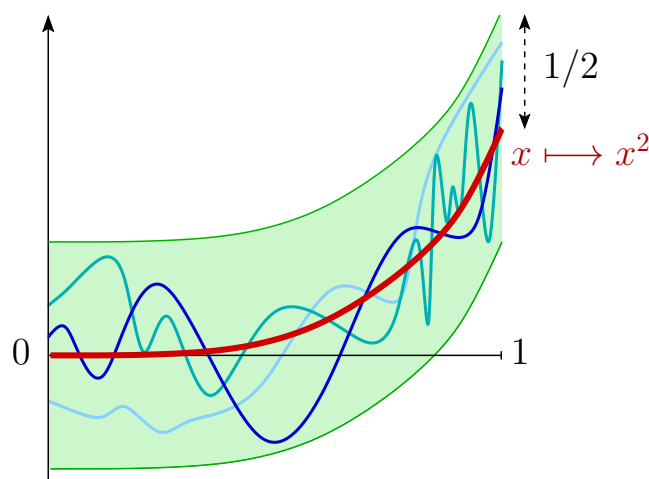


FIGURE 3.3 – On considère $\mathcal{C}^0([0,1])$ muni de la norme infini. La boule de centre $f : x \mapsto x^2$ et de rayon $1/2$ contient toutes les fonctions continues dont le graphe reste entre $x^2 - 1/2$ et $x^2 + 1/2$.

On peut aussi munir $\mathcal{C}^0([a,b],\mathbb{R})$ d'autres normes.

Proposition 3.19

Soient $a < b$ deux réels. L'espace $\mathcal{C}^0([a,b],\mathbb{R})$ muni de la norme 1

$$\|f\|_1 = \int_a^b |f(x)| dx$$

est un espace vectoriel normé.

Démonstration : Commençons par noter que sur un compact $[a,b]$, l'intégrale d'une fonction continue est toujours bien définie. L'inégalité triangulaire et l'homogénéité se déduisent des propriétés de la valeur absolue et de l'intégrale. Par exemple

$$\|f + g\|_1 = \int_a^b |f(x) + g(x)| dx \leq \int_a^b (|f(x)| + |g(x)|) dx$$

par l'inégalité triangulaire sur \mathbb{R} et car l'intégrale est monotone. Puis la linéarité de l'intégrale implique que

$$\int_a^b (|f(x)| + |g(x)|) dx = \int_a^b |f(x)| dx + \int_a^b |g(x)| dx = \|f\|_1 + \|g\|_1 .$$

La subtilité ici est de montrer que si $\|f\|_1 = 0$, alors $f \equiv 0$. Pour cela, il nous faut utiliser que f est continue. Supposons que $f \not\equiv 0$, il existe $c \in [a,b]$ tel que $f(c) \neq 0$. Soit $\varepsilon = |f(c)|/2 > 0$, par définition de la continuité, il existe $\delta > 0$ tel que pour tout $x \in [a,b]$ avec $|x - c| < \delta$, on a $|f(x) - f(c)| < \varepsilon$. On obtient une minoration de $|f(x)|$ pour tout $x \in [a,b] \cap]c - \delta, c + \delta[$ en écrivant

$$|f(x)| = |f(c) + f(x) - f(c)| \geq |f(c)| - |f(x) - f(c)| \geq 2\varepsilon - \varepsilon = \varepsilon.$$

Donc la fonction $|f|$ est positive sur $[a,b]$ et plus grande que $\varepsilon > 0$ sur un intervalle $[a,b] \cap]c - \delta, c + \delta[$ de longueur $\eta > 0$. Donc $\|f\|_1 \geq \varepsilon\eta > 0$. Par contraposée, si $\|f\|_1 = 0$, c'est que f est identiquement nulle. \square

1.4 Espaces préhilbertiens

Les espaces préhilbertiens sont des structures importantes de l'analyse et de la géométrie. Dans ce cours, nous ne les étudierons pas en détail mais nous les verrons comme des cas particuliers d'espaces vectoriels normés.

Définition 3.20

Soit E un espace vectoriel. On dit que la fonction

$$(x,y) \in E^2 \longmapsto \langle x|y \rangle \in \mathbb{R}$$

est un **produit scalaire** sur E si

- symétrie : pour tous x et y dans E , $\langle y|x \rangle = \langle x|y \rangle$.
- bilinéarité : pour tout $y \in E$, la fonction $x \mapsto \langle x|y \rangle$ est linéaire. Ceci entraînant par symétrie que pour tout $x \in E$, la fonction $y \mapsto \langle x|y \rangle$ est aussi linéaire.
- positivité : pour tout x dans E , $\langle x|x \rangle \geq 0$.
- séparation : pour tout x dans E , $\langle x|x \rangle = 0$ si et seulement si $x = 0$.

Définition 3.21

Un espace vectoriel E muni d'un produit scalaire est appelé un **espace préhilbertien**.

Un espace préhilbertien est un cas particulier d'espace vectoriel normé car le produit scalaire engendre une norme naturelle. Sauf mention du contraire, cette norme associée sera la norme sur l'espace préhilbertien considéré.

Proposition 3.22

Soit E un espace préhilbertien. La fonction

$$x \in E \longmapsto \|x\| = \sqrt{\langle x|x \rangle}$$

est une norme sur E .

Démonstration : Par positivité, l'expression $\sqrt{\langle x|x \rangle}$ est bien définie. La séparation de la norme vient directement de la séparation du produit scalaire. L'homogénéité se déduit de la bilinéarité puisque

$$\|\lambda x\| = \sqrt{\langle \lambda x|\lambda x \rangle} = \sqrt{\lambda \langle x|\lambda x \rangle} = \sqrt{\lambda^2 \langle x|x \rangle} = |\lambda| \sqrt{\langle x|x \rangle} = |\lambda| \|x\| .$$

Avant de considérer l'inégalité triangulaire, nous allons généraliser l'inégalité de Cauchy-Schwarz à tous les produits scalaires. Soit t un réel et x et y deux vecteurs. On sait que $\|x + ty\|^2$ est positif. En utilisant la bilinéarité du produit scalaire, on développe

$$\begin{aligned} \|x + ty\|^2 &= \langle (x + ty)|(x + ty) \rangle = \langle x|(x + ty) \rangle + t \langle y|(x + ty) \rangle \\ &= \langle x|x \rangle + t \langle x|y \rangle + t \langle y|x \rangle + t^2 \langle y|y \rangle \\ &= \|x\|^2 + 2t \langle x|y \rangle + t^2 \|y\|^2 . \end{aligned}$$

Le développement montre qu'il s'agit d'un polynôme du second degré en t . S'il reste positif pour tout $t \in \mathbb{R}$, son discriminant doit être négatif ou nul. Donc

$$4 \langle x|y \rangle^2 - 4 \|x\|^2 \|y\|^2 \leq 0$$

ce qui donne

$$|\langle x|y \rangle| \leq \|x\| \|y\| .$$

Pour montrer l'inégalité triangulaire, on reprend le développement ci-dessus pour $t = 1$ et l'inégalité qu'on vient de prouver et on écrit

$$\|x + y\|^2 = \|x\|^2 + 2 \langle x|y \rangle + \|y\|^2 \leq \|x\|^2 + 2 \|x\| \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2 .$$

Il ne reste plus qu'à prendre la racine carrée de tous les termes (qui sont positifs). \square

Au passage de la démonstration ci-dessus, nous avons généralisé une inégalité qui nous est familière dans \mathbb{R}^d .

Proposition 3.23 (inégalité de Cauchy-Schwarz)

Soit $(E, \langle \cdot | \cdot \rangle)$ un espace préhilbertien. Pour tout x et y dans E , on a

$$|\langle x | y \rangle| \leq \|x\| \|y\| .$$

Il y a égalité si et seulement si les deux vecteurs x et y sont colinéaires.

Démonstration : Il nous reste à regarder les cas d'égalité. Si x et y sont colinéaires, disons que $x = \lambda y$, alors un calcul direct nous montre qu'il y a égalité

$$|\langle x | y \rangle| = |\langle \lambda y | y \rangle| = |\lambda \langle y | y \rangle| = |\lambda| \|y\|^2 = \|\lambda y\| \|y\| = \|x\| \|y\| .$$

S'il y a égalité, c'est que le polynôme du second degré apparaissant dans la preuve de l'inégalité a un discriminant nul. Dans ce cas, on sait que le polynôme a exactement un zéro, c'est-à-dire qu'il existe t tel que $\|x + ty\|^2 = 0$. Par propriété de séparation, on a $x + ty = 0$ et donc x et y sont colinéaires. \square

Les deux exemples importants pour nous sont les suivants (démonstrations laissées au lecteur).

Proposition 3.24

Soit $d \geq 1$. L'espace \mathbb{R}^d muni du produit scalaire

$$\langle x | y \rangle = \sum_{i=1}^d x_i y_i$$

est un espace préhilbertien dont la norme associée est la norme $\|\cdot\|_2$.

Proposition 3.25

Soient $a < b$ deux réels. L'espace $\mathcal{C}^0([a,b], \mathbb{R})$ muni du produit scalaire

$$\langle f | g \rangle = \int_a^b f(x)g(x)dx$$

est un espace préhilbertien dont la norme associée est la norme

$$\|f\|_2 = \left(\int_a^b |f(x)|^2 dx \right)^{1/2} .$$

Ce ne sont pas les seuls exemples que l'on a de produits scalaires. Par exemple, on peut simplement changer de base canonique dans \mathbb{R}^d et/ou mettre des poids sur les coordonnées, comme prendre $\langle x | y \rangle = 2x_1y_1 + (x_1 + x_2)(y_1 + y_2)$. Mais la plupart

des normes ne sont pas associées à des produits scalaires. Par exemple sur \mathbb{R}^d , alors il est impossible de trouver un produit scalaire induisant la norme p pour $p \neq 2$.

Il faut comprendre les espaces préhilbertiens comme des espaces vectoriels normés dans lesquels il y a une structure en plus donnée par le produit scalaire. Cette structure permet de parler d'angle, d'orthogonalité, de projection orthogonale... On peut par exemple avoir une version du théorème de Pythagore sur les fonctions : si f et g sont orthogonales, alors $\|f + g\|^2 = \|f\|^2 + \|g\|^2$. Ces notions sont absentes des espaces vectoriels normés sans produit scalaire. C'est pour cela que les normes 2 sont souvent privilégiés. Ainsi, on fait des interpolations suivant les moindres carrés, c'est-à-dire en mesurant les distances selon une norme 2. Si on mesure les distances avec la norme 1 ou infini, d'apparence plus simple, on perd l'outil important qui est la projection orthogonale et cela devient plus compliqué en pratique. Avec la norme 2, on aura au contraire des formules simples pour la projection.

Finissons avec une question de vocabulaire. Si ces espaces sont *préhilbertien*, c'est qu'il existe des espaces dits *de Hilbert*. Ce sont des espaces préhilbertiens qui ont en plus la propriété d'être complets pour la norme du produit scalaire. Nous verrons cette notion en fin de cours. Par ce qui est de \mathbb{R}^d muni d'un produit scalaire, on parle aussi d'espace *euclidien* (même si on n'a pas pris le produit scalaire classique). Ce nom correspond évidemment au fait que *Les Éléments* d'Euclide (probablement vers 300 av.J.C. en Grèce) forment un livre fondateur des mathématiques, incluant en particulier toute la géométrie du plan et de l'espace connue alors... qui correspond évidemment à la géométrie de la distance $\|\cdot\|_2$.

2 Topologie des espaces vectoriels normés : définitions de base

De façon générale, une topologie sur un ensemble est la donnée de ses ouverts. On définit ensuite les notions comme la convergence des suites, la continuité des fonctions etc. à l'aide des ouverts. Dans un espace vectoriel normé, il y a une notion d'ouvert naturellement associée aux boules ouvertes de la norme et de même qu'il y a une notion naturelle de convergence des suites. Ces deux notions sont compatibles et on aura souvent plusieurs définitions équivalentes : une utilisant la notion de boule ouverte et une utilisant les suites. Si on travaille dans des espaces plus généraux, seule la première est la « bonne » caractérisation. On fera parfois des remarques en ce sens pour les lecteurs qui ont déjà vu ou verront une notion de topologie plus générale. Mais pour ce cours, les deux façons de faire seront équivalentes et toutes les deux correctes.

Dans toute cette partie, on travaille dans un espace vectoriel E muni d'une norme $\|\cdot\|$. Comme dit plus haut, mathématiquement parlant, une topologie est la donnée des ouverts d'un ensemble. Les fermés sont par définition les complémentaires des ouverts.

Définition 3.26

Un sous-ensemble $\mathcal{O} \subset E$ est dit **ouvert** si pour tout point x de \mathcal{O} , il existe un rayon $r > 0$ tel que la boule ouverte $B(x,r)$ soit incluse dans \mathcal{O} .
 Un sous-ensemble $\mathcal{F} \subset E$ est dit **fermé** si son complémentaire \mathcal{F}^C est ouvert.

Nos premiers exemples sont assez simples mais universels.

Proposition 3.27

Les ensembles E et \emptyset sont à la fois ouverts et fermés dans E .

Démonstration : Comme toute boule $B(x,r)$ est incluse dans l'espace E par définition, E est clairement ouvert. Par ailleurs, l'ensemble vide vérifie toute proposition commençant par « pour tout x dans l'ensemble » puisqu'il n'y a aucun x dedans et donc aucune condition à vérifier. Donc \emptyset est aussi un ouvert. Par complémentaire, E et \emptyset sont aussi des fermés. \square

Proposition 3.28

Soit $x \in E$ et $r > 0$. La boule ouverte $B(x,r)$ est un ouvert de E et la boule fermée $\overline{B}(x,r)$ est un fermé de E .

Démonstration : Soit $y \in B(x,r)$. On pose $\rho = r - \|x - y\|$. Par définition de la boule ouverte, on a $\rho > 0$. Soit $z \in B(y,\rho)$. L'inégalité triangulaire implique que

$$\|x - z\| = \|(x - y) + (y - z)\| \leq \|x - y\| + \|y - z\| < \|x - y\| + \rho = r .$$

Donc z est dans la boule $B(x,r)$. Ceci montre que $B(y,\rho) \subset B(x,r)$ et donc que $B(x,r)$ est un ouvert de E .

Soit y qui est dans le complémentaire de $\overline{B}(x,r)$. Par définition, on doit avoir $\|x - y\| > r$. On pose $\rho = \|x - y\| - r$. Soit $z \in B(y,\rho)$. La deuxième version de l'inégalité triangulaire implique que

$$\|x - z\| = \|(x - y) + (y - z)\| \geq \|x - y\| - \|y - z\| > \|x - y\| - \rho = r .$$

Donc z n'est pas dans la boule $\overline{B}(x,r)$. Ceci montre que $B(y,\rho) \subset \overline{B}(x,r)^C$ et donc que $\overline{B}(x,r)^C$ est un ouvert de E . Par définition, $\overline{B}(x,r)$ est donc un fermé de E . \square

Nous allons tout de suite donner une caractérisation équivalente des fermés à l'aide des suites convergentes.

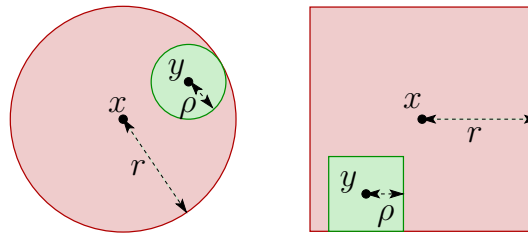


FIGURE 3.4 – Une illustration de la preuve que la boule ouverte est ouverte : à gauche dans \mathbb{R}^2 muni de la norme euclidienne, à droite dans \mathbb{R}^2 muni de la norme infini.

Définition 3.29

Une suite (x_n) d'un espace vectoriel normé E **converge vers** $\ell \in E$ si

$$\|x_n - \ell\| \xrightarrow{n \rightarrow +\infty} 0.$$

Une suite qui converge vers un point $\ell \in E$ est dite **convergente**. Une suite qui ne converge pas est dite **divergente**.

Exemple :

Dans \mathbb{R}^d muni de la norme infini, une suite $(x_n) \subset \mathbb{R}^d$ tend vers ℓ si et seulement si chaque coordonnée de x_n converge vers la coordonnée correspondante de ℓ . Plus précisément, pour tout n , on écrit $x_n = (x_n^1, x_n^2, \dots, x_n^d)$ et $\ell = (\ell^1, \ell^2, \dots, \ell^d)$. Si $x_n \rightarrow \ell$, alors on a $\|x_n - \ell\|_\infty \rightarrow 0$ par définition. Comme pour tout i , on a $|x_n^i - \ell^i| \leq \|x_n - \ell\|_\infty$, cela montre la convergence coordonnée par coordonnée. Supposons maintenant que la $x_n^i \rightarrow \ell^i$ pour tout $i = 1, \dots, d$. Soit $\varepsilon > 0$. Pour tout i , il existe un rang N_i tel que $|x_n^i - \ell^i| < \varepsilon$ (convergence dans \mathbb{R}). On pose $N = \max_i N_i$, on a alors pour tout $n \geq N$, $\|x_n - \ell\|_\infty = \max_i |x_n^i - \ell^i| < \varepsilon$.

La caractérisation suivante pourra être utilisée comme définition équivalente dans les espaces vectoriels normés (mais pas dans les espaces topologiques généraux). Comme son énoncé fait intervenir un test par des suites, on parle de *caractérisation séquentielle* des fermés.

Proposition 3.30

Un ensemble $\mathcal{F} \subset E$ est fermé si et seulement si toute limite $\ell \in E$ d'une suite convergente $(x_n) \subset \mathcal{F}$ est dans \mathcal{F} .

Démonstration : Soit \mathcal{F} un fermé de E , par définition \mathcal{F}^C est ouvert. Soit (x_n) une suite qui converge vers $\ell \in E$. Si $\ell \in \mathcal{F}^C$, il existe $r > 0$ tel que la boule $B(\ell, r)$ est incluse dans \mathcal{F}^C . Par définition de la convergence, il existe $N \in \mathbb{N}$ tel que pour tout $n \geq N$, $\|x_n - \ell\| < r$. Ceci montre que la suite (x_n) finit dans la boule $B(\ell, r)$ et donc n'est plus dans \mathcal{F} . Si la suite (u_n) est toute

incluse dans \mathcal{F} , sa limite est donc aussi dans \mathcal{F} (c'est la contraposée de ce qu'on vient de démontrer).

Soit maintenant $\mathcal{F} \subset E$ un ensemble tel que toute suite $(u_n) \subset \mathcal{F}$ qui converge dans E a sa limite dans \mathcal{F} . Soit $x \notin \mathcal{F}$. Supposons qu'il n'existe aucune boule $B(x, 1/n)$ incluse dans \mathcal{F}^C . Alors pour tout n , il doit exister au moins un point $x_n \in B(x, 1/n) \cap \mathcal{F}$. Mais comme $\|x_n - x\| < 1/n$, la suite (x_n) converge vers x tout en étant dans \mathcal{F} . Par hypothèse, on doit avoir $x \in \mathcal{F}$, ce qui est absurde. Donc il existe forcément n tel que $B(x, 1/n) \subset \mathcal{F}^C$. On vient de montrer que \mathcal{F}^C est ouvert et donc que \mathcal{F} est fermé. \square

Notons bien qu'on prend dans la proposition ci-dessus une suite convergente de \mathcal{F} mais dont la limite n'est pas a priori dans \mathcal{F} . Mais pour un ensemble fermé, la suite ne peut s'en « échapper » et la limite reste dans \mathcal{F} .

Exemples :

- Dans un espace vectoriel normé E , un singleton $\{x\} \subset E$ est toujours un fermé de E . En effet, la seule suite possible incluse dans E est la suite constante égale à x qui converge vers x qui est bien dans E .
- Remontrons qu'une boule fermée $\overline{B}(x, r)$ est fermée par la caractérisation séquentielle. Soit (x_n) une suite de points de $\overline{B}(x, r)$ qui converge vers $\ell \in E$. On a

$$\|\ell - x\| = \|(\ell - x_n) + (x_n - x)\| \leq \|\ell - x_n\| + \|x_n - x\| \leq \|\ell - x_n\| + r.$$

En passant à la limite $n \rightarrow +\infty$, on trouve $\|\ell - x\| \leq r$ et donc ℓ est bien dans la boule $\overline{B}(x, r)$.

- La droite horizontale $\mathbb{R} \times \{0\}$ est un fermé de \mathbb{R}^2 muni de la norme infini. En effet, si une suite $u_n = (x_n, 0)$ converge vers $\ell = (x, y)$, on doit avoir $\max(|x - x_n|, |y|) \rightarrow 0$ et donc $y = 0$. La limite est donc bien dans $\mathbb{R} \times \{0\}$.
- On se place dans \mathbb{R} . Un intervalle $]a, b[$ est un ouvert car c'est la boule ouverte $B(\frac{a+b}{2}, \frac{b-a}{2})$. De la même façon, un intervalle $[a, b]$ est un fermé car c'est une boule fermée. Montrons que $[a, b[$ n'est ni ouvert ni fermé. La suite $u_n = b - (b - a)2^{-n}$ est dans $[a, b[$ et tend vers $b \notin [a, b[$, donc $[a, b[$ n'est pas fermé. Mais la suite $u_n = a - 2^{-n}$ est dans $[a, b[$ et tend vers $a \in [a, b[$, donc $[a, b[$ n'est pas fermé et donc son complémentaire $[a, b]$ n'est pas ouvert.

Le statut d'ouvert ou fermé passe bien aux intersections et unions finies. Par contre, il faut faire attention aux intersections et unions infinies. Le plus simple est de se rappeler qu'il y a des pièges et de retenir deux contre-exemples pour les unions et intersections infinies (voir ci-dessous), ce qui permet de savoir lesquels ne marchent pas.

Proposition 3.31

Soit $(\mathcal{O}_j)_{j \in J}$ une famille d'ouverts de cardinal quelconque. Alors l'union $\bigcup_{j \in J} \mathcal{O}_j$ est un ouvert de E .

Soit $p \in \mathbb{N}$ et $\mathcal{O}_n, n = 0, \dots, p$ un nombre fini d'ouverts. Alors l'intersection $\mathcal{O}_0 \cap \mathcal{O}_1 \cap \dots \cap \mathcal{O}_p$ est un ouvert.

Soit $(\mathcal{F}_j)_{j \in J}$ une famille de fermés de cardinal quelconque. Alors l'intersection $\bigcap_{j \in J} \mathcal{F}_j$ est un fermé de E .

Soit $p \in \mathbb{N}$ et $\mathcal{F}_n, n = 0, \dots, p$ un nombre fini de fermés. Alors l'union $\mathcal{F}_0 \cup \mathcal{F}_1 \cup \dots \cup \mathcal{F}_p$ est un fermé.

Démonstration : Soit $(\mathcal{O}_j)_{j \in J}$ une famille d'ouverts et soit $x \in \bigcup_{j \in J} \mathcal{O}_j$. Par définition, x appartient à un \mathcal{O}_{j_0} qui est ouvert. Il existe donc une boule $B(x, r)$ incluse dans \mathcal{O}_{j_0} et donc a fortiori dans $\bigcup_{j \in J} \mathcal{O}_j$. Ceci montre la première assertion.

Soit $\mathcal{O}_n, n = 0, \dots, p$ un nombre fini d'ouverts et soit $x \in \mathcal{O}_0 \cap \dots \cap \mathcal{O}_p$. Pour tout $n = 0, \dots, p$, il existe un rayon $r_n > 0$ tel que $B(x, r_n) \subset \mathcal{O}_n$. Soit $r = \min_{n=1 \dots p} r_n$. Comme il n'y a qu'un nombre fini de rayons, le min est bien atteint et r est strictement positif. Par construction, la boule $B(x, r)$ est dans tous les \mathcal{O}_n et donc dans leur intersection. Ceci montre la deuxième assertion. Les propriétés sur les fermés s'obtiennent en passant au complémentaire les propriétés sur les ouverts. \square

Exemples :

- Un ensemble fini $\{x_1, \dots, x_p\}$ de points de E est un fermé de E par union finie de singleton.
- L'ensemble $\{2^{-j}, j \in \mathbb{N}\}$ de \mathbb{R} est une union dénombrables de fermés qui n'est pas fermée. En effet, la suite $u_n = 2^{-n}$ est incluse dans l'ensemble mais sa limite 0 n'y est pas.
- L'ensemble $\bigcup_{n \in \mathbb{Z}}]n, n + 1[$ est un ouvert comme union (infinie) d'ouverts. Son complémentaire \mathbb{Z} est donc fermé (ne pas déduire que \mathbb{Z} est fermé comme l'union de singletons fermés $\bigcup_{n \in \mathbb{Z}} \{n\}$ puisqu'il s'agit d'une union infinie!).
- On voit facilement que l'intersection infinie d'intervalles ouverts $\bigcap_{n \in \mathbb{N}}] - 2^{-n}, 2^{-n}[$ se réduit à $\{0\}$. Or $\{0\}$ n'est pas ouvert puisqu'aucune boule $B(0, r) =] - r, r[$ avec $r > 0$ n'est entièrement incluse dans $\{0\}$. Donc une intersection infinie d'ouverts n'est pas forcément ouverte.

À chaque ensemble quelconque, on peut associer une partie ouverte et une partie fermée comme suit.

Définition 3.32

Soit $A \subset E$ un ensemble. On appelle **intérieur** de A et on note $\overset{\circ}{A}$ ou $\text{Adh}(A)$ le plus grand ouvert de E inclus dans A , c'est-à-dire $\overset{\circ}{A}$ est l'union de tous les ouverts inclus dans A .

On appelle **adhérence** de A et on note \overline{A} ou $\text{Int}(A)$ le plus grand fermé de E contenant A , c'est-à-dire \overline{A} est l'intersection de tous les fermés contenant A .

Soit $A \subset E$. On appelle **frontière** de A et on note ∂A (prononcer « d rond A ») l'ensemble

$$\partial A := \overline{A} \setminus \overset{\circ}{A}.$$

Cette définition contient en fait une propriété à démontrer : que l'union de tous les ouverts $(\mathcal{O}_j)_{j \in J}$ tels que $\mathcal{O}_j \subset A$ est bien le plus grand ouvert inclus dans A . Il s'agit bien d'un ouvert comme union d'ouverts et il est inclus dans A car chaque \mathcal{O}_j est inclus dans A . La notion de « plus grand » vient de l'ordre de l'inclusion : si \mathcal{O} est un ouvert contenu dans A , c'est un des \mathcal{O}_j et donc il est plus petit que l'union totale. La propriété analogue pour l'adhérence se fait en passant au complémentaire.

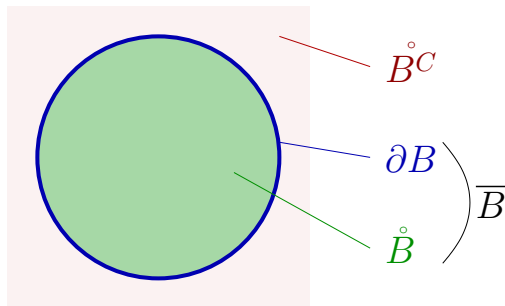


FIGURE 3.5 – Les différents ensembles topologiques associés à une boule B , qu'elle soit ouverte ou fermée.

De manière assez évidente, on a aussi la proposition suivante.

Proposition 3.33

Un ensemble A est ouvert si et seulement si $\overset{\circ}{A} = A$. Un ensemble A est fermé si et seulement si $\overline{A} = A$.

Démonstration : Si $A = \overset{\circ}{A}$, A est ouvert car son intérieur est par définition un ouvert. Si A est ouvert, il est difficile de faire un plus grand ouvert inclus dans A que A . Le cas fermé/adhérence est similaire. \square

Il existe des caractérisations pratiques pour les espaces vectoriels normés.

Proposition 3.34

Soit A un sous-ensemble d'un espace vectoriel normé E . L'intérieur de A est l'ensemble des centres de boules incluses dans A , i.e.

$$\overset{\circ}{A} = \{x \in A \mid \exists r > 0, B(x,r) \subset A\}.$$

L'adhérence de A est par symétrie

$$\overline{A} = \{x \in E \mid \forall r > 0, B(x,r) \cap A \neq \emptyset\}.$$

C'est aussi l'ensemble des limites de suites de A , i.e.

$$\overline{A} = \{x \in E \mid \exists (a_n)_{n \in \mathbb{N}} \subset A \ a_n \rightarrow x\}.$$

On peut aussi montrer qu'un ensemble F est l'adhérence de A en montrant qu'il s'agit d'un fermé contenant A dont tous les points sont limite d'une suite de A .

Concernant la frontière, on a que

$$\partial A = \overline{A} \cap \overline{A^C} = E \setminus (\overset{\circ}{A} \cup \overset{\circ}{A^C}).$$

et qu'elle se caractérise aussi comme l'ensemble des points x qui sont à la fois limite d'une suite de A et d'une suite de son complémentaire ou bien par

$$\partial A = \{x \in E, \forall r > 0, B(x,r) \cap A \neq \emptyset \text{ et } B(x,r) \cap A^C \neq \emptyset\}.$$

Démonstration : Si x est le centre d'une boule ouverte $B(x,r)$ incluse dans A , alors $B(x,r)$ étant un ouvert inclus dans A , on a $B(x,r) \subset \overset{\circ}{A}$. Ceci montre que $x \in \overset{\circ}{A}$. Réciproquement, supposons que $x \in \overset{\circ}{A}$. Comme $\overset{\circ}{A}$ est un ouvert, on doit avoir un rayon $r > 0$ tel que $B(x,r) \subset \overset{\circ}{A}$. Comme $\overset{\circ}{A}$ est inclus dans A , on a bien $B(x,r) \subset A$.

La première caractérisation de \overline{A} se fait par le complémentaire. En effet, \overline{A} est le complémentaire de l'intérieur de A^C . Donc $x \in \overline{A}$ si et seulement si il n'existe pas de rayon $r > 0$ tel que $B(x,r) \subset A^C$, i.e. si et seulement si $B(x,r) \cap A$ est non vide pour tout $r > 0$.

Montrons la deuxième caractérisation de \overline{A} . Soit x la limite d'une suite (a_n) de points de A . Comme \overline{A} contient A , on a $(a_n) \subset \overline{A}$. Mais comme \overline{A} est fermé, on trouve que la limite x est dans \overline{A} . Réciproquement, supposons que $x \in \overline{A}$. On considère la boule $B(x, 2^{-n})$. D'après la caractérisation précédente, il existe $a_n \in A \cap B(x, 2^{-n})$ pour tout $n \in \mathbb{N}$. On a donc une suite $(a_n) \subset A$ qui converge vers x par construction.

Si F est un fermé contenant A alors il contient \overline{A} . Mais si tous ses points sont limites d'une suite de A , alors la caractérisation précédente montre que $\overline{A} \subset F$. Donc $F = \overline{A}$.

Pour la frontière, on utilise les propriétés des intérieurs et adhérences. Par définition $\partial A = \overline{A} \setminus \overset{\circ}{A}$ ce qui est la même chose que $\overline{A} \cap (\overset{\circ}{A})^C = \overline{A} \cap \overline{(A^C)}$.

L'expression avec les intérieurs se trouve en passant au complémentaire. La deuxième partie se montre avec les caractérisations de l'adhérence de la proposition 3.34. \square

Enfin, énonçons les règles sur les unions et intersections. Nous donnerons des contre-exemples à toutes les propositions non-énoncées ci-dessous.

Proposition 3.35

Soient A et B deux ensembles de E . On a :

- si $A \subset B$ alors $\overset{\circ}{A} \subset \overset{\circ}{B}$ et $\overline{A} \subset \overline{B}$,
- $\text{Int}(A)^C = \text{Adh}(A^C)$ et $\text{Adh}(A)^C = \text{Int}(A^C)$.
- $\text{Int}(A \cap B) = \text{Int}(A) \cap \text{Int}(B)$ et $\text{Int}(A) \cup \text{Int}(B) \subset \text{Int}(A \cup B)$,
- $\text{Adh}(A \cup B) = \text{Adh}(A) \cup \text{Adh}(B)$ et $\text{Adh}(A \cap B) \subset \text{Adh}(A) \cap \text{Adh}(B)$.

Toutes les règles ci-dessus pour deux ensembles A et B passent par récurrences à un nombre fini d'ensembles.

Démonstration : Supposons que $A \subset B$. L'intérieur de A est un ouvert inclus dans A et a fortiori dans B . Il est donc inclus dans l'intérieur de B qui est le plus grand ouvert ayant cette propriété. L'adhérence de B est un fermé qui contient B et donc il contient a fortiori A . Il contient alors l'adhérence de A qui est le plus petit fermé ayant cette propriété.

L'ensemble $\text{Int}(A)^C$ est le complémentaire d'un ouvert contenu dans A , c'est donc un fermé contenant A^C . Soit F un autre fermé contenant A^C , F^C est un ouvert contenu dans A et donc il doit être contenu dans $\text{Int}(A)$. Donc F doit contenir $\text{Int}(A)^C$. Ceci montre que $\text{Int}(A)^C = \text{Adh}(A^C)$. L'autre propriété est similaire.

Soient A et B quelconques. On a $\text{Int}(A) \subset A$ et $\text{Int}(B) \subset B$. Donc $(\text{Int}(A) \cap \text{Int}(B)) \subset (A \cap B)$. Comme il s'agit d'un ouvert comme intersection de deux ouverts, on a $(\text{Int}(A) \cap \text{Int}(B)) \subset \text{Int}(A \cap B)$. Par ailleurs, $(A \cap B) \subset A$ donc $\text{Int}(A \cap B) \subset \text{Int}(A)$ et de même $\text{Int}(A \cap B) \subset \text{Int}(B)$. Donc $\text{Int}(A \cap B) \subset (\text{Int}(A) \cap \text{Int}(B))$. On vient de montrer que $\text{Int}(A \cap B) = \text{Int}(A) \cap \text{Int}(B)$.

Si $x \in \text{Int}(A) \cup \text{Int}(B)$, soit $x \in \text{Int}(A) \subset A$, soit $x \in \text{Int}(B) \subset B$. Donc $(\text{Int}(A) \cup \text{Int}(B)) \subset (A \cup B)$. Comme $\text{Int}(A) \cup \text{Int}(B)$ est un ouvert comme union d'ouverts, on a forcément $\text{Int}(A) \cup \text{Int}(B) \subset \text{Int}(A \cup B)$.

Nous laissons la démonstration des propriétés de l'adhérence en exercice au lecteur (ou au passage par complémentaire). Le passage à un nombre fini d'ensembles se fait par récurrence sur le mode $\text{Int}(A \cap B \cap C) = \text{Int}(A \cap B) \cap \text{Int}(C) = \text{Int}(A) \cap \text{Int}(B) \cap \text{Int}(C)$. \square

Exemples :

- Comme \emptyset et E sont tous les deux à la fois ouverts et fermés, ils sont leur propres intérieurs et adhérences.
- On a vu qu'un singleton $\{x\}$ est fermé, son adhérence est donc lui-même. Si E

n'est pas l'espace vectoriel trivial réduit à $\{0\}$, il existe au moins un point $y \neq 0$. La suite $u_n = x + 2^{-n}y$ vérifie $u_n \neq x$ et $\|u_n - x\| = \|2^{-n}y\| = 2^{-n}\|y\| \rightarrow 0$. Donc $(u_n) \subset \{x\}^C$ mais (u_n) converge vers $x \notin \{x\}^C$. Donc le complémentaire de $\{x\}$ n'est pas fermé et $\{x\}$ n'est pas ouvert. Donc $\text{Int}(\{x\}) \neq \{x\}$. Comme l'intérieur est inclus dans l'ensemble, il ne reste plus que le choix $\text{Int}(\{x\}) = \emptyset$.

- Soit $A = \{x_1, x_2, \dots\}$ un nombre dénombrables de points. Si on n'est pas dans le cas trivial $E = \{0\}$, toute boule $B(x_i, r)$ est de cardinal indénombrable (voir les compléments à la fin du chapitre). Ceci montre que A ne peut contenir aucune boule et donc que $\overset{\circ}{A} = \emptyset$. Si $A = \{x_1, \dots, x_p\}$ est un ensemble fini de points, il est fermé et $\overline{A} = A$. Mais si A n'est pas fini, on ne peut rien dire de général : $A = \mathbb{N}$ est un fermé de \mathbb{R} mais $A = \{2^{-n}\}$ n'est pas un fermé de \mathbb{R} .
- Soit $B(x, r)$ une boule ouverte de E . Comme c'est un ouvert, son intérieur est elle-même. Montrons que son adhérence est la boule fermée $\overline{B(x, r)}$. Cette dernière est un fermé qui contient $B(x, r)$, donc on a $\overline{B(x, r)} \subset \overline{B(x, r)}$. Il suffit de montrer que tous les points de $\overline{B(x, r)}$ sont dans l'adhérence, par exemple en montrant qu'ils sont limites de points de la boule ouverte. Soit x^* un point de $\overline{B(x, r)}$ qui n'est pas dans la boule $B(x, r)$ (sinon c'est déjà fini). On a forcément $\|x^* - x\| = r$. Soit $x_n = 2^{-n}x + (1 - 2^{-n})x^*$. On a $\|x_n - x\| = (1 - 2^{-n})\|x^* - x\| = (1 - 2^{-n})r < r$ et donc (x_n) est bien une suite de la boule ouverte. Comme $\|x^* - x_n\| = \|2^{-n}(x - x^*)\| = 2^{-n}r$, la suite (x_n) tend bien vers x^* . D'après la Proposition 3.34, x^* est bien dans l'adhérence de la boule ouverte.
- Soit $\overline{B(x, r)}$ une boule fermée de E . Comme c'est un fermé, elle est sa propre adhérence. Montrons que son intérieur est la boule ouverte. Comme la boule ouverte est ouverte, on a forcément que $B(x, r) \subset \text{Int}(\overline{B(x, r)})$. Montrons qu'un point x^* avec $\|x^* - x\| = r$ ne peut être dans l'intérieur de la boule fermée. Prenons une boule $B(x^*, \varepsilon)$ avec $\varepsilon > 0$. Cette boule contient le point $y = x + (1 + \varepsilon/2r)(x^* - x)$ car $\|y - x^*\| = \varepsilon/2r\|x^* - x\| = \varepsilon/2$. Mais y n'est pas dans $\overline{B(x, r)}$ car $\|y - x\| = (1 + \varepsilon/2r)\|x^* - x\| = r + \varepsilon/2 \geq r$. D'après la Proposition 3.34, x^* ne peut être dans l'intérieur de la boule fermée.
- Les deux exemples précédents montre que la frontière d'une boule ouverte ou fermée est la sphère (cf figure 3.3.5).
- L'ensemble \mathbb{Q} est dense dans \mathbb{R} et donc son adhérence est \mathbb{R} . Mais comme \mathbb{Q} est dénombrable, l'argument plus haut montre que son intérieur est vide. La frontière de \mathbb{Q} est donc \mathbb{R} . On obtient au passage un exemple montrant que l'intérieur de l'adhérence n'est pas l'intérieur de l'ensemble en général (il ne s'agit pas d'opérations réciproques ou même qui pourraient commuter).
- Voici les contre-exemples pour les propriétés manquantes de la proposition 3.35 pour deux ensembles. Soient $A = [0, 1]$ et $B = [1, 2]$. Comme $\text{Int}([a, b]) =]a, b[$ (les intervalles sont des boules de \mathbb{R}), on a $\text{Int}(A) \cup \text{Int}(B) =]0, 1[\cup]1, 2[\neq \text{Int}(A \cup B) =]0, 2[$. De façon similaire, soit $A =]0, 1[$ et $B =]1, 2[$. Comme $\text{Adh}(]a, b]) = [a, b]$, on a $\text{Adh}(A \cap B) = \text{Adh}(\emptyset) = \emptyset$ et $\text{Adh}(A) \cap \text{Adh}(B) = \{1\}$.

- Les égalités de la proposition 3.35 ne passent pas aux unions ou intersections infinies. Par exemple sur \mathbb{R} , si on considère l'intersection $\{0\} = \bigcap_n]-1/n, 1/n[$, alors l'intérieur de $\{0\}$ est vide mais l'intersection des intérieurs de $] -1/n, 1/n[$ est encore $\{0\}$. Inversement, si on regarde l'union $[0, 1[= \bigcup_n [0, 1 - 1/n[$, son adhérence est $[0, 1]$ alors que l'union des adhérences des $[0, 1 - 1/n[$ reste $[0, 1[$.

On peut facilement généraliser aux espaces vectoriels normés les deux définitions suivantes qui étaient connues sur \mathbb{R} .

Définition 3.36

Un ensemble A d'un espace vectoriel normé E est **borné** si

$$\exists M > 0, \forall a \in A, \|a\| \leq M.$$

Définition 3.37

Un ensemble $A \subset E$ est dit **dense** dans E si $\bar{A} = E$.

On obtient facilement les caractérisations équivalentes suivantes.

Proposition 3.38

Un ensemble $A \subset E$ est borné si et seulement s'il est inclus dans une boule ouverte ou fermée de E .

Démonstration : Si A est borné avec la borne $M > 0$, il est inclus dans la boule ouverte $B(0, M)$ et dans la boule fermée $\bar{B}(0, M)$. Soit A inclus dans la boule ouverte (ou fermée) $B(x, r)$. Pour tout $a \in A$, on a $\|a\| = \|(a-x) + x\| \leq \|a-x\| + \|x\| \leq r + \|x\|$. Donc l'ensemble A est borné avec la borne $M = r + \|x\|$ (ou une borne plus grande). \square

Proposition 3.39

Soit $A \subset E$. Les trois propriétés suivantes sont équivalentes :

- A est dense dans E
- tout $x \in E$ est limite d'une suite $(a_n) \subset A$ de points de A .
- toute boule ouverte $B(x, r)$ de E contient un élément de A .

Démonstration : Il s'agit d'appliquer les caractérisations de l'adhérence de la proposition 3.34. \square

Exemples :

- Montrons que \mathbb{Q}^d est dense dans \mathbb{R}^d muni de la norme 1. Soit $x = (x_1, \dots, x_d) \in \mathbb{R}^d$ et $r > 0$. Par densité de \mathbb{Q} dans \mathbb{R} (chapitre précédent), il existe d rationnels

$q_i \in \mathbb{Q}$ tels que $q_i \in]x_i - \frac{r}{d}, x_i + \frac{r}{d}[$. On pose $q = (q_1, \dots, q_d)$ qui est dans \mathbb{Q}^d . On a

$$\|x - q\|_1 = \sum_{i=1}^d |x_i - q_i| < \sum_{i=1}^d \frac{r}{d} = r .$$

Donc q est dans la boule $B(x, r)$, ce qui montre que \mathbb{Q}^d est dense dans \mathbb{R}^d .

- Les polynômes $\mathbb{R}[X]$ sont denses dans l'espace des fonctions $\mathcal{C}^0([0,1], \mathbb{R})$ muni de la norme infinie, voir le théorème de Weierstrass (théorème 2.42).
- Les polynômes trigonométriques sont denses dans l'espace des fonctions $\mathcal{C}^0([0,1], \mathbb{R})$ muni de la norme 2 (voir les théorèmes de convergence des séries de Fourier).

3 Suites d'un espace vectoriel normé

On a déjà vu plus haut la définition de convergence des suites au sens de la norme, qui est pour rappel comme suit.

Définition 3.40

Une suite (x_n) d'un espace vectoriel normé E **converge vers** $\ell \in E$ si

$$\|x_n - \ell\| \xrightarrow{n \rightarrow +\infty} 0 .$$

Si on imagine bien la même définition quand on n'a pas de norme mais seulement une distance, c'est moins évident si on a un espace topologique sans notion de distance. Dans ce cas, on utilise comme définition la caractérisation suivante qui n'utilise que les ouverts.

Proposition 3.41

Soit $(E, \|\cdot\|)$ un espace vectoriel normé. Une suite (x_n) de E converge vers $\ell \in E$ si et seulement si pour tout $r > 0$, il existe un rang $N \in \mathbb{N}$ tel que $u_n \in B(\ell, r)$ pour tout $n \geq N$.

De manière plus générale, une suite (x_n) de E converge vers $\ell \in E$ si et seulement si pour tout ouvert \mathcal{O} de E contenant ℓ , il existe un rang $N \in \mathbb{N}$ tel que $u_n \in \mathcal{O}$ pour tout $n \geq N$.

Démonstration : La caractérisation avec les boules ouvertes est juste une traduction géométrique de

$$\forall \varepsilon > 0 , \exists N \in \mathbb{N} , \forall n \geq N , \|x_n - \ell\| < \varepsilon$$

(où $r > 0$ remplace $\varepsilon > 0$ comme variable muette). Si (x_n) entre à partir d'un certain rang dans tout ouvert contenant ℓ , alors elle entre a fortiori dans toute

boule ouverte contenant ℓ . Pour montrer la réciproque, supposons que la suite entre dans toute boule ouverte contenant ℓ . Soit \mathcal{O} un ouvert contenant ℓ . Par définition des ouverts, il existe une boule ouverte $B(\ell, r) \subset \mathcal{O}$. Comme (u_n) finit dans $B(\ell, r)$, elle finit bien par entrer dans \mathcal{O} . \square

Exemple :

On retrouve la caractérisation séquentielle des fermés. En effet, si F est fermé, alors F^C est ouvert. Si $\ell \in F^C$ est la limite d'une suite (x_n) , alors la suite (x_n) doit finir dans l'ouvert F^C . Par contraposée, si ℓ est la limite d'une suite $(x_n) \subset F$, alors $\ell \in F$.

Dans les espaces munis d'une distance, on peut parler de la limite d'une suite.

Proposition 3.42

Soit $(E, \|\cdot\|)$ un espace vectoriel normé et soient ℓ et ℓ' deux limites d'une même suite (x_n) . Alors $\ell = \ell'$.

Démonstration : On a

$$\|\ell - \ell'\| = \|(\ell - x_n) - (\ell' - x_n)\| \leq \|\ell - x_n\| + \|\ell' - x_n\| \xrightarrow{n \rightarrow +\infty} 0.$$

Donc $\|\ell - \ell'\| = 0$ et $\ell = \ell'$ par propriété de séparation. \square

La propriété suivante peut s'interpréter comme la continuité de la norme pour sa propre topologie.

Proposition 3.43

Soit (x_n) une suite convergente d'un espace vectoriel normé E . Alors

$$\lim_{n \rightarrow +\infty} \|x_n\| = \left\| \lim_{n \rightarrow +\infty} x_n \right\|.$$

Démonstration : Soit ℓ la limite de la suite. La variante minorante de l'inégalité triangulaire nous dit que $\|x_n - \ell\| \geq \left| \|x_n\| - \|\ell\| \right|$. Comme $\|x_n - \ell\|$ tend vers 0, on a que $\|x_n\| - \|\ell\| \rightarrow 0$ par théorème des gendarmes sur \mathbb{R} . \square

Définition 3.44

Une suite $(x_n) \subset E$ est dite **bornée** si $\{x_n\}$ est borné, c'est-à-dire s'il existe M tel que $\|x_n\| \leq M$ pour tout n .

Proposition 3.45

Une suite convergente d'un espace vectoriel normé est bornée.

Démonstration : La norme de la suite converge vers la norme de la limite et les suites réelles convergentes sont bornées. \square

Les résultats sur les manipulations des limites de suites ne cachent pas de surprises. N'oublions pas ici que pour un espace vectoriel, les deux seules manipulations sont les sommes et les homothéties.

Proposition 3.46

Soit $(E, \|\cdot\|)$ un espace vectoriel normé. Soient (u_n) et (v_n) deux suites de E qui convergent vers u_∞ et v_∞ . Alors la suite $(u_n + v_n)$ converge vers la somme des limites $u_\infty + v_\infty$.

Si de plus (λ_n) est une suite de scalaires convergeant vers $\lambda_\infty \in \mathbb{R}$, alors $(\lambda_n u_n)$ est une suite convergente vers $\lambda_\infty u_\infty$.

En corollaire, la limite d'une combinaison linéaire (finie) est la combinaison linéaire des limites.

Démonstration : La première propriété vient de l'inégalité triangulaire

$$\|(u_n + v_n) - (u_\infty + v_\infty)\| = \|(u_n - u_\infty) - (v_n - v_\infty)\| \leq \|u_n - u_\infty\| + \|v_n - v_\infty\| \rightarrow 0.$$

Pour la deuxième propriété, nous allons utiliser que les suites de \mathbb{R} et de E qui convergent sont bornées. Soit M tel que $\|u_n\| \leq M$ et $|\lambda_n| \leq M$ pour tout n . On a

$$\begin{aligned} \|\lambda_n u_n - \lambda_\infty u_\infty\| &= \|\lambda_n(u_n - u_\infty) + (\lambda_n - \lambda_\infty)u_\infty\| \\ &\leq \|\lambda_n(u_n - u_\infty)\| + \|(\lambda_n - \lambda_\infty)u_\infty\| \\ &\leq |\lambda_n| \|u_n - u_\infty\| + \|u_n - u_\infty\| \|u_\infty\| \\ &\leq M(\|u_n - u_\infty\| + \|u_n - u_\infty\|) \rightarrow 0. \end{aligned}$$

Finalement, ces propriétés peuvent se combiner un nombre fini de fois (par récurrence par exemple) pour obtenir les limites de combinaisons linéaires. \square

La proposition suivante peut se démontrer par des estimations directes mais nous pouvons utiliser aussi les propriétés ci-dessus pour aller plus vite.

Proposition 3.47

Dans \mathbb{R}^d muni d'une des normes $\|\cdot\|_p$, avec $p \in [1, +\infty]$, la convergence d'une suite est équivalente à la convergence coordonnée par coordonnée.

Démonstration : Nous avons déjà vu ce résultat pour la norme $\|\cdot\|_\infty$, cf l'exemple sous la définition 3.29. Soit $p \in [1, +\infty[$. On peut décomposer chaque x_n sur la base canonique

$$x_n = x_n^1 \cdot (1, 0, \dots, 0) + x_n^2 \cdot (0, 1, \dots, 0) + \dots + x_n^d \cdot (0, 0, \dots, 1) .$$

La proposition précédente nous montre que si, pour tout i , la suite $(x_n^i)_n$ converge, alors la suite x_n converge aussi comme combinaison linéaire de suites convergentes. Si maintenant la suite (x_n) converge vers une limite ℓ , on a

$$\forall i \in \llbracket 1, d \rrbracket , \quad |x_n^i - \ell^i| \leq \left(\sum_{j=1}^d |x_n^j - \ell^j|^p \right)^{1/p} = \|x_n - \ell\|_p \rightarrow 0 .$$

Ceci montre que chaque coordonnée $(x_n^i)_n$ converge vers ℓ^i . □

On peut proposer des généralisations des définitions concernant les suites réelles.

Définition 3.48

Soit $(u_n)_{n \in \mathbb{N}}$ une suite d'un espace vectoriel normé $(E, \|\cdot\|)$. Une fonction monotone $\varphi : \mathbb{N} \rightarrow \mathbb{N}$ est appelée une extraction et la suite $(u_{\varphi(n)})_{n \in \mathbb{N}}$ est appelée une **suite extraite** ou une **sous-suite** de (u_n) .

Un point $x \in E$ est appelé **valeur d'adhérence** de la suite (u_n) s'il est limite d'une sous-suite de (u_n) .

Par des constructions astucieuses, on peut montrer que l'ensemble des valeurs d'adhérence peut être n'importe quoi, du moment qu'il s'agit d'un fermé. Cette contrainte vient du résultat suivant.

Proposition 3.49

L'ensemble des valeurs d'adhérence d'une suite (u_n) est

$$\bigcap_{N \in \mathbb{N}} \overline{\bigcup_{n \geq N} \{u_n\}} .$$

C'est donc un fermé de E .

Démonstration : Une fois l'expression montrée, on aura que l'ensemble des valeurs d'adhérence est fermé comme intersections de fermés.

Si ℓ est une valeur d'adhérence de (u_n) , il existe une sous-suite $(u_{\varphi(n)})$ qui converge vers ℓ . Pour tout $N \in \mathbb{N}$, la suite $(u_{\varphi(n)})_{n \geq N}$ est une suite d'éléments de $\bigcup_{n \geq N} \{u_n\}$ qui converge vers ℓ qui est donc dans $\overline{\bigcup_{n \geq N} \{u_n\}}$. Comme cela est vrai pour tout n , ℓ est bien dans $\bigcap_{N \in \mathbb{N}} \overline{\bigcup_{n \geq N} \{u_n\}}$.

Soit ℓ un point de $\bigcap_{N \in \mathbb{N}} \overline{\bigcup_{n \geq N} \{u_n\}}$. On construit l'extraction suivante. On pose $\varphi(0) = 0$. Comme $\ell \in \overline{\bigcup_{n \geq 1} \{u_n\}}$, la boule $B(\ell, 1)$ rencontre $\bigcup_{n \geq 1} \{u_n\}$ et

il existe un rang $n_1 =: \varphi(1)$ tel que $\|u_{n_1} - \ell\| \leq 1$. On repart de ce nouveau rang n_1 : comme $\ell \in \overline{\bigcup_{n \geq n_1+1} \{u_n\}}$, la boule $B(\ell, 1/2)$ rencontre $\bigcup_{n > n_1} \{u_n\}$ et il existe un rang $n_2 =: \varphi(2) > n_1$ tel que $\|u_{n_2} - \ell\| \leq 1/2$. Puis de nouveau, comme $\ell \in \overline{\bigcup_{n \geq n_2+1} \{u_n\}}$, la boule $B(\ell, 1/3)$ rencontre $\bigcup_{n > n_2} \{u_n\}$ et il existe un rang $n_3 =: \varphi(3) > n_2$ tel que $\|u_{n_3} - \ell\| \leq 1/3$. . . On construit ainsi une suite extraire $(u_{\varphi(n)})$ qui converge vers ℓ . Donc ℓ est bien une valeur d'adhérence. \square

Exemples :

- Dans \mathbb{R} , la suite $u_n = (-1)^n$ a deux valeurs d'adhérence ± 1 . Une démonstration rapide utilise le résultat précédent puisque $\bigcup_{n \geq N} \{u_n\}$ vaut toujours $\{-1, 1\}$ qui est un fermé comme union de deux singletons.
- Dans \mathbb{R}^2 , la suite $x_n = ((-1)^n, (-1)^n)$ a pour valeurs d'adhérence $\{(1, 1), (-1, -1)\}$, alors que la suite $x_n = ((-1)^n, (-1)^{n+1})$ a pour valeurs d'adhérence $\{(1, -1), (-1, 1)\}$.
- Soit (q_n) une énumération des rationnels d'un intervalle $[a, b]$ (resp. de \mathbb{R}). Cette suite existe puisque \mathbb{Q} est dénombrable comme toutes ses sous-parties. On peut montrer que les valeurs d'adhérence de (q_n) sont tout l'intervalle $[a, b]$ (resp. tout \mathbb{R}). Comme le produit de deux ensembles dénombrables est dénombrables, on peut aussi énumérer les rationnels du carré $[0, 1]^2$ et obtenir une suite dans les valeurs d'adhérence forme tout le carré.

Nous verrons dans un chapitre suivant la notion de « compacité » qui généralisera le théorème de Bolzano-Weierstrass (cf théorème 2.25). Mais nous allons voir déjà une première généralisation dans le cas particulier de \mathbb{R}^d muni d'une norme usuelle.

Théorème 3.50 (Bolzano-Weierstrass pour les normes p)

Dans \mathbb{R}^d muni d'une des normes p ($p \in [1, +\infty]$), toute suite bornée admet une sous-suite convergente.

Démonstration : Soit (x_n) une suite bornée : il existe M tel que $\|x_n\|_p \leq M$ pour tout n . On décompose x_n sur ses coordonnées : $x_n = (x_n^1, \dots, x_n^d)$. Le caractère explicite des normes p nous montre que chaque suite de composantes (x_n^i) de (x_n) est aussi bornée par M dans \mathbb{R} . D'après le théorème de Bolzano-Weierstrass sur \mathbb{R} (cf corollaire 2.26), on peut extraire une sous-suite $(x_{\varphi_1(n)}^1)$ de la première coordonnées qui converge vers $x_*^1 \in \mathbb{R}$. On regarde maintenant la suite $(x_{\varphi_1(n)}^2)$ qui est la suite des deuxièmes composantes des termes qu'on a gardés. C'est toujours une suite réelle bornée, donc on peut en extraire une sous-suite $(x_{\varphi_1 \circ \varphi_2(n)}^2)$ qui converge vers $x_*^2 \in \mathbb{R}$. Notons que la sous-suite $(x_{\varphi_1 \circ \varphi_2(n)}^1)$ étant une sous-suite de $(x_{\varphi_1(n)}^1)$ converge toujours vers x_*^1 . On continue ainsi de suite : de la suite $(x_{\varphi_1 \circ \varphi_2(n)}^3)$ on peut extraire une sous-suite $(x_{\varphi_1 \circ \varphi_2 \circ \varphi_3(n)}^3)$ convergente etc. À chaque étape, on retire de nouveaux termes de la suite pour permettre à la composante i de converger. Comme on n'a qu'un

nombre fini de composantes, il reste encore une infinité de termes après les d extractions et pour chaque i , la suite $(x_{\varphi_1 \circ \dots \circ \varphi_d(n)}^i)$ converge vers un réel (x_*^i) . Nous avons déjà remarqué que la convergence dans \mathbb{R}^d muni d'une norme p est équivalente à la convergence composante par composante. Donc la suite extraite $(x_{\varphi_1 \circ \dots \circ \varphi_d(n)})$ converge vers $x_* = (x_*^1, \dots, x_*^d)$ dans \mathbb{R}^d . \square

4 Équivalence de normes

Nous avons vu plusieurs exemples de normes. Elles mesurent des distances différentes, mais on peut se poser la question de savoir si elles induisent des topologies différentes ou non, c'est-à-dire si la notion d'ouvert ou de suite convergente sera différente ou non selon la norme. Nous avons vu plus haut que la notion de convergence dans \mathbb{R}^d ne dépend pas du choix de la norme p et donc toutes les normes p de \mathbb{R}^d induisent la même notion de fermés et donc d'ouverts. On aimerait généraliser ce type de comparaison entre normes.

Définition 3.51

Soit E un espace vectoriel et soient $\|\cdot\|_*$ et $\|\cdot\|_{\#}$ deux normes sur E . On dit que $\|\cdot\|_*$ **domine** $\|\cdot\|_{\#}$ s'il existe une constante $M > 0$ telle que

$$\forall x \in E, \quad \|x\|_{\#} \leq M\|x\|_* .$$

On dit que les normes $\|\cdot\|_*$ et $\|\cdot\|_{\#}$ sont **équivalentes** sur E s'il existe des constantes m et M strictement positives telles que

$$\forall x \in E, \quad m\|x\|_* \leq \|x\|_{\#} \leq M\|x\|_* .$$

Exemples :

- Sur \mathbb{R}^d , les normes 1 et ∞ sont équivalentes. En effet, comme le maximum des $|x_i|$ apparaît au moins une fois dans la norme, on a

$$\|x\|_1 = \sum_{i=1}^d |x_i| \geq \max_{i \in \llbracket 1, d \rrbracket} |x_i| = \|x\|_{\infty} .$$

Et en majorant chaque $|x_i|$ par leur maximum, on obtient

$$\|x\|_1 = \sum_{i=1}^d |x_i| \leq \sum_{i=1}^d \|x\|_{\infty} = d\|x\|_{\infty} .$$

- Les normes 1 et ∞ ne sont pas équivalentes sur l'espace de fonctions continues $\mathcal{C}^0([0,1], \mathbb{R})$. En effet, on considère la suite de fonctions $f_n : x \mapsto x^n$. On a

$$\|f_n\|_{\infty} = \max_{x \in [0,1]} |x|^n = 1 \quad \text{et} \quad \|f_n\|_1 = \int_0^1 |x^n| dx = \left[\frac{1}{n+1} x^{n+1} \right]_0^1 = \frac{1}{n+1} .$$

S'il existait une constante M telle que $\|f_n\|_\infty \leq M\|f_n\|_1$, le fait que $\|f_n\|_1 \rightarrow 0$ impliquerait que $\|f_n\|_\infty \rightarrow 0$, ce qui n'est pas le cas.

Par contre, la norme ∞ domine la norme 1 car pour tout $f \in \mathcal{C}^0([0,1],\mathbb{R})$, le calcul similaire au cas \mathbb{R}^d donne

$$\|f_n\|_1 = \int_0^1 |x^n| dx \leq \int_0^1 \|f\|_\infty dx = \|f\|_\infty \int_0^1 dx = \|f\|_\infty .$$

Il est utile d'avoir une vision géométrique de la domination d'une norme... mais attention au sens de l'inclusion! Comme on utilise plusieurs normes sur le même espace vectoriel, il est important de bien distinguer avec quelle norme on mesure une distance, on définit une boule etc. Nous utiliserons de façon naturelle une notation mettant la norme en indice.

Proposition 3.52

Soit E un espace vectoriel et soient $\|\cdot\|_*$ et $\|\cdot\|_\#$ deux normes sur E . La domination

$$\forall x \in E, \quad \|x\|_\# \leq M\|x\|_*$$

est équivalente à l'inclusion des boules

$$B_{\|\cdot\|_*}(0,1) \subset B_{\|\cdot\|_\#}(0,M)$$

Démonstration : Supposons que $\|x\|_\# \leq M\|x\|_*$ pour tout x . Soit $x \in B_{\|\cdot\|_*}(0,1)$, on a par définition $\|x\|_* < 1$. La domination implique que $\|x\|_\# < M$ et donc que $x \in B_{\|\cdot\|_\#}(0,M)$.

Supposons maintenant que $B_{\|\cdot\|_*}(0,1) \subset B_{\|\cdot\|_\#}(0,M)$. Soit $x \in E$. Si $x = 0$, l'estimation est triviale. Sinon, on normalise $\|x\|$ en posant $\tilde{x}_n = \frac{1-2^n}{\|x\|_*}x$. On a alors

$$\|\tilde{x}_n\|_* = \left\| \frac{1-2^n}{\|x\|_*}x \right\|_* = \frac{1-2^n}{\|x\|_*} \|x\|_* = 1 - 2^{-n} .$$

Notons qu'on a pris exactement ce qu'il fallait pour avoir \tilde{x}_n dans la boule ouverte $B_{\|\cdot\|_*}(0,1)$ (en général, on normalise en un vecteur de taille 1, mais on a ici des boules ouvertes donc ce n'est pas exactement convenable). Par hypothèse d'inclusion, \tilde{x}_n appartient aussi à $B_{\|\cdot\|_\#}(0,M)$. Donc pour tout n on a $\|\tilde{x}_n\|_\# < M$ et donc $\frac{1-2^n}{\|x\|_*} \|x\|_\# < M$. En faisant tendre n vers l'infini et en multipliant par $\|x\|_*$, on trouve bien $\|x\|_\# \leq M\|x\|_*$. \square

Exemple :

On se place dans le plan. On vérifie facilement géométriquement les inclusions

$$B_{\|\cdot\|_1}(0,1) \subset B_{\|\cdot\|_\infty}(0,1)$$

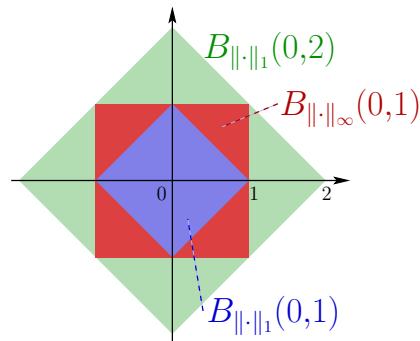
et

$$B_{\|\cdot\|_\infty}(0,1) \subset B_{\|\cdot\|_1}(0,2).$$

La proposition 52 montre donc que

$$\forall x \in \mathbb{R}^2, \quad \|x\|_\infty \leq \|x\|_1 \leq 2\|x\|_\infty.$$

On retrouve géométriquement ce que l'on avait obtenu par le calcul ci-dessus.



L'intérêt central de cette notion d'équivalence de normes, c'est de voir quand deux normes induisent la même topologie.

Proposition 3.53

Soit E un espace vectoriel et soient $\|\cdot\|_*$ et $\|\cdot\|_\#$ deux normes sur E . Les propriétés suivantes sont toutes équivalentes :

- (i) les normes $\|\cdot\|_*$ et $\|\cdot\|_\#$ sont équivalentes,
- (ii) un ensemble $\mathcal{O} \subset E$ est ouvert pour la topologie de $\|\cdot\|_*$ si et seulement s'il est ouvert pour la topologie de $\|\cdot\|_\#$,
- (iii) un ensemble $\mathcal{F} \subset E$ est fermé pour la topologie de $\|\cdot\|_*$ si et seulement s'il est fermé pour la topologie de $\|\cdot\|_\#$,
- (iv) un ensemble $\mathcal{B} \subset E$ est borné pour la topologie de $\|\cdot\|_*$ si et seulement s'il est borné pour la topologie de $\|\cdot\|_\#$,
- (v) une suite $(x_n) \subset E$ converge vers $x \in E$ pour la topologie de $\|\cdot\|_*$ si et seulement si elle une suite $(x_n) \subset E$ converge vers x pour la topologie de $\|\cdot\|_\#$.

Démonstration : Il est clair que (ii) et (iii) sont équivalentes en passant au complémentaire $\mathcal{F} = \mathcal{O}^c$.

Supposons que (v) soit vraie. On utilise la caractérisation séquentielle des fermés. Si \mathcal{F} est un fermé pour la norme $\|\cdot\|_*$ et qu'on regarde une suite $(x_n) \subset \mathcal{F}$ convergeant vers $x \in E$ pour la norme $\|\cdot\|_\#$ alors (v) implique que la suite converge aussi pour la norme $\|\cdot\|_*$ et par fermeture que $x \in \mathcal{F}$. Ceci montre qu'un fermé pour $\|\cdot\|_*$ est aussi un fermé pour $\|\cdot\|_\#$. Par symétrie, l'autre implication est aussi vraie et donc (v) implique (iii).

Supposons que (ii) soit vraie. On utilise la caractérisation de la convergence par les ouverts. Supposons que (x_n) tend vers x pour $\|\cdot\|_*$ et que x est dans \mathcal{O} ouvert pour $\|\cdot\|_\#$. Alors \mathcal{O} est encore un ouvert pour $\|\cdot\|_*$ et donc x_n doit

appartenir à \mathcal{O} pour n assez grand. Ceci montre que (x_n) converge aussi pour $\|\cdot\|_{\#}$. Par symétrie, l'autre implication est aussi vraie et donc (ii) implique (v). Au total nous avons pour le moment que (ii), (iii) et (v) sont équivalentes. Si (i) est vraie et que (x_n) tend vers $x \in E$ pour la norme $\|\cdot\|_*$, alors par définition $\|x_n - x\|_* \rightarrow 0$. Mais comme $\|\cdot\|_*$ domine $\|\cdot\|_{\#}$, on doit avoir par théorème des gendarmes $\|x_n - x\|_{\#} \rightarrow 0$. Donc (x_n) tend aussi vers $x \in E$ pour la norme $\|\cdot\|_{\#}$. Encore une fois, l'autre cas est symétrique et donc (i) implique (v).

Si (ii) est vraie alors la boule ouverte $B_{\|\cdot\|_*}(0,1)$ doit être un ouvert pour la norme $\|\cdot\|_{\#}$. Donc il existe $r > 0$ tel que la boule $B_{\|\cdot\|_{\#}}(0,r)$ soit incluse dans $B_{\|\cdot\|_*}(0,1)$. Par propriété d'homothétie des normes (cf Proposition 3.6), on a alors $B_{\|\cdot\|_{\#}}(0,1) \subset B_{\|\cdot\|_*}(0,1/r)$. La proposition 3.52 conclut que $\|\cdot\|_{\#}$ domine $\|\cdot\|_*$. L'autre domination se montre symétriquement et donc (ii) implique (i). Il ne reste plus qu'à s'occuper de (iv). Si les deux normes sont équivalentes, il est clair qu'un ensemble borné pour l'un doit être borné pour l'autre par simple inégalité. Supposons que (iv) soit vraie. La boule $B_{\|\cdot\|_*}(0,1)$ qui est bornée pour $\|\cdot\|_*$ doit aussi être bornée pour $\|\cdot\|_{\#}$. Donc il existe M assez grand tel que $B_{\|\cdot\|_*}(0,1) \subset B_{\|\cdot\|_{\#}}(0,M)$ et la proposition 52 conclut que $\|\cdot\|_*$ domine $\|\cdot\|_{\#}$. Comme toujours, on obtient l'autre implication par symétrie et on a bien que (iv) implique (i). \square

Le résultat principal de cette partie est l'équivalence des normes en dimension finie. Ainsi, tous les résultats que nous avons déjà obtenus pour les normes p sur \mathbb{R}^d seront généralisés.

Théorème 3.54

Sur l'espace vectoriel \mathbb{R}^d , toutes les normes sont équivalentes.

Démonstration : Soit $\|\cdot\|$ une norme quelconque. Il suffit de montrer qu'elle est équivalente à la norme infini car la relation d'équivalence des normes est transitive. Soit (e_1, \dots, e_d) la base canonique de \mathbb{R}^d . Soit $x \in \mathbb{R}^d$ décomposé en $x = x_1e_1 + \dots + x_de_d$. Par inégalité triangulaire, on a

$$\|x\| = \|x_1e_1 + \dots + x_de_d\| \leq |x_1|\|e_1\| + \dots + |x_d|\|e_d\| \leq \max_i |x_i| \max_j \|e_j\| .$$

Si on pose $M = \max_j \|e_j\|$, on vient de montrer que $\|x\| \leq M\|x\|_{\infty}$.

Pour montrer l'autre domination, nous raisonnons par l'absurde. S'il n'existe pas de constante $M' > 0$ telle que $\|x\|_{\infty} \leq M'\|x\|$, c'est que pour tout $n \in \mathbb{N}^*$, il existe $x_n \in E$ tel que $\|x_n\|_{\infty} > n\|x_n\|$. Nous normalisons la suite en posant $\tilde{x}_n = x_n/\|x_n\|_{\infty}$. On a alors

$$\|\tilde{x}_n\|_{\infty} = \left\| \frac{1}{\|x_n\|_{\infty}} x_n \right\|_{\infty} = \frac{1}{\|x_n\|_{\infty}} \|x_n\|_{\infty} = 1$$

et pour tout $n \geq 1$,

$$\|\tilde{x}_n\| = \frac{1}{\|x_n\|_{\infty}} \|x_n\| < \frac{1}{n} .$$

Donc \tilde{x}_n tend vers 0 pour la norme $\|\cdot\|$. Concernant la norme infini, la suite est contenue dans la sphère unité. D'après le théorème 3.50, on peut en extraire une sous-suite $(\tilde{x}_{\varphi(n)})$ convergeant vers x_* . Par la proposition 3.43, la normalisation $\|\tilde{x}_n\|_\infty = 1$ implique que $\|x_*\|_\infty = 1$. Finalement, nous utilisons la première domination pour écrire

$$\|x_* - \tilde{x}_{\varphi(n)}\| \leq M \|x_* - \tilde{x}_{\varphi(n)}\|_\infty \xrightarrow{n \rightarrow +\infty} 0.$$

Donc $(\tilde{x}_{\varphi(n)})$ converge vers x_* pour la norme $\|\cdot\|$ et par unicité de la limite, comme on avait déjà montré la convergence vers 0, on a forcément $x_* = 0$. Or ceci contredit le fait que $\|x_*\|_\infty = 1$. On vient donc d'obtenir l'existence d'un entier N tel que $\|x\|_\infty \leq N\|x\|$ ce qui conclut l'équivalence des normes. \square

Ce théorème permet de généraliser à tout espace de dimension finie ce que nous avons vu pour les normes p .

Exemple :

En fin de partie 2, nous avons démontré que \mathbb{Q}^d est dense dans \mathbb{R}^d pour la norme 1. Comme la densité peut se caractériser en terme de convergence de suites, la proposition 3.53 et le théorème 3.54 nous donne directement que \mathbb{Q}^d est dense dans \mathbb{R}^d pour toutes les normes sur \mathbb{R}^d , en particulier les autres normes p avec $p \in [1, +\infty]$.

En résumé :

Dans un espace vectoriel E de dimension finie :

- Toutes les normes sont équivalentes et ce n'est pas utile de préciser dans quelle norme on considère une propriété topologique.
- La convergence des suites est équivalente à la convergence coordonnée par coordonnée dans une base.

Dans un espace vectoriel E de dimension infinie :

- Il existe des normes non équivalentes. Il faut toujours bien préciser avec quelle norme on travaille.

5 Compléments

5.1 Espaces de matrices

Une matrice A de $\mathcal{M}_d(\mathbb{R})$ est la donnée de d^2 nombres réels $(a_{ij})_{1 \leq i, j \leq d}$. L'espace vectoriel des matrices est donc \mathbb{R}^{d^2} , sur lequel on a ajouté une structure supplémentaire de multiplication. On peut munir $\mathcal{M}_d(\mathbb{R})$ de n'importe quelle norme car toutes sont équivalentes (dimension finie). On peut utiliser par exemple la norme

infinie

$$\|A\|_\infty = \sup_{1 \leq i, j \leq d} |a_{i,j}|.$$

Il peut être intéressant de voir la relation entre cette structure de norme et la multiplication.

Proposition 3.55

Si A et B sont des matrices de $\mathcal{M}_d(\mathbb{R})$, alors $\|AB\|_\infty \leq d\|A\|_\infty\|B\|_\infty$.

Démonstration : Il suffit de voir que les coefficients de AB sont de la forme $c_{ij} = \sum_{k=1}^d a_{ik}b_{kj}$. Une majoration brutale donne alors

$$\begin{aligned} \|AB\|_\infty &\leq \sup_{ij} \left| \sum_{k=1}^d a_{ik}b_{kj} \right| \leq \sup_{ij} \sum_{k=1}^d |a_{ik}| \cdot |b_{kj}| \\ &\leq \sup_{ij} \sum_{k=1}^d \|A\|_\infty \|B\|_\infty = d\|A\|_\infty \|B\|_\infty. \end{aligned}$$

□

Dans la proposition précédente, le facteur d est un peu gênant et peu esthétique. On aurait aimé avoir une *norme d'algèbre* pour laquelle la norme du produit est majorée par le produit des normes. Cela est possible avec une norme plus complexe que nous verrons au prochain chapitre. Une fois que l'on a muni notre espace d'une topologie, on peut montrer des résultats qui sont intéressants.

Proposition 3.56

L'ensemble des matrices inversibles est dense parmi les matrices.

Démonstration : Un calcul direct par développement montre que si A est une matrice, la fonction $\lambda \in \mathbb{R} \mapsto \chi_A(\lambda) = \det(A - \lambda Id)$ est un polynôme non nul en lambda (il s'agit du polynôme caractéristique qui commence par $(-\lambda)^d + \dots$). Par ailleurs, A est inversible si et seulement si $\det A \neq 0$. Soit A une matrice non inversible. Le polynôme χ_A a au plus d racines. Donc pour tout $\varepsilon > 0$, il existe $\lambda \in]0, \varepsilon[$ qui n'est pas une racine de χ_A . La matrice $B = A - \lambda Id$ est alors une matrice inversible et

$$\|B - A\|_\infty = \|\lambda Id\|_\infty = |\lambda| < \varepsilon.$$

Donc B est une matrice inversible dans la boule $B(A, \varepsilon)$. Ceci montre que toute boule contient au moins une matrice inversible et donc que l'ensemble des matrices inversibles est dense. □

5.2 Les boules sont des ensembles indénombrables

On montre ici un résultat qui peut être utile, typiquement pour montrer qu'un ensemble de cardinal dénombrable est d'intérieur vide. On suppose que $(E, \|\cdot\|)$ est un espace vectoriel normé non trivial, c'est-à-dire que E n'est pas réduit à $\{0\}$.

Proposition 3.57

Soit $x \in E$ et $r > 0$. La boule ouverte $B(x,r)$ contient un nombre indénombrable de points. C'est aussi a fortiori le cas de la boule fermée $\overline{B}(x,r)$.

Démonstration : Soit $y \neq 0$ un point de E . Le vecteur $\tilde{y} = \frac{1}{\|y\|}y$ est de norme 1. On considère le segment $I = \{x + \theta\tilde{y}, \theta \in [0,r[$ qui est par construction en bijection avec $[0,r[$ et est donc de cardinal indénombrable (on rappelle que tous les intervalles de \mathbb{R} sont de cardinal indénombrable). Pour tout $z \in I$, il existe $\theta \in [0,r[$ tel que $z = x + \theta\tilde{y}$. On a alors

$$\|x - z\| = \|x + \theta\tilde{y} - x\| = \theta\|\tilde{y}\| = \theta < r.$$

Donc I est inclus dans la boule $B(x,r)$ et cette dernière contient aussi un nombre indénombrable de points.

Le cas de la boule fermée découle directement puisque $B(x,r) \subset \overline{B}(x,r)$. \square

Corollaire 3.58

Si $A \subset E$ est dénombrable, alors A est d'intérieur vide.

Démonstration : Aucune boule $B(x,r)$ ne peut être incluse dans A car sinon on aurait un ensemble indénombrable inclus dans un ensemble dénombrable, ce qui est absurde. Par la caractérisation de l'intérieur de la proposition 3.34, cet intérieur est donc vide. \square

5.3 Les ouverts-fermés d'un espace vectoriel

Théorème 3.59

Dans un espace vectoriel normé E , les seuls sous-ensembles qui sont à la fois fermés et ouverts dans E sont l'ensemble vide et E tout entier.

Démonstration : Soit U un ouvert-fermé de E qui n'est pas \emptyset . Il existe donc un point $x \in U$. Soit $y \in E$ un autre point. On regarde le segment $[x,y] = \{z_\theta = \theta x + (1-\theta)y, \theta \in [0,1]\}$. Soit $A \subset [0,1]$ l'ensemble des paramètres θ tels que $z_\theta \in U$ pour tout $\theta' \in [0,\theta]$ (on regarde les segments $[x,z_\theta] \subset [x,y] \cap U$). L'ensemble A est non vide car $0 \in A$ et il est majoré par 1. Il admet donc une

borne supérieure $\theta = \sup A$. On pose $u_n = z_{(1-2^{-n})\theta}$. Comme

$$\|u_n - z_\theta\| = \|-2^{-n}\theta x + 2^{-n}\theta y\| = 2^{-n}\theta\|y - x\|,$$

la suite u_n tend vers z_θ tout en étant dans U par construction. Comme U est fermé, on a $z_\theta \in U$. Mais comme U est ouvert, il existe $r > 0$ tel que $B(z_\theta, r) \subset U$. Si $\theta < 1$, on aurait $z_{\theta'} \in B(z_\theta, r) \subset U$ pour tout $\theta' > \theta$ assez proche de θ . Cela contredirait la définition de z_θ . Donc $\theta = 1$ et $z_\theta = y \in U$. On vient de montrer que tout $y \in E$ est dans U , donc $U = E$. \square

5.4 Utilisation d'un lemme « à la Sperner »

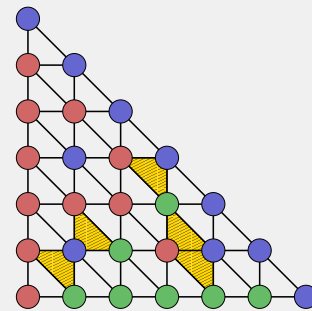
Le lemme de Sperner est un lemme sur les coloriages d'une triangulation. Son énoncé est simple et il existe plusieurs démonstrations. Mais ce résultat est bien plus profond qu'il en a l'air car il dépend en fait de la topologie du triangle. On verra ainsi certaines de ces conséquences dont l'énoncé dépend vraiment de la forme du domaine que l'on considère.

Nous n'allons pas montrer le lemme de Sperner dans toute sa généralité mais uniquement une version simplifiée.

Proposition 3.60

On considère un triangle rectangle isocèle découpé en petits triangles rectangles isocèles comme dans la figure ci-contre. On suppose que chaque sommet du découpage est colorié comme suit :

- sur le bord ouest, les sommets sont rouge
- sur le bord sud, les sommets sont verts
- sur le bord nord-est, les sommets sont bleus
- les sommets du grand triangle ont au choix une des deux couleurs des bords associés (rouge ou bleu pour le sommet nord etc.).
- les sommets à l'intérieur du triangle ont n'importe quelle des trois couleurs rouge, vert ou bleu.



Alors il existe au moins un petit triangle tricolore dont les trois sommets sont de couleurs différentes.

Démonstration : Comptons le nombre d'arêtes de petits triangles qui sont vert-rouge. Pour chaque triangle bicoloré, il y en a 0 ou 2 alors qu'il y en a exactement une pour un triangle tricolore. Si on dénombre ainsi les arêtes, on note que les arêtes internes sont comptées deux fois car elles sont dans deux petits triangles alors que les arêtes externes sont comptées une seule fois. On

obtient l'équation :

$$\begin{aligned}
 & 2 \times \text{nombre de triangles vert-vert-rouge} \\
 & + 2 \times \text{nombre de triangles vert-rouge-rouge} \\
 & \quad + \text{nombre de triangles tricolores} \\
 & = \\
 & 2 \times \text{nombre d'arêtes vert-rouge internes} \\
 & + \text{nombre d'arêtes vert-rouge externes}
 \end{aligned}$$

Or il y a exactement une arête vert-rouge externe dans le coin sud-ouest. Donc le nombre de petits triangles tricolores est impair. \square

Application au partage équitable d'une colocation :

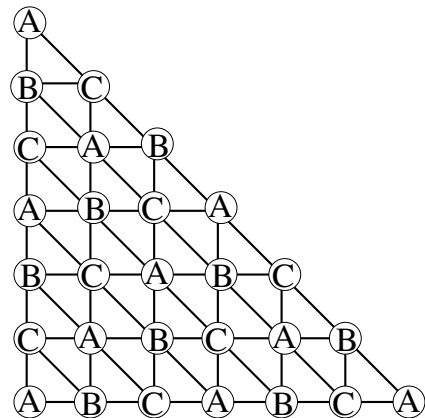
Trois colocataires doivent se répartir les 3 chambres d'un appartement. Celles-ci ne sont pas du tout semblables et chaque colocataire peut avoir des préférences différentes. Une façon de rendre le partage équitable est que le loyer soit partagé de façon à compenser les écarts de qualité des chambres. On note A , B et C les trois colocataires et on fait correspondre à chaque chambre une couleur rouge, vert ou bleu. Une répartition du loyer est la donnée de deux pourcentages : x , qui est le pourcentage de loyer attribué à la chambre rouge, et y qui est le pourcentage de loyer attribué à la chambre verte. Évidemment, la chambre bleue aura le loyer $100 - (x + y)$. Une répartition des loyers par chambre est donc un point du triangle rectangle isocèle

$$\{(x,y) \in \mathbb{R}^2, x \geq 0, y \geq 0 \text{ et } x + y \leq 100\}.$$

On suppose que le choix des colocataires suit les règles :

- Un colocataire préfère toujours une chambre gratuite à une payante,
- Si un colocataire préfère une chambre pour toutes les répartitions de loyer (x_n, y_n) et que (x_n, y_n) tend vers (x_*, y_*) , alors le colocataire préfère toujours la même chambre pour la répartition (x_*, y_*) (il peut préférer plusieurs chambres de manière égale).

On découpe le triangle représentant la répartition des loyers comme ci-contre : à chaque sommet on demande au colocataire marqué sur le sommet une chambre qu'il préfère pour cette répartition de loyers. On vérifie que ce coloriage respecte la règle de Sperner. Par exemple, si $x = 0$, alors la chambre rouge est gratuite et si $y \neq 0$ et $y \neq 100$ c'est la seule gratuite et donc la chambre rouge sera toujours choisie. La proposition 5.60 montre qu'il existe un triangle tricolore et on note (x_1, y_1) un de ses sommets. On remarque que le fait que le triangle est tricolore correspond à ce que chaque colocataire a choisi une chambre



sommet correspondant à une répartition différente). On recommence alors avec un découpage en triangles plus petits et on obtient un triangle tricolore dont un sommet est (x_2, y_2) etc. On a ainsi une suite de répartition des loyers pour des triangles tricolores de plus en plus petits.

Nous allons extraire de cette suite de sommets une sous-suite adéquate. Tout d'abord, par le théorème 3.50, nous pouvons supposer, quitte à extraire, que la suite (x_n, y_n) a une limite (x_*, y_*) . Ensuite, chaque (x_n, y_n) correspond à un triangle tricolore et une certaine répartition des chambres. Comme il n'y a que 6 répartitions possibles et une infinité de (x_n, y_n) , on peut aussi extraire une sous-suite qui correspondra toujours à la même répartition des chambres. Puis on utilise la règle de continuité des choix pour montrer que cette répartition des chambres, avec la répartition des loyers (x_*, y_*) , donne une répartition de la colocation telle qu'aucun des colocataires n'est jaloux de la chambre du voisin.

Chapitre 4 : Fonctions sur les espaces vectoriels normés

1 Définitions générales

Une fonction $f : E \longrightarrow F$ est de façon intuitive une boîte noire qui envoie tout $x \in E$ sur une image $f(x) \in F$. On peut donner de façon plus rigoureuse les définitions suivantes.

Définition 4.1

Soient E et F deux ensembles. Une **fonction** $f : E \longrightarrow F$ est la donnée d'un sous-ensemble $\mathcal{G}_f \subset E \times F$ appelé **graphe** de f tel que pour tout $x \in E$, il existe un unique couple (x, y) dans \mathcal{G}_f . On dit alors que y est l'image de x et on note $y = f(x)$.

La définition abstraite ci-dessus permet de dépasser l'idée qu'une fonction est une formule calculatoire. On voit aussi l'équivalence entre la donnée du graphe et celle de la fonction. Notons que le graphe est donc un sous-ensemble de $E \times F$ avec exactement un seul point « au-dessus » de chaque $x \in E$. Le graphe d'une fonction de \mathbb{R} dans \mathbb{R}^2 est donc une courbe uni-dimensionnelle dans \mathbb{R}^3 (enfin s'il y a plus ou moins de continuité, sinon c'est un ensemble de points très irrégulier). Le graphe d'une fonction de \mathbb{R}^2 dans \mathbb{R} est donc une surface bi-dimensionnelle dans \mathbb{R}^3 (avec la même restriction qu'il faut une certaine régularité pour avoir une vraie surface). Le graphe d'une fonction de \mathbb{R}^2 dans \mathbb{R}^2 est un ensemble de \mathbb{R}^4 qu'il sera plus difficile de visualiser etc.

Définition 4.2

Soit $f : E \longrightarrow F$ une fonction et $E' \subset E$ un sous-ensemble de E . L'**image** de E' par f est

$$f(E') := \{ y \in F, \exists x' \in E', f(x') = y \} .$$

Soit $F' \subset F$ un sous-ensemble de F . L'**image réciproque** de F' par f est

$$f^{-1}(F') := \{ x \in E, f(x) \in F' \} .$$



La notation $f^{-1}(F')$ ne signifie pas que f est bijective ni même injective. Il s'agit d'une opération sur les ensembles et si $F' = \{y'\}$ est un singleton, $f^{-1}(\{y'\})$ peut contenir une infinité de points. Et même si f est bijective avec $f(x) = y'$, remarquons que $f^{-1}(\{y'\})$ vaut $\{x\}$ et non pas x . De fait, les notations ci-dessus sont conventionnelles mais un peu abusives puisque formellement, on ne regarde pas la fonction f mais une fonction $\tilde{f} : \mathcal{P}(E) \rightarrow \mathcal{P}(F)$ qui envoie un sous-ensemble de E sur un sous-ensemble de F .

Exemple :

On considère la fonction $f : \theta \in \mathbb{R} \mapsto (\cos \theta, \sin \theta) \in \mathbb{R}^2$. L'image de f est exactement le cercle unité $\mathbb{S}^1 = \{(x, y) \in \mathbb{R}^2, x^2 + y^2 = 1\}$. Par ailleurs,

$$f^{-1}(\{(1,0)\}) = 2\pi\mathbb{Z} \quad \text{et} \quad f^{-1}(\{0\} \times \mathbb{R}) = \frac{\pi}{2} + \pi\mathbb{Z} .$$

On rappelle les définitions suivantes normalement très familières.

Définition 4.3

Soient E et F deux ensembles et soit $f : E \rightarrow F$ une fonction. On dit que f est **surjective** si $f(E) = F$. On dit que f est **injective** si $x \neq x'$ implique $f(x) \neq f(x')$. Si f est à la fois injective et surjective, on dit qu'elle est **bijective**.

Définition 4.4

Soient E et F deux ensembles et soit $f : E \rightarrow F$ une fonction. Soit $\tilde{E} \subset E$ un sous-ensemble de E . On dit que $\tilde{f} : \tilde{E} \rightarrow F$ est la **restriction** de f à \tilde{E} si on a $f(x) = \tilde{f}(x)$ pour tout $x \in \tilde{E}$. On note alors $\tilde{f} = f|_{\tilde{E}}$.

Dans ce cours, nous nous concentrons sur les espaces vectoriels. Soit $(X, \|\cdot\|_X)$ et $(Y, \|\cdot\|_Y)$ deux espaces vectoriels normés. Une fonction entre X et Y est une fonction du type

$$f : \mathcal{D}_f \subset X \rightarrow Y$$

où \mathcal{D}_f est un sous-ensemble de X . On aura souvent affaire à des fonctions définies sur tout X , mais ce n'est pas forcément toujours le cas, ne serait-ce qu'à cause des multiples fonctions réelles pas définies partout sur \mathbb{R} . Comme Y n'admet pas forcément de notion d'ordre, on ne peut plus parler de fonction majorée ou minorée, ni de fonction croissante etc. On peut par contre conserver une notion de borne.

Définition 4.5

Une fonction $f : \mathcal{D}_f \subset X \rightarrow Y$ définie entre deux espaces vectoriels normés est **bornée** si son image $f(\mathcal{D}_f)$ est un borné de Y , c'est-à-dire s'il existe $M \in \mathbb{R}$ tel que

$$\forall x \in \mathcal{D}_f, \quad \|f(x)\|_Y \leq M .$$

2 Limites

Nous allons généraliser la notion de limite que l'on connaît sur \mathbb{R} . Soit $(X, \|\cdot\|_X)$ et $(Y, \|\cdot\|_Y)$ deux espaces vectoriels normés et soit une fonction

$$f : \mathcal{D}_f \subset X \longrightarrow Y .$$

On peut chercher à obtenir la limite de la fonction en un point de \mathcal{D}_f mais aussi sur un point sur le bord du domaine de définition.

Définition 4.6

Soit f comme ci-dessus et soit $x_* \in \overline{\mathcal{D}_f}$. On dit que $f(x)$ **converge vers** $\ell \in Y$ ou **a pour limite** ℓ quand x tend vers x_* si

$$\forall \varepsilon > 0, \exists \delta > 0, \forall x \in \mathcal{D}_f, \|x - x_*\|_X < \delta \Rightarrow \|f(x) - \ell\|_Y < \varepsilon .$$

On note alors

$$\lim_{x \rightarrow x_*} f(x) = \ell \quad \text{ou} \quad f(x) \xrightarrow{x \rightarrow x_*} \ell .$$

Comme pour le cas réel, la définition est alourdie à cause du domaine de définition mais il s'agit juste d'être sûr que $f(x)$ a un sens. Et comme pour le cas réel, il existe plusieurs façon d'écrire la même notion de limite.

Proposition 4.7

Soit $f : \mathcal{D}_f \subset X \rightarrow Y$, soit $x_* \in \overline{\mathcal{D}_f}$ et soit $\ell \in Y$. Les propositions suivantes sont équivalentes.

- (i) $f(x)$ tend vers ℓ quand x tend vers x_* .
- (ii) Pour toute boule ouverte $B(\ell, \varepsilon)$, il existe une boule $B(x_*, \delta)$ telle que $f(B(x_*, \delta) \cap \mathcal{D}_f) \subset B(\ell, \varepsilon)$.
- (iii) Pour tout ouvert \mathcal{O} de Y contenant ℓ , il existe un ouvert \mathcal{U} de X contenant x_* tel que $f(\mathcal{U} \cap \mathcal{D}_f) \subset \mathcal{O}$.
- (iv) Pour toute suite $(x_n) \subset \mathcal{D}_f$ qui converge vers x_* dans X , la suite $f(x_n)$ converge dans Y vers ℓ .

Démonstration : Pour commencer par ce qui est évident, on se convainc facilement que (ii) n'est qu'une réécriture géométrique de (i) qui remplace les estimations de distances par l'appartenance à des boules.

Montrons maintenant que (ii) implique (iii). Soit \mathcal{O} un ouvert de Y contenant ℓ . Par définition d'un ouvert, il existe une petite boule $B(\ell, \varepsilon)$ incluse dans \mathcal{O} . (ii) fournit une boule $B(x_*, \delta)$ telle que $f(B(x_*, \delta) \cap \mathcal{D}_f) \subset B(\ell, \varepsilon)$ et $B(x_*, \delta) \cap \mathcal{D}_f$ est un ouvert du domaine comme recherché.

Montrons que (iii) implique (iv). Soit $(x_n) \subset \mathcal{D}_f$ une suite qui converge vers x_* dans X . Soit $\varepsilon > 0$. La boule $B(\ell, \varepsilon)$ est un ouvert de Y , donc il existe un

ouvert \mathcal{U} de X contenant x_* tel que $f(\mathcal{U} \cap \mathcal{D}_f) \subset B(\ell, \varepsilon)$. Comme x_* est dans \mathcal{U} ouvert, il existe une petite boule $B(x_*, \delta)$ incluse dans \mathcal{U} et donc qui a une image dans $B(\ell, \varepsilon)$. La suite (x_n) tendant vers x_* , il existe un rang N tel que pour tout $n \geq N$, $x_n \in B(x_*, \delta)$. On a alors $f(x_n) \subset B(\ell, \varepsilon)$ pour tout $n \geq N$. Ceci montre que $f(x_n)$ tend vers ℓ .

Enfin, montrons que (iv) implique (ii) par contraposée, c'est-à-dire montrons que non (ii) implique non (iv). Si (ii) est faux, c'est qu'il existe $\varepsilon > 0$ tel que, pour tout $\delta > 0$, il existe un point $x_\delta \in B(x_*, \delta) \cap \mathcal{D}_f$ tel que $f(x_\delta) \notin B(\ell, \varepsilon)$. On applique cette propriété pour $\delta = 1/n$ et on trouve une suite de points $(x_n) \subset \mathcal{D}_f$ tels que $\|x_n - x_*\| < 1/n$ et $f(x_n) \notin B(\ell, \varepsilon)$. Cela donne un contre-exemple à la propriété (iv) qui est donc fausse. \square

Les règles de manipulation sur les limites de fonctions se déduisent directement de la proposition ci-dessus et des règles sur les limites de suites. Évidemment, on ne peut pas parler de la multiplication ou division car ces opérations n'ont généralement pas de sens dans un espace vectoriel.

Proposition 4.8

Soient $(X, \|\cdot\|_X)$ et $(Y, \|\cdot\|_Y)$ deux espaces vectoriels normés. Soient $f : \mathcal{D}_f \subset X \rightarrow Y$, $g : \mathcal{D}_g \subset X \rightarrow Y$ et $\lambda \in \mathbb{R}$. Pour tout $x_* \in \overline{\mathcal{D}_f \cap \mathcal{D}_g}$, si f et g ont des limites en x_* , alors $\lambda f + g : \mathcal{D}_f \cap \mathcal{D}_g \rightarrow Y$ a aussi une limite en x_* et

$$\lim_{x \rightarrow x_*} (\lambda f + g)(x) = \lambda \lim_{x \rightarrow x_*} f(x) + \lim_{x \rightarrow x_*} g(x).$$

Proposition 4.9

Soient $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ et $(Z, \|\cdot\|_Z)$ trois espaces vectoriels normés. Soient $f : \mathcal{D}_f \subset X \rightarrow Y$, $g : \mathcal{D}_g \subset Y \rightarrow Z$. Quitte à restreindre f , on suppose que $f(\mathcal{D}_f) \subset \mathcal{D}_g$. Soit $x_* \in \overline{\mathcal{D}_f}$. Si f a une limite y_* en x_* et si g a une limite en y_* , alors $g \circ f$ a une limite en x_* et

$$\lim_{x \rightarrow x_*} g(f(x)) = \lim_{y \rightarrow \left(\lim_{x \rightarrow x_*} f(x)\right)} g(y).$$



On fera attention qu'il n'existe plus de concept de « tendre vers l'infini » dans un espace vectoriel normé. Dans \mathbb{R} on voit bien ce que veut dire tendre vers $+\infty$ ou $-\infty$. Mais même dans le plan \mathbb{R}^2 , on peut aller vers l'infini dans plein de directions différentes, mais aussi en tourbillonnant ou avec des mouvements plus tordus. On pourrait juste parler de $\|f\|_Y$ qui tend vers l'infini, au sens des réels.

En dimension finie, nous savons que toutes les normes sont équivalentes et que la convergence des suites est équivalente à la convergence composante par composante. On obtient donc le résultat suivant.

Proposition 4.10

Soit $(X, \|\cdot\|_X)$ un espace vectoriel normé et soit $Y = \mathbb{R}^d$ muni de n'importe quelle norme. On considère une fonction $f : \mathcal{D}_f \subset X \rightarrow \mathbb{R}^d$ avec $f(x) = (f_1(x), \dots, f_d(x))$ et chaque f_i est à valeur dans \mathbb{R} . Alors $f(x)$ converge en x_* si et seulement si chaque $f_i(x)$ converge et on a

$$\lim_{x \rightarrow x_*} f(x) = \left(\lim_{x \rightarrow x_*} f_1(x), \lim_{x \rightarrow x_*} f_2(x), \dots, \lim_{x \rightarrow x_*} f_d(x) \right).$$

Dans \mathbb{R} , il existe les notions de convergence à droite et à gauche. Dans un espace plus général, il y a plein de façons d'approcher un point. Une généralisation raisonnable est la suivante.

Définition 4.11

Soient $f : \mathcal{D}_f \subset X \rightarrow Y$, $D \subset \mathcal{D}_f$ et $x_* \in \overline{D}$. On dit que $f(x)$ **converge quand** $x \rightarrow x_*$ **selon l'ensemble** D si la restriction $f|_D$ converge quand $x \rightarrow x_*$. En particulier, si $D = x_* + \mathbb{R}e$ est une droite affine de E de direction $e \in E$, on dit que $f(x)$ converge en x_* **selon la direction** e .

Exemples :

- La fonction $f : x \in \mathbb{R} \mapsto (x^2, x \ln x) \in \mathbb{R}^2$ est définie sur $]0, +\infty[$. Quand $x \rightarrow 0$, la convergence composante par composante montre que $f(x) \rightarrow (0, 0)$.
- La fonction $f : (x, y) \in \mathbb{R}^2 \mapsto (x + \cos y, x \ln x + \sin y)$ est définie sur le demi-plan ouvert $\mathbb{R}_+^* \times \mathbb{R}$. On veut regarder ce qu'il se passe quand (x, y) tend vers un point du bord $(0, y_*)$. On prend deux suites réelles (x_n) et (y_n) qui tendent respectivement vers 0 et y_* . On a par composition des limites $f(x_n, y_n) \rightarrow (\cos y_*, \sin y_*)$. D'après la caractérisation (iv) de la proposition 4.7, cela montre que $f(x, y) \rightarrow (\cos y_*, \sin y_*)$ quand $(x, y) \rightarrow (0, y_*)$.
- La fonction $f : (x, y) \in \mathbb{R} \mapsto y \ln x \in \mathbb{R}$ est définie sur $\mathbb{R}_+^* \times \mathbb{R}$. On a $f(1/n, 0) = 0$ mais $f(e^{-n}, 1/n) = -1$. On a deux suites qui tendent vers 0 mais dont les images ne convergent pas vers la même limite. D'après la caractérisation (iv) de la proposition 4.7, cela montre que f n'a pas de limite en 0.
- Dans un e.v.n. E quelconque, la fonction $f : x \in E \setminus \{0\} \mapsto \frac{x}{\|x\|^{1/2}}$ est bien définie pour $x \neq 0$. Quand $x \rightarrow 0$, on a

$$\|f(x)\| = \left\| \frac{1}{\|x\|^{1/2}} x \right\| = \frac{1}{\|x\|^{1/2}} \|x\| = \|x\|^{1/2} \rightarrow 0.$$

Donc f converge vers 0 quand x tend vers 0.

- On regarde maintenant la fonction $f : x \in E \setminus \{0\} \mapsto \frac{x}{\|x\|}$ qui est à valeur dans la sphère unité $S = \{x \in E, \|x\| = 1\}$: c'est ce qu'on appelle la normalisation d'un vecteur x . Quand $x \rightarrow 0$, f n'a pas de limite. Par exemple, si on prend un $x \in E$ non nul, la suite $x_n = \frac{1}{n}x$ tend vers 0 et $f(x_n) = x \rightarrow x$ mais la

suite $\tilde{x}_n = -\frac{1}{n}x$ tend aussi vers 0 et $f(\tilde{x}_n) = -x \rightarrow -x \neq x$. Par contre, la fonction f a une limite selon chaque demi-droite $\mathbb{R}_+^* \cdot x$.

- Enfin, on considère la fonction $f : x \in E \setminus \{0\} \mapsto \frac{x}{\|x\|^2}$. Quand $x \rightarrow 0$, on a

$$\|f(x)\| = \left\| \frac{1}{\|x\|^2} x \right\| = \frac{1}{\|x\|} \rightarrow +\infty .$$

La norme de f « explose » près de 0 mais on ne peut pas parler de convergence de f vers un certain objet « infini » car la direction de $f(x)$ dépend de la direction de x .

3 Continuité

Nous suivons la même démarche que pour les fonctions réelles : une fois comprise la notion de limite précédente, la continuité devient naturelle.

Définition 4.12

Soient $(X, \|\cdot\|_X)$ et $(Y, \|\cdot\|_Y)$ deux espaces vectoriels normés. Soient $f : \mathcal{D}_f \subset X \rightarrow Y$ et soit $x_* \in \mathcal{D}_f$. On dit que f est **continue en** x_* si

$$\forall \varepsilon > 0, \exists \delta > 0, \forall x \in \mathcal{D}_f, \|x - x_*\|_X < \delta \Rightarrow \|f(x) - f(x_*)\|_Y < \varepsilon ,$$

c'est-à-dire si $f(x)$ converge vers $f(x_*)$ quand $x \rightarrow x_*$. On dit que f est **continue sur un ensemble** $D \subset \mathcal{D}_f$ si f est continue en chaque $x_* \in D$. On note alors $f \in \mathcal{C}^0(D, Y)$.

Il suffit d'appliquer la proposition 4.7 pour obtenir les caractérisations suivantes.

Proposition 4.13

Soit $f : \mathcal{D}_f \subset X \rightarrow Y$, soit $x_* \in \overline{\mathcal{D}_f}$, les propositions suivantes sont équivalentes.

- (i) f est continue en x_* .
- (ii) Pour toute boule ouverte $B(f(x), \varepsilon)$, il existe une boule $B(x_*, \delta)$ telle que $f(B(x_*, \delta) \cap \mathcal{D}_f) \subset B(f(x), \varepsilon)$.
- (iii) Pour tout ouvert \mathcal{O} de Y contenant $f(x_*)$, il existe un ouvert \mathcal{U} de X contenant x_* tel que $f(\mathcal{U} \cap \mathcal{D}_f) \subset \mathcal{O}$.
- (iv) Pour toute suite $(x_n) \subset \mathcal{D}_f$ qui converge vers x_* dans X , la suite $f(x_n)$ converge dans Y vers $f(x_*)$.

En corollaire immédiat de la définition et des propriétés des limites, on obtient les règles suivantes.

Proposition 4.14

Soient $(X, \|\cdot\|_X)$ et $(Y, \|\cdot\|_Y)$ deux espaces vectoriels normés. Soient $f : \mathcal{D}_f \subset X \rightarrow Y$, $g : \mathcal{D}_g \subset X \rightarrow Y$ et $\lambda : \mathcal{D}_\lambda \subset \mathbb{R} \rightarrow \mathbb{R}$. Pour tout $x_* \in \mathcal{D}_f \cap \mathcal{D}_g \cap \mathcal{D}_\lambda$, si f , g et λ sont continues en x_* , alors $\lambda f + g$ est aussi continue en x_* .

Proposition 4.15

Soient $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ et $(Z, \|\cdot\|_Z)$ trois espaces vectoriels normés. Soient $f : \mathcal{D}_f \subset X \rightarrow Y$, $g : \mathcal{D}_g \subset Y \rightarrow Z$. Soit $x_* \in \mathcal{D}_f$ tel que $f(x_*) \in \mathcal{D}_g$. Si f est continue en x_* et si g est continue en $f(x_*)$ alors $g \circ f$ est continue en x_* .

Exemples :

- La fonction identité $\text{Id} : x \in X \mapsto x \in X$ est toujours continue si les espaces vectoriels de départ et d'arrivée sont les mêmes (par exemple en appliquant la caractérisation (iv) de la proposition 4.13).
- La norme $x \in X \mapsto \|x\|_X \in \mathbb{R}$ est toujours continue pour sa propre topologie, c'est la proposition 3.43.
- La fonction de normalisation $f : x \in X \setminus \{0\} \mapsto \frac{1}{\|x\|_X} x \in X$ est continue en dehors de 0 par combinaison des exemples précédents. Par contre, le travail plus haut sur les limites en 0 de cet exemple montre qu'on ne peut pas prolonger la fonction par continuité en 0 puisqu'elle n'a pas de limite en 0.
- On considère l'espace des polynômes $\mathbb{R}[X]$ muni de la norme $\|a_0 + a_1X + a_2X^2 + \dots + a_pX^p\| = \max_{i=1, \dots, p} |a_i|$. (cf le corollaire 3.18. La dérivation $P \mapsto P'$ n'est pas continue partout. Par exemple, $P_n = \frac{1}{n}X^n$ est une suite qui tend vers 0, la dérivée de 0 est 0 mais $P'_n = X^{n-1}$ est de norme 1 et ne tend donc pas vers 0.

En particulier, quand on combine avec la caractérisation de la convergence composante par composante et l'équivalence des normes qui sont vraies en dimension finie, on obtient la règle suivante.

Soit f une fonction définie d'une partie de \mathbb{R}^d dans $\mathbb{R}^{d'}$ (munis des normes que l'on veut). Si les composantes de $f(x)$ sont définies par des formules utilisant les fonctions réelles continues usuelles appliquées aux coordonnées de x , alors f est continue là où elle est définie.

Exemples :

- La fonction $f : \theta \in \mathbb{R} \mapsto (\cos \theta, \sin \theta) \in \mathbb{R}^2$ est continue et a pour image le cercle unité. En effet, si (θ_n) est une suite convergeant vers θ_* , alors pour continuité des formules trigonométriques, $\cos \theta_n$ tend vers $\cos \theta_*$ et $\sin \theta_n$ tend vers $\sin \theta_*$. La convergence des deux coordonnées donne la convergence de tout le vecteur et donc la continuité.

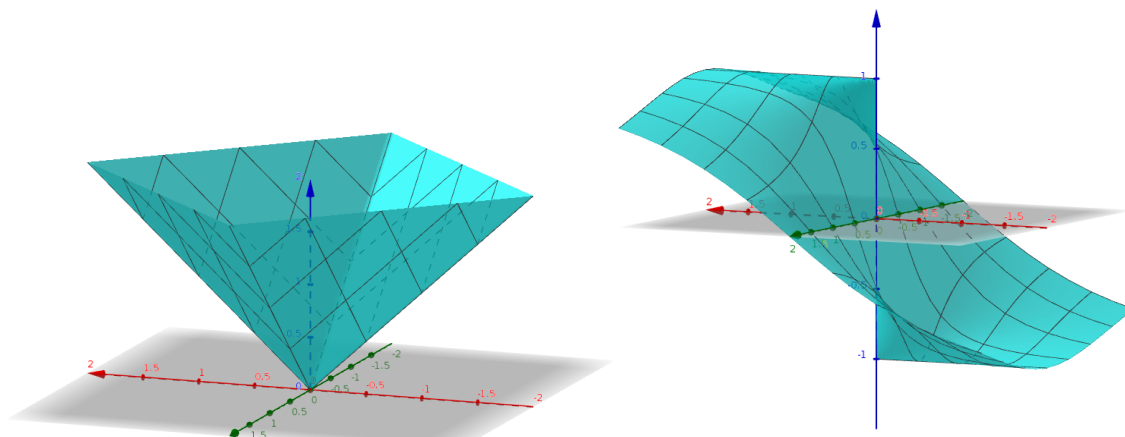


FIGURE 4.1 – À gauche, la fonction $(x,y) \mapsto \|(x,y)\|_1$ qui est continue en 0. À droite, la fonction $(x,y) \mapsto x/\|(x,y)\|_2$ qui est la première coordonnée de la normalisation. On voit qu'en 0, la fonction a toutes les limites possibles entre -1 et 1 suivant la direction selon laquelle on tend vers 0.

- La fonction $f : (x,y) \in \mathbb{R}^2 \mapsto (x^2, e^y \sin x, \frac{y^3}{1+x^2})$ est continue de \mathbb{R}^2 dans \mathbb{R}^3 . En effet, si (x_n, y_n) tend vers (x_*, y_*) , par continuité des fonctions usuelles et propriété des suites réelles, on a $e^{y_n} \sin x_n$ qui tend vers $e^{y_*} \sin y_*$. Cela montre la convergence de la deuxième composante et les autres se traitent pareil.
- On reprend l'exemple plus haut sur la dérivation dans l'espace des polynômes. On se limite à $X = \mathbb{R}_d[X]$ l'espace des polynômes de degré au plus d . Celui-ci est un espace identifiable à \mathbb{R}^{d+1} muni de $\|\cdot\|_\infty$ par l'identification de $P = a_0 + \dots + a_d X^d$ au vecteur (a_0, \dots, a_d) . La dérivation n'est rien d'autre que la fonction $\partial : (a_0, a_1, a_2, \dots, a_d) \mapsto (a_1, 2a_2, \dots, da_d, 0)$ qui est continue comme polynôme en les coefficients.
- On munit l'espace des matrices carrées $\mathcal{M}_d(\mathbb{R})$ d'une norme quelconque. C'est un espace de dimension n^2 donc toutes les normes sont équivalentes. Le déterminant d'une matrice $A = (a_{i,j})$ est un réel qui peut s'écrire selon la formule

$$\det A = \sum_{\sigma \in \mathfrak{S}_d} \text{sign}(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} \dots a_{d,\sigma(d)} .$$

Cette formule dépend polynomialement des coefficients, elle est donc continue par rapport à la matrice.

De manière générale, on pourra constater que l'application des critères de convergence et continuité ainsi que les propriétés des normes équivalentes montrent qu'une fonction est continue ou non indépendamment du fait de remplacer les normes de départ et d'arrivée par des normes équivalentes. Par contre, attention quand les normes ne sont pas équivalentes!

Exemple :

On considère l'identité sur l'espace de fonction $\mathcal{C}^0([0,1],\mathbb{R})$. Si on prend les mêmes normes au départ et à l'arrivée, alors elle est continue par application triviale du critère des suites. Par contre, si on met maintenant des normes différentes au départ et à l'arrivée en considérant

$$\psi : f \in (\mathcal{C}^0([0,1],\mathbb{R}), \|\cdot\|_1) \longmapsto f \in (\mathcal{C}^0([0,1],\mathbb{R}), \|\cdot\|_\infty).$$

On a vu dans le chapitre précédent quand la suite $f_n : x \mapsto x^n$ tend vers 0 pour la norme 1 car $\|f_n\|_1 = 1/(n+1)$ mais $\|\psi(f_n) - \psi(0)\|_\infty = \|f_n\|_\infty = 1 \not\rightarrow 0$. Donc ψ n'est pas continue en 0. Cela montre que quand les normes ne sont pas équivalentes, la continuité d'une fonction peut dépendre des normes choisies.

Il existe une caractérisation topologique de la continuité de f en tout point. Dans le cas des fonctions définies partout, celle-ci est surtout utile pour prouver que des ensembles sont des ouverts ou des fermés.

Proposition 4.16

Soient $(X, \|\cdot\|_X)$ et $(Y, \|\cdot\|_Y)$ deux espaces vectoriels normés et soit $f : X \rightarrow Y$ une fonction définie partout. Alors f est continue si et seulement si l'image réciproque $f^{-1}(\mathcal{O})$ de tout ouvert \mathcal{O} de Y est un ouvert de X . De même, f est continue si et seulement si l'image réciproque $f^{-1}(\mathcal{O})$ de tout fermé \mathcal{O} de Y est un fermé de X .

Démonstration : Soit f une fonction continue sur X et soit \mathcal{O} un ouvert de Y . Soit $x \in f^{-1}(\mathcal{O})$. Par définition $f(x) \in \mathcal{O}$ et donc comme \mathcal{O} est ouvert, il existe une petite boule $B(f(x), r)$ entièrement incluse dans \mathcal{O} . En utilisant la caractérisation (ii) de la proposition 4.13, on obtient une boule $B(x, \delta)$ telle que $f(B(x, \delta)) \subset B(f(x), r) \subset \mathcal{O}$. Mais ceci implique que $B(x, \delta) \subset f^{-1}(\mathcal{O})$. On vient de trouver, pour un point $x \in f^{-1}(\mathcal{O})$ quelconque, une boule $B(x, \delta)$ restant dans $f^{-1}(\mathcal{O})$, ce qui montre que $f^{-1}(\mathcal{O})$ est un ouvert.

Si on suppose maintenant que f est telle que l'image réciproque d'un ouvert est toujours ouvert, alors la caractérisation (iii) de la proposition 4.13 avec $\mathcal{U} = f^{-1}(\mathcal{O})$ prouve que f est continue partout.

L'équivalence pour les fermés se trouve en passant au complémentaire (ou en faisant les démonstrations à l'aide des caractérisations séquentielles). \square



Ne pas se tromper de sens : l'image réciproque d'un ouvert par une fonction continue est un ouvert. Mais l'image directe d'un ouvert n'est pas forcément ouvert, voir l'exemple plus bas.

Exemples :

- La fonction $f : (x, y) \in \mathbb{R}^2 \mapsto y - x^2 \in \mathbb{R}$ est continue car polynomiale en les coordonnées. Comme $]0, +\infty[$ est un ouvert de \mathbb{R} , l'ensemble $\{(x, y) \in \mathbb{R}^2, y >$

x^2 au-dessus de la parabole est un ouvert de \mathbb{R}^2 car c'est $f^{-1}(]0, +\infty[)$.

- Dans tout espace vectoriel normé $(X, \|\cdot\|)$, la sphère $S = \{x \in X, \|x\| = 1\}$ est un fermé de X comme image réciproque de $\{1\}$ par la fonction continue $\|\cdot\|$.
- On a vu que le déterminant est une fonction continue en la matrice. Comme $\{0\}$ est un singleton et est donc fermé dans \mathbb{R} , on en déduit que son image réciproque par le déterminant est fermée, c'est-à-dire que l'ensemble des matrices de déterminant nul est un fermé. En conséquence, l'ensemble des matrices inversibles est un ouvert des matrices (notons qu'on ne précise pas la norme puisqu'elles sont toutes équivalentes en dimension finie et que la propriété d'ouverture ou fermeture ne dépend donc pas du choix de la norme). On a montré dans les compléments du chapitre précédent sa densité. Donc on trouve que les matrices inversibles forment un ouvert dense de l'espace des matrices, ce qui veut dire que « quasiment toutes » les matrices sont inversibles.
- Une fonction constante f qui à tout $x \in X$ associe le même $c \in Y$ est évidemment continue sur X . L'image réciproque $f^{-1}(\mathcal{O})$ d'un ensemble \mathcal{O} est soit X soit \emptyset suivant que c est dans \mathcal{O} ou pas. Comme X et \emptyset sont des ouverts et des fermés, on retrouve bien que f est continue. Mais $f(X) = \{c\}$, X est ouvert dans X et $\{c\}$ n'est pas ouvert (si Y non trivial). Donc l'image d'un ouvert par une fonction continue n'est pas forcément un ouvert.

Pour finir cette partie, nous allons voir deux notions plus fortes de continuité.

Définition 4.17

Soient $(X, \|\cdot\|_X)$ et $(Y, \|\cdot\|_Y)$ deux espaces vectoriels normés. Soit $f : \mathcal{D}_f \subset X \rightarrow Y$ et soit $D \subset \mathcal{D}_f$. On dit que f est **uniformément continue sur D** si

$$\forall \varepsilon > 0, \exists \delta > 0, \forall x, x' \in D, \|x - x'\|_X < \delta \Rightarrow \|f(x) - f(x')\|_Y < \varepsilon .$$

Définition 4.18

Soient $(X, \|\cdot\|_X)$ et $(Y, \|\cdot\|_Y)$ deux espaces vectoriels normés. Soit $f : \mathcal{D}_f \subset X \rightarrow Y$ et soit $K \geq 0$. On dit que f est **K -lipschitzienne** si

$$\forall x, x' \in \mathcal{D}_f, \|f(x) - f(x')\|_Y \leq K \|x - x'\|_X .$$

Plus généralement, on dit que f est **lipschitzienne** si elle est K -lipschitzienne pour un certain K . Dans le cas particulier où $K \in [0,1[$ (resp. $K \in [0,1]$), on dit que f est **une contraction** (resp. est **strictement contractante**).

Il y a un ordre strict de force pour ces différentes propriétés.

Proposition 4.19

Toute fonction lipschitzienne est uniformément continue et toute fonction uniformément continue est continue. Les réciproques sont fausses.

Démonstration : Supposons f K -lipschitzienne. Si $K = 0$, c'est une fonction constante et elle est clairement uniformément continue. Sinon, pour tout $\varepsilon > 0$, on pose $\delta = \varepsilon/K$ et pour tout x et x' vérifiant $\|x - x'\|_X < \delta$, on a

$$\|f(x) - f(x')\|_Y \leq K\|x - x'\| < K\delta = \varepsilon .$$

Donc f est aussi uniformément continue.

Si f est uniformément continue, il suffit de fixer x' à l'avance, sans lui laisser la liberté du « pour tout » pour qu'on obtienne la définition de la continuité en x' . Donc f est continue en tout x' et donc continue partout.

Nous avons déjà vu sur \mathbb{R} que $x \mapsto x^2$ est continue mais pas uniformément continue. Nous savons aussi par le théorème de Heine que $f : x \in [0,1] \mapsto \sqrt{x}$ est uniformément continue sur $[0,1]$. Mais si on prend $x = 0$ et $x' = 1/n$, on a $\frac{|f(x) - f(x')|}{|x - x'|} = \sqrt{n}$ qui ne peut pas être borné par une constante K quand n tend vers l'infini. Donc f n'est pas lipschitzienne sur $[0,1]$. \square

Dans \mathbb{R} , on utilisera souvent le critère suivant.

Proposition 4.20

Soit $f : I \rightarrow \mathbb{R}$ une fonction dérivable sur un intervalle réel $I \subset \mathbb{R}$. Alors f est K -lipschitzienne si et seulement si sa dérivée f' est bornée par $|f'| \leq K$ sur I . En particulier, $f \in \mathcal{C}^1([a,b], \mathbb{R})$ est forcément lipschitzienne.

Démonstration : Si x et x' sont deux points de I , disons $x < x'$ (s'il y a égalité, tout est trivial), alors le théorème des accroissements finis dit que

$$|f(x) - f(x')| \leq \left(\max_{\theta \in]x, x'[} |f'(\theta)| \right) \cdot |x - x'| .$$

Une borne K sur la dérivée donne directement que f est K -lipschitzienne. Supposons maintenant que f est dérivable et K -lipschitzienne. On a alors

$$\left| \frac{f(x+h) - f(x)}{h} \right| = \frac{|f(x+h) - f(x)|}{|(x+h) - x|} \leq K .$$

En prenant la limite $h \rightarrow 0$, on obtient la borne sur $|f'(x)|$.

Enfin, si $f \in \mathcal{C}^1([a,b], \mathbb{R})$, elle est lipschitzienne puisque sa dérivée est bornée car continue sur un intervalle compact. \square

Exemples :

- La proposition précédente montre que les fonctions sinus et cosinus sont 1-lipschitziennes (i.e. contractantes) sur \mathbb{R} .
- La fonction $x \mapsto x^2$ n'étant pas uniformément continue sur \mathbb{R} , elle ne peut pas être lipschitzienne. Mais sur chaque intervalle $[-M, M]$, elle est $2M$ -lipschitzienne.
- Sur tout espace vectoriel X muni d'une norme $\|\cdot\|$, la fonction norme, définie de X dans \mathbb{R} , est contractante en application des inégalités triangulaires :

$$|f(x) - f(x')| = \left| \|x\| - \|x'\| \right| \leq \|x - x'\| .$$

La norme est donc 1-lipschitzienne et donc uniformément continue sur X et a fortiori continue sur X .

- De façon triviale, une homothétie $x \mapsto \lambda x$ avec $\lambda \in \mathbb{R}$ est $|\lambda|$ -lipschitzienne sur tout espace vectoriel normé et une translation $x \mapsto x + x_0$ est 1-lipschitzienne.
- L'intégrale $I(f) = \int_a^b f(x)dx$ sur l'espace $\mathcal{C}^0([a, b], \mathbb{R})$ est une fonction 1-lipschitzienne pour la norme 1. En effet

$$|I(f) - I(g)| = \left| \int_a^b (f(x) - g(x))dx \right| \leq \int_a^b |f(x) - g(x)|dx = \|f - g\|_1 .$$

- L'intégrale $I(f) = \int_a^b f(x)dx$ sur l'espace $\mathcal{C}^0([a, b], \mathbb{R})$ est une fonction $(b - a)$ -lipschitzienne pour la norme infini. En effet

$$|I(f) - I(g)| \leq \int_a^b |f(x) - g(x)|dx \leq \int_a^b \max_{t \in [a, b]} |f(t) - g(t)|dx = (b - a)\|f - g\|_\infty .$$

- Les deux exemples précédents nous montre que l'intégrale est une fonction continue sur $\mathcal{C}^0([a, b], \mathbb{R})$ muni d'une de ces normes. Donc l'ensemble des fonctions à intégrale nulle est un fermé de $\mathcal{C}^0([a, b], \mathbb{R})$ pour ces topologies car il s'agit de $I^{-1}(\{0\})$.

On démontre aisément les propriétés suivantes.

Proposition 4.21

La somme de fonctions K - et K' -lipschitziennes est $(K + K')$ -lipschitzienne. La composition de fonctions K - et K' -lipschitziennes est KK' -lipschitzienne.

Démonstration : Supposons f et g respectivement K - et K' -lipschitziennes

de X dans Y . Pour tout x et x' dans X ,

$$\begin{aligned} \|(f+g)(x) - (f+g)(x')\|_Y &= \|(f(x) - f(x')) + (g(x) - g(x'))\|_Y \\ &\leq \|f(x) - f(x')\|_Y + \|g(x) - g(x')\|_Y \\ &\leq K\|x - x'\|_X + K'\|x - x'\|_X \\ &\leq (K + K')\|x - x'\|_X . \end{aligned}$$

Si maintenant $f : X \rightarrow Y$ est K -lipschitzienne et $g : Y \rightarrow Z$ est K' -lipschitzienne, alors

$$\|(g \circ f)(x) - (g \circ f)(x')\|_Z \leq K'\|f(x) - f(x')\|_Y \leq K'K\|x - x'\|_X . \quad \square$$

Proposition 4.22

Si f est une fonction K -lipschitzienne, alors il existe a et b réels tels que $\|f(x)\|_Y \leq a\|x\|_X + b$.

Démonstration : On prend le cas particulier $x' = 0$ et on obtient $\|f(x) - f(0)\|_Y \leq K\|x - 0\|_X$. Par ailleurs, on sait que $\|f(x) - f(0)\|_Y \geq \|f(x)\|_Y - \|f(0)\|_Y$. Cela montre la proposition pour $a = K$ et $b = \|f(0)\|_Y$. \square

4 Applications linéaires

Parmi les fonctions entre espaces vectoriels normés, les applications linéaires ont une place privilégiée puisque ce sont celles qui respectent la structure d'espace vectoriel.

Définition 4.23

Une application **linéaire** $f : X \rightarrow Y$ entre deux espaces vectoriels réels est une fonction telle que

$$\forall x, x' \in X , \forall \lambda \in \mathbb{R} , f(\lambda x + x') = \lambda f(x) + f(x') .$$

On introduit une norme naturelle pour les applications linéaires.

Définition 4.24

Soit $f : X \rightarrow Y$ une fonction linéaire entre deux espaces vectoriels réels. S'il existe et est fini, on appelle **norme triple** ou **norme d'application** de f le nombre

$$\| \|f\| \|_{X \rightarrow Y} := \sup_{x \in X \setminus \{0\}} \frac{\|f(x)\|_Y}{\|x\|_X} .$$

S'il n'y a pas d'ambiguïté, on notera simplement $\| \|f\| \|$.

La norme triple est géométriquement le facteur d'étirement maximal que présente la fonction. Sa finitude est équivalente à la continuité de la fonction.

Théorème 4.25

Soit $f : X \rightarrow Y$ une fonction linéaire entre deux espaces vectoriels réels non triviaux. Les propositions suivantes sont équivalentes.

- (i) la fonction f est continue de $(X, \|\cdot\|_X)$ dans $(Y, \|\cdot\|_Y)$
- (ii) la fonction f est continue en un point $x_* \in X$
- (iii) la fonction f est continue en 0
- (iv) la norme triple $\|f\|$ de f existe et est finie
- (v) la fonction f est K -lipschitzienne avec $K = \|f\|$

Démonstration : Nous allons procéder par boucle d'équivalence. Il est clair que (i) implique (ii) puisque la continuité en un point est un cas particulier de la continuité partout.

Supposons que (ii) est vraie. Soit (ε_n) une suite qui tend vers 0 dans X . On pose $x_n = x_* + \varepsilon_n$ qui est une suite qui tend vers x_* . Par continuité en x_* , on a que $f(x_n)$ tend vers $f(x_*)$. Mais $f(x_n) = f(x_* + \varepsilon_n) = f(x_*) + f(\varepsilon_n)$ par linéarité. Donc $f(\varepsilon_n)$ tend vers 0 dans Y . Comme f est linéaire, $0_Y = f(0_X)$ et on vient donc de montrer que $f(\varepsilon_n)$ tend vers $f(0)$. Donc (iii) est vraie.

Supposons que (iii) est vraie. Pour $\varepsilon = 1$, il existe une boule $B_X(0, \delta)$ telle que pour tout $x \in B_X(0, \delta)$, $\|f(x) - f(0)\|_Y < 1$ et donc $\|f(x)\|_Y < 1$ puisque $f(0) = 0$. La borne supérieure définissant la norme triple est sur un ensemble non vide et bien défini. Il suffit donc de montrer que cet ensemble est majoré. Soit $x \neq 0$. On pose $\tilde{x} = \frac{\delta}{2\|x\|_X}x$, on a $\|\tilde{x}\|_X = \frac{\delta}{2\|x\|_X}\|x\|_X < \delta$. Donc on sait que $\|f(\tilde{x})\|_Y < 1$, ce qui donne que $\left\|f\left(\frac{\delta}{2\|x\|_X}x\right)\right\|_Y < 1$. En utilisant la linéarité de f et l'homogénéité de la norme, on obtient $\frac{\delta\|f(x)\|_Y}{2\|x\|_X} < 1$ et donc $\frac{\|f(x)\|_Y}{\|x\|_X} < \frac{2}{\delta}$. Donc $2/\delta$ est un majorant de notre ensemble et sa borne supérieure existe et est finie.

Supposons (iv) vraie et posons $K = \|f\|$. Soit x et x' dans X . Si $x = x'$, il est trivial que $\|f(x) - f(x')\|_Y = 0 \leq K\|x - x'\|_X$. Sinon, on a par définition de K que $\frac{\|f(x-x')\|_Y}{\|x-x'\|_X} \leq K$ et donc que $\|f(x) - f(x')\|_Y = \|f(x-x')\|_Y \leq K\|x - x'\|_X$, où on a utilisé la linéarité de f . Au total, f est bien K -lipschitzienne.

Finalement, on a déjà vu que l'implication f lipschitzienne implique f continue est toujours vraie. Donc (v) implique (i). \square

On a plusieurs écritures possibles de la norme triple.

Proposition 4.26

On a les égalités suivantes

$$\begin{aligned} \| \|f\| \| &:= \sup_{x \in X \setminus \{0\}} \frac{\|f(x)\|_Y}{\|x\|_X} = \sup_{x \in B_X(0,1)} \|f(x)\|_Y \\ &= \sup_{x \in \overline{B}_X(0,1)} \|f(x)\|_Y = \sup_{x \in X, \|x\|_X=1} \|f(x)\|_Y . \end{aligned}$$

Démonstration : Il est clair que plus l'ensemble est petit, plus la borne supérieure est petite. Le reste des arguments est basé sur le principe de normalisation : si $x \neq 0$, on pose $\tilde{x} = x/\|x\|_X$. On a $\|\tilde{x}\|_X = 1$ par homogénéité de la norme. Par linéarité de f et homogénéité de la norme, on a aussi

$$\frac{\|f(x)\|_Y}{\|x\|_X} = \|f(\tilde{x})\|_Y$$

donc le premier sup est le même que le dernier sup sur la sphère.

Pour montrer la dernière égalité, on constate simplement que si $\|x\|_X < 1$, alors soit $x = 0$ et $f(x) = 0$, soit $\tilde{x} = x/\|x\|_X$ est de norme 1 et $\|f(x)\|_Y = \|f(\tilde{x})\|_Y \|x\|_X < \|f(\tilde{x})\|_Y$. Donc la valeur en x ne dépasse pas la valeur en \tilde{x} et le sup n'est pas influencé par les valeurs de f sur la boule ouverte.

La seconde égalité consiste à passer à la limite. Si x est de norme 1, on pose $x_n = (1 - 1/n)x$ qui est dans la boule ouverte. Par linéarité de f et homogénéité de la norme, on a $\|f(x_n)\|_Y = (1 - 1/n)\|f(x)\|_Y \rightarrow \|f(x)\|_Y$. Donc $\sup_n \|f(x_n)\|_Y$ est $\|f(x)\|_Y$ et on retrouve ainsi les valeurs sur la sphère omise dans le sup. \square

Corollaire 4.27

Soient $(X, \|\cdot\|_X)$ et $(Y, \|\cdot\|_Y)$ deux espaces vectoriels normés. Si X est de dimension finie, alors toute fonction linéaire $f : X \rightarrow Y$ est continue.

Démonstration : Soit (e_1, \dots, e_p) une base de X . Pour tout $x \in X$, avec $x = x_1 e_1 + \dots + x_p e_p$, on pose $N(x) = \max_i |x_i|$. On vérifie qu'il s'agit d'une norme sur X (c'est la norme infinie dans la base donnée) et donc par équivalence des normes en dimension finie, il existe $M > 0$ tel que $N(x) \leq M\|x\|_X$. On a alors

$$\begin{aligned} \|f(x)\|_Y &= \|f(x_1 e_1 + \dots + x_p e_p)\|_Y = \|x_1 f(e_1) + \dots + x_p f(e_p)\|_Y \\ &\leq \sum_{i=1}^p |x_i| \|f(e_i)\|_Y \leq p \max_i |x_i| \max_i \|f(e_i)\|_Y \\ &\leq pM (\max_i \|f(e_i)\|_Y) \|x\|_X . \end{aligned}$$

Cette majoration est équivalente à dire que $\| \|f\| \| \leq pM (\max_i \|f(e_i)\|_Y)$ et

donc f est continue d'après les résultats précédents. \square

On fera attention que, comme souvent dans ce cours, le cas de la dimension infinie est plus subtil.

Exemples :

- On se place dans l'espace des polynômes $\mathbb{R}[X]$ muni de la norme $\|a_0 + a_1X + \dots + a_dX^d\| = \sup_i |a_i|$. Pour tout $P = a_0 + a_1X + \dots + a_dX^d$, on pose $f(P) = a_0X + \frac{1}{2}a_1X^2 + \dots + \frac{1}{d+1}a_dX^{d+1}$ qui est la primitive de P . Il s'agit d'une opération linéaire et comme, clairement, $\|f(P)\| \leq \|P\|$, cela montre que la primitivation est une opération continue sur cet espace de polynômes.
- Dans le même cadre, on regarde maintenant la dérivation, qui est aussi une opération linéaire. On constate que $\|X^n\| = 1$ et $\|nX^{n-1}\| = n$. Donc $\|\partial X^n\|/\|X^n\| \rightarrow +\infty$ quand $n \rightarrow +\infty$ et le rapport dans la définition de la norme triple de la dérivation n'est pas majoré. La dérivation n'est donc pas une application linéaire continue.

La norme triple est plus qu'un test pratique et rapide pour montrer qu'une fonction est continue. Il s'agit, comme son nom l'indique, d'une norme.

Définition 4.28

Pour X et Y deux espaces vectoriels normés, on note $\mathcal{L}(X, Y)$ l'ensemble des applications linéaires continues de X dans Y .

Proposition 4.29

Soit X et Y deux espaces vectoriels normés. La norme triple $f \in \mathcal{L}(X, Y) \mapsto \| \|f\| \|$ est une norme sur $\mathcal{L}(X, Y)$.

Démonstration : Il est clair que la norme triple est un nombre positif qui est bien défini par hypothèse sur f . Supposons que $\| \|f\| \| = 0$, on a pour tout $x \neq 0$, $\|f(x)\|_Y \leq \| \|f\| \| \cdot \|x\|_X = 0$ et donc $f \equiv 0$. Si f et g sont deux fonctions linéaires continues, alors leur somme est forcément linéaire et continue et on a

$$\begin{aligned} \| \|f + g\| \| &= \sup_{x \in B_X(0,1)} \|f(x) + g(x)\|_Y \leq \sup_{x \in B_X(0,1)} (\|f(x)\|_Y + \|g(x)\|_Y) \\ &\leq \sup_{x \in B_X(0,1)} \|f(x)\|_Y + \sup_{x \in B_X(0,1)} \|g(x)\|_Y = \| \|f\| \| + \| \|g\| \| . \end{aligned}$$

Il nous reste à vérifier l'homogénéité :

$$\| \|\lambda f\| \| = \sup_{x \in B_X(0,1)} \|\lambda f(x)\|_Y = \sup_{x \in B_X(0,1)} |\lambda| \|f(x)\|_Y = |\lambda| \cdot \| \|f\| \| .$$

\square

On peut composer facilement les différents normes triples. Si $Y = X$, on parle alors de norme d'algèbre c'est-à-dire une norme telle que $\|AB\| \leq \|A\| \cdot \|B\|$.

Proposition 4.30

Si X, Y et Z sont trois espaces vectoriels normés, si $f \in \mathcal{L}(X, Y)$ a pour norme triple $\|f\|_{X \rightarrow Y}$ et si $g \in \mathcal{L}(Y, Z)$ a pour norme triple $\|g\|_{Y \rightarrow Z}$, alors $g \circ f$ est une fonction de $\mathcal{L}(X, Z)$ et sa norme triple vérifie

$$\|g \circ f\|_{X \rightarrow Z} \leq \|f\|_{X \rightarrow Y} \|g\|_{Y \rightarrow Z} .$$

Démonstration : Pour tout $x \in X$, on a

$$\|(g \circ f)(x)\|_Z \leq \|g\|_{Y \rightarrow Z} \|f(x)\|_Y \leq \|g\|_{Y \rightarrow Z} \|f\|_{X \rightarrow Y} \|x\|_X . \quad \square$$

Application aux matrices et au conditionnement :

On munit \mathbb{R}^d de la norme $\|\cdot\|_\infty$. Les applications linéaires de \mathbb{R}^d dans \mathbb{R}^d sont représentées par les matrices dans le sens où $A \in \mathcal{M}_d(\mathbb{R})$ code l'application linéaire $x \mapsto Ax$. Comme on est en dimension finie, ces applications sont toutes continues et on munit $\mathcal{M}_d(\mathbb{R})$ de la norme triple associée. Cette norme triple peut se calculer explicitement. Soit $A = (a_{ij}) \in \mathcal{M}_d(\mathbb{R})$, on a

$$\|Ax\|_\infty = \max_i \left| \sum_j a_{ij} x_j \right| \leq \max_i \sum_j |a_{ij}| |x_j| \leq \left(\max_i \left(\sum_j |a_{ij}| \right) \right) \|x\|_\infty .$$

Ceci montre que $\|A\| \leq \max_i (\sum_j |a_{ij}|)$. Si la ligne i_0 est une ligne où le max est atteint et si $x_j = +1$ si $a_{i_0 j} \geq 0$ et $x_j = -1$ sinon, alors on a

$$|(Ax)_{i_0}| = \left| \sum_j a_{i_0 j} x_j \right| = \sum_j |a_{i_0 j}| = \max_i \left(\sum_j |a_{ij}| \right) .$$

Donc $\|Ax\|_\infty \geq \max_i (\sum_j |a_{ij}|) \|x\|_\infty$ et au final

$$\|A\| = \max_i \left(\sum_j |a_{ij}| \right)$$

On souhaite maintenant résoudre l'équation linéaire consistant à trouver $x \in X$ tel que $Ax = y$. Si une approximation δy est commise sur y , quelle erreur δx obtient-on sur x ? Ceci est lié au *conditionnement* de A , défini par

$$\text{Condi}(A) = \|A\| \times \|A^{-1}\| .$$

En effet, on a $A(x + \delta x) = y + \delta y$ et donc $A\delta x = \delta y$. D'où

$$\|\delta x\|_\infty = \|A^{-1} \delta y\|_\infty \leq \|A^{-1}\| \|\delta y\|_\infty .$$

Par ailleurs,

$$\|y\|_\infty = \|Ax\|_\infty \leq \|A\| \|x\|_\infty .$$

On obtient donc l'estimation suivante pour les erreurs relatives :

$$\frac{\|\delta x\|_\infty}{\|x\|_\infty} \leq \text{Condi}(A) \frac{\|\delta y\|_\infty}{\|y\|_\infty}.$$

Comme le conditionnement est symétrique si on change A en A^{-1} , l'estimation précédente reste vraie si on échange x et y , c'est-à-dire si on a plutôt une erreur en x et qu'on cherche à contrôler l'erreur relative en y . En analyse numérique, ce conditionnement est important car si un problème linéaire est mal conditionné, c'est-à-dire qu'il fait intervenir des A pour lesquels $\text{Condi}(A)$ est grand, les erreurs ont toutes les chances d'exploser pendant les calculs. Il est souvent important de *bien conditionner* le problème en reformulant le problème en un problème équivalent mais qui fait intervenir des opérateurs avec des conditionnements plus petits.

5 Compléments

5.1 Topologie induite

Nous avons vu qu'une fonction $f : X \rightarrow Y$ est continue si et seulement si l'image réciproque d'un ouvert est un ouvert, cf. la proposition 4.16. Cette caractérisation est pratique mais ne marche pas si f est définie seulement sur un sous-ensemble \mathcal{D}_f de X . Il est alors utile d'introduire la notion suivante, qui est plus généralement une façon de définir les ouverts et fermés d'un sous-ensemble d'un espace vectoriel normé.

Définition 4.31

Soit $\mathcal{D}_f \subset X$ une partie d'un espace vectoriel normé. On dit que $\mathcal{O} \subset \mathcal{D}_f$ est un **ouvert du domaine** \mathcal{D}_f ou que \mathcal{O} est **la trace d'un ouvert** sur \mathcal{D}_f s'il existe un ouvert \mathcal{O}' de X tel que $\mathcal{O} = \mathcal{O}' \cap \mathcal{D}_f$. De même, on dit que $\mathcal{F} \subset \mathcal{D}_f$ est un **fermé du domaine** \mathcal{D}_f ou que \mathcal{F} est **la trace d'un fermé** sur \mathcal{D}_f s'il existe un fermé \mathcal{F}' de X tel que $\mathcal{F} = \mathcal{F}' \cap \mathcal{D}_f$.

Exemples :

- On se place sur \mathbb{R} avec $\mathcal{D}_f = [0, +\infty[$. L'intervalle $[0,1[$ n'est pas un ouvert de \mathbb{R} mais c'est un ouvert du domaine \mathcal{D}_f car $[0,1[=] - 1,1[\cap \mathcal{D}_f$.
- On se place sur un espace vectoriel normé E quelconque et on suppose que \mathcal{D}_f contient un point e qui est isolé dans le sens où il existe $r > 0$ tel que $B(e,r) \cap \mathcal{D}_f = \{e\}$. Alors $\{e\}$ est à la fois un ouvert et un fermé du domaine car $\{e\} = B(e,r) \cap \mathcal{D}_f = \overline{B}(e,r/2) \cap \mathcal{D}_f$.

On obtient alors la généralisation de la proposition 4.16. La démonstration est la même mutatis mutandis puisque la définition ci-dessus colle parfaitement aux caractérisations de la proposition 4.13.

Proposition 4.32

Soient $(X, \|\cdot\|_X)$ et $(Y, \|\cdot\|_Y)$ deux espaces vectoriels normés et soit $f : X \rightarrow Y$ une fonction définie partout. Alors f est continue si et seulement si l'image réciproque $f^{-1}(\mathcal{O})$ de tout ouvert \mathcal{O} de Y est un ouvert de X . De même, f est continue si et seulement si l'image réciproque $f^{-1}(\mathcal{O})$ de tout fermé \mathcal{O} de Y est un fermé de X .

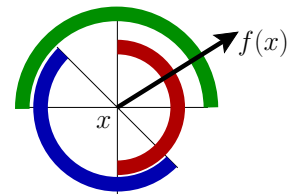
5.2 Le point fixe de Brouwer par un lemme « à la Sperner »

On a énoncé dans la proposition 3.60 un lemme « à la Sperner » sur les coloriages d'une triangulation. Nous allons l'utiliser pour démontrer un résultat qui généralise le théorème de point fixe 2.41 dans un triangle.

Proposition 4.33 (point fixe de Brouwer dans un triangle)

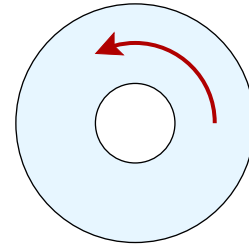
Soit $T \subset \mathbb{R}^2$ le triangle rectangle isocèle de sommets $(0,0)$, $(1,0)$ et $(0,1)$, bords compris. Toute fonction continue $f : T \rightarrow T$ admet un point fixe.

Démonstration : On découpe le triangle en petits triangles comme auparavant et on regarde chaque sommet de la triangulation. Si le sommet est un point fixe, c'est gagné. Sinon, le vecteur $\overrightarrow{(x, f(x))}$ est non nul et pointe donc une direction. On colorie le sommet par une couleur possible dans le découpage en secteur ci-contre, sauf sur les bords où on prend parmi les choix la couleur qui donnera un coloriage selon les règles de la proposition 3.60. On vérifie que cela est bien possible : par exemple sur le bord ouest, l'image $f(x)$ est forcément à droite de x (pour ne pas sortir du triangle) et donc le choix de la couleur rouge est possible. Si aucun sommet n'est un point fixe, on retrouve un coloriage auquel on peut appliquer la proposition 3.60 et il existe un triangle de sommets x_1 , y_1 et z_1 tricolore. On réapplique le procédé avec une triangulation de plus en plus fine. Soit on tombe à un moment sur un point fixe et c'est gagné, soit on obtient des suites de sommets rouges (x_n) , verts (y_n) et bleus (z_n) avec $\|x_n - y_n\| \rightarrow 0$ et $\|x_n - z_n\| \rightarrow 0$. Par le théorème 3.50, on peut supposer, quitte à en extraire une sous-suite, que (x_n) converge vers x_* , qui est aussi limite de (y_n) et (z_n) . Si $f(x_*) \neq x_*$, le vecteur $\overrightarrow{(x_*, f(x_*))}$ a une direction avec laquelle une des couleurs n'est pas compatible. Par continuité, cela contredirait la convergence d'une des suites (x_n) , (y_n) ou (z_n) . Donc x_* est bien un point fixe de f . \square



Nous verrons tout à la fin de ce cours qu'il est possible de généraliser le résultat à d'autres formes géométrique F que le triangle T . En particulier, ce résultat de point fixe est vrai pour le disque. Là où ce résultat montre une grande profondeur

topologique, c'est qu'il n'est pas vrai pour toutes les formes F . Ainsi, si F est un anneau, une rotation non-triviale autour du centre est une fonction continue sans point fixe. Le « trou » dans F fait que l'anneau est différent du triangle et du disque. Commencer à classifier les formes du plan en « comptant » le nombre de « trous » et en regardant les propriétés que cela implique, c'est s'engouffrer dans le domaine des *topologues*.



Chapitre 5 : Complétude

1 Suites de Cauchy et espaces complets

On peut généraliser la notion de suites de Cauchy de façon naturelle.

Définition 5.1

Une suite (x_n) d'un espace vectoriel normé E est dite **de Cauchy** si pour tout $\varepsilon > 0$, il existe un rang $N \in \mathbb{N}$ tel que pour tous $p \geq N$ et $q \geq N$, on a $\|x_p - x_q\| < \varepsilon$.

Une interprétation géométrique de cette définition est la suivante.

Proposition 5.2

Une suite (x_n) est de Cauchy si et seulement si, pour tout $r > 0$, la suite finit par entrer dans une boule de rayon r et ne plus en sortir.

Démonstration : Si la suite est dans une boule $B(x, r)$ à partir du rang N , alors $\|x_p - x_q\| \leq \|x_p - x\| + \|x - x_q\| < 2r$ pour tout $p, q \geq N$. Réciproquement, si $\|x_p - x_q\| < \varepsilon$ pour tout $p, q \geq N$, alors x_p est dans $B(x_N, \varepsilon)$ pour tout $p \geq N$. \square

Un des liens liant le concept de suite de Cauchy et la convergence est toujours vrai.

Proposition 5.3

Si une suite (x_n) d'un espace vectoriel normé E est convergente, alors elle est de Cauchy.

Démonstration : Soit $\ell \in E$ la limite de la suite et soit $\varepsilon > 0$. Comme la suite converge, il existe un rang N tel que les termes x_n avec $n \geq N$ sont dans la boule $B(\ell, \varepsilon/2)$. Dans ce cas, pour tout p et q plus grands que N , on a

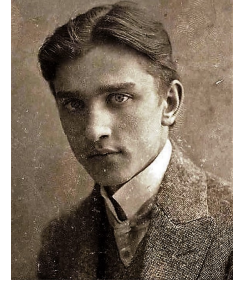
$$\|x_p - x_q\| = \|(x_p - \ell) - (x_q - \ell)\| \leq \|x_p - \ell\| + \|x_q - \ell\| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \quad \square$$

Dans \mathbb{R} , la réciproque est vraie mais ce n'est pas le cas dans tous les espaces. Il s'agit d'une propriété importante et que les espaces qui la vérifient sont appelés *complets*.

Définition 5.4

Soit $(E, \|\cdot\|)$ un espace vectoriel normé. On dit qu'il est **complet** si toutes ses suites de Cauchy convergent dans E . On dit aussi que E est un **espace de Banach**.

Les deux définitions d'espace complet et d'espace de Banach seront équivalentes pour nous. En fait, il y a une distinction : un espace topologique dont toutes les suites de Cauchy convergent est dit complet. Quand cet espace est de plus un espace vectoriel normé, alors on parle d'espace de Banach. Comme ce cours ne regarde que ce cas particulier d'espace topologique, les deux définitions se recouvrent. Le nom de Banach revient souvent en analyse fonctionnelle.



Stefan Banach
1892-1945
Pologne

Proposition 5.5

L'espace \mathbb{R}^d est complet : il s'agit d'un espace de Banach.

Démonstration : Comme toutes les normes sont équivalentes en dimension finie, on peut considérer par exemple la norme infinie. Si (x_n) est une suite de Cauchy dans \mathbb{R}^d , alors chaque coordonnée est une suite de Cauchy dans \mathbb{R} . En effet, pour tout $\varepsilon > 0$, il existe un rang N tel que $\|x_n - x_m\|_\infty < \varepsilon$ pour tout $n, m \geq N$ et on a alors

$$\forall n, m \geq N, \forall i \in \llbracket 1, d \rrbracket, |x_n^i - x_m^i| \leq \|x_n - x_m\|_\infty < \varepsilon.$$

La complétude de la droite réelle nous donne que chaque coordonnée converge vers un réel ℓ^i puis la proposition 3.47 montre que (x_n) tend vers $\ell = (\ell^1, \dots, \ell^d)$.

□

Est-ce qu'il y a des exemples d'espaces non-complets ? Oui, dès qu'on passe en dimension infinie, les choses sont moins triviales.

Proposition 5.6

On considère $\mathbb{R}[X]$ l'espace des polynômes réels muni de la norme

$$P = a_p X^p + a^{p-1} X^{p-1} + \dots + a_1 X + a_0 \in \mathbb{R}[X] \mapsto \|P\| = \max_i |a_i|.$$

Cet espace vectoriel normé n'est pas complet.

Démonstration : On a vu dans le corollaire 3.18 qu'il s'agit bien d'un espace vectoriel normé. Montrons que cet espace n'est pas complet. Soit $(P_n)_{n \in \mathbb{N}}$ la suite de polynômes

$$P_n = X + \frac{1}{2}X^2 + \frac{1}{3}X^3 + \dots + \frac{1}{n}X^n .$$

Il s'agit d'une suite de Cauchy car si $p \geq q \geq N$, on a

$$\|P_p - P_q\| = \left\| \frac{1}{p}X^p + \frac{1}{p-1}X^{p-1} + \dots + \frac{1}{q+1}X^{q+1} \right\| = \max_{i \in [q+1, p]} \frac{1}{i} \leq \frac{1}{N} .$$

Donc pour tout $\varepsilon > 0$, on peut trouver N assez grand tel que $\|P_q - P_q\| < \varepsilon$ pour tout $p, q \geq N$. Mais est-ce que la suite (P_n) a une limite? Si oui, alors sa limite P devrait avoir le coefficient a_i de X^i égal à $1/i$ pour tout i car

$$\forall i \in \mathbb{N} , \forall n \geq i , \|P_n - P\| \geq \left| \frac{1}{i} - a_i \right| .$$

Mais dans ce cas, cela ne peut être un polynôme car il y a un nombre infini de coefficients non nuls. Donc le seul candidat à être la limite de (P_n) n'existe pas dans l'espace $\mathbb{R}[X]$ et donc $\mathbb{R}[X]$ n'est pas complet. \square

Dans l'exemple précédent, on peut noter qu'on a envie de dire qu'il manque simplement des éléments à $\mathbb{R}[X]$ pour le rendre complet. Le procédé consistant à « boucher les trous » en ajoutant toutes les limites virtuelles des suites de Cauchy à l'espace de départ est ce qu'on appelle la *complétion*. L'espace obtenu est appelé le *complété* de l'espace de départ. C'est exactement le procédé qui donne la construction de Cauchy de \mathbb{R} comme complété de \mathbb{Q} .

Avoir une dimension infinie n'est pas toutefois un obstacle à être complet.

Proposition 5.7

L'espace $\mathcal{C}^0([0,1], \mathbb{R})$ muni de la norme $\|f\|_\infty = \max_{x \in [0,1]} |f(x)|$ est un espace complet.

Démonstration : Voir la fin du cours. \square

2 Convergence normale des séries et applications

Nous nous plaçons dans un espace vectoriel normé $(E, \|\cdot\|)$. Soit $(x_n) \subset E$ une suite de E , on considère la série $(\sum_{n \geq 0} x_n)$ qui est définie comme la suite des sommes partielles

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

selon le même procédé que pour les séries réelles. En particulier, on dit que la série converge si la suite des sommes partielles $(S_N)_{N \geq 0}$ converge dans E vers un point $S \in E$. On notera alors

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

Savoir si une série converge ou pas peut être délicat (cf cours de L2 pour les séries simplement dans $E = \mathbb{R}$). Mais il existe un critère utile pour prouver la convergence qui repose sur la convergence normale (i.e. « en norme »).

Définition 5.8

On dit que la série $(\sum_{n \geq 0} x_n)$ converge normalement si la série de termes positifs réels $(\sum_{n \geq 0} \|x_n\|)$ converge dans \mathbb{R} .

Dans \mathbb{R} , la norme est la valeur absolue. La définition ci-dessus correspond donc à la convergence absolue. L'exemple des séries à termes de signe quelconque dans \mathbb{R} montre qu'une série peut converger sans pour autant converger normalement. Par contre, l'implication réciproque est vraie dans les espaces complets.

Proposition 5.9

Si $(E, \|\cdot\|)$ est un espace vectoriel complet, alors toute série normalement convergente est convergente dans E .

Démonstration : Soit $U_N = \sum_{n=0}^N \|x_n\|$ la somme partielle de la série des normes. Comme elle converge dans \mathbb{R} , elle est de Cauchy et pour tout $\varepsilon > 0$, il existe N_0 tel que, pour tout $p \geq q \geq N_0$,

$$|U_p - U_q| = \sum_{n=q+1}^p \|x_n\| < \varepsilon .$$

Mais alors, on a aussi par inégalité triangulaire

$$\|S_p - S_q\| = \left\| \sum_{n=q+1}^p x_n \right\| \leq \sum_{n=q+1}^p \|x_n\| < \varepsilon .$$

Ceci montre que la suite (S_N) est de Cauchy. Comme E est complet, elle converge. \square

Applications aux séries de matrices :

Pour le reste de cette partie, on se place sur \mathbb{R}^d muni d'une certaine norme $\|\cdot\|$. Les applications linéaires de $\mathcal{L}(\mathbb{R}^d, \mathbb{R}^d)$ peuvent être représentées par des matrices $A \in \mathcal{M}_d(\mathbb{R})$ et on note $\|A\|$ la norme triple associée.

Commençons par voir que l'on peut définir l'exponentielle d'une matrice.

Proposition 5.10 (exponentielle de matrice)

Soit $A \in \mathcal{M}_d(\mathbb{R})$. La série $(\sum_{n \geq 0} \frac{1}{n!} A^n)$ est normalement convergente. Sa limite est appelée **exponentielle de A** et on note

$$e^A = \sum_{n \geq 0} \frac{1}{n!} A^n .$$

Démonstration : Par propriété de composition de la norme triple, on sait que $|||AB||| \leq |||A||| \cdot |||B|||$ et donc

$$|||\frac{1}{n!} A^n||| = \frac{1}{n!} |||A^n||| \leq \frac{1}{n!} |||A|||^n \leq \frac{1}{n!} |||A|||^n .$$

La série $(\sum |||A|||^n/n!)$ est convergente (application du critère de D'Alembert ou simplement en notant que c'est la série de $e^{|||A|||}$) et par comparaison de séries à termes positifs, la série de l'exponentielle de matrice est bien normalement convergente. \square

L'exponentielle de matrice est un objet important dans l'analyse des équations différentielles et en géométrie différentielle. On peut ainsi montrer que $X(t) = e^{tA} X_0$ est la solution du système d'équations différentielles linéaires $X'(t) = AX(t)$ avec $X(0) = X_0 \subset \mathbb{R}^d$ et les exponentielles de matrices permettent de résoudre tous les systèmes de ce type.

Si A est une matrice de $\mathcal{M}_d(\mathbb{R})$ qui est diagonalisable de valeurs propres $\lambda_1, \dots, \lambda_d$, alors on appelle *rayon spectral* le nombre $\rho(A) = \max |\lambda_i|$.

Proposition 5.11

Soit $A \in \mathcal{M}_d(\mathbb{R})$ une matrice diagonalisable. Il existe une constante M telle que

$$\forall n \geq 0, \quad |||A^n||| \leq M \rho(A)^n . \quad (5.1)$$

En particulier, si $\rho(A) < 1$, la suite A^n tend exponentiellement vite vers 0.

Démonstration : Pour démontrer (5.1), on suppose que la diagonalisation est sous la forme $A = PDP^{-1}$. On pose $N : x \in \mathbb{R}^d \mapsto \|P^{-1}x\|_\infty$. On vérifie directement que N est bien une norme sur \mathbb{R}^d . Comme toutes les normes sont équivalentes en dimension finie, il existe une constante $K \geq 1$ telle que $K^{-1}\|x\| \leq N(x) \leq K\|x\|$. Enfin, on note que $Dx = (\lambda_1 x_1, \dots, \lambda_d x_d)$ et donc que $\|Dx\|_\infty \leq \rho(A)\|x\|_\infty$. On a pour tout $x \in \mathbb{R}^d$,

$$\begin{aligned} \|A^n x\| &\leq KN(A^n x) = K\|P^{-1}A^n x\|_\infty = K\|D^n P^{-1}x\|_\infty \\ &\leq K\rho(A)^n \|P^{-1}x\|_\infty = K\rho(A)^n N(x) \leq K^2 \rho(A)^n \|x\| . \end{aligned}$$

Ceci montre bien (5.1) avec $M = K^2$. \square

Corollaire 5.12

Soit $A \in \mathcal{M}_d(\mathbb{R})$ une matrice diagonalisable telle que $\rho(A) < 1$. Alors $\text{Id} - A$ est inversible et son inverse est donné par la série

$$(\text{Id} - A)^{-1} = \text{Id} + A + A^2 + A^3 + \dots$$

Démonstration : On utilise la convergence normale de la proposition 5.9 dans $\mathcal{M}_d(\mathbb{R})$ muni de la norme triple, qui est de dimension finie et donc complet. Comme $\rho(A) < 1$, la proposition 5.11 montre que la série $\sum_{n \geq 0} A^n$ est normalement convergente et donc converge vers une matrice B . Par ailleurs,

$$(\text{Id} + A + A^2 + A^3 + \dots + A^N) \cdot (\text{Id} - A) = \text{Id} - A^{N+1} \xrightarrow{N \rightarrow +\infty} \text{Id}$$

et en passant à la limite, on trouve bien que $B(\text{Id} - A) = \text{Id}$ (on utilise ici que la norme triple étant une norme d'algèbre, la multiplication par $(\text{Id} - A)$ est continue et on peut passer à la limite dedans). \square

Applications aux séries de fonctions :

On munit $\mathcal{C}^0([a,b],\mathbb{R})$ de la norme $\|\cdot\|_\infty$. On a admis qu'il s'agit bien d'un espace complet. Soit $\alpha > 1$ et $\beta \in \mathbb{R}$. Comme $(\sum_{n \geq 1} \frac{1}{n^\alpha})$ est convergente (série de Riemann) et comme $\|\sin(n^\beta \cdot)\|_\infty \leq 1$, on obtient que la somme

$$f_{\alpha,\beta}(x) = \sum_{n=1}^{\infty} \frac{1}{n^\alpha} \sin(n^\beta x)$$

correspond bien à une fonction continue car la série est normalement convergente. Pour $\beta = 1$, on retrouve des séries de type Fourier. La série définissant $f_{\alpha,\beta}$ converge toujours pour la norme uniforme, mais les dérivées des termes sont du type $n^{\beta-\alpha} \cos(n^\beta x)$ ce qui peut donner une série explosant rapidement si β est grand. C'est avec ce type d'idée que Karl Weierstrass et Leopold Kronecker ont découvert en 1872 la première fonction continue partout mais dérivable nul part.



FIGURE 5.1 – La fonction $f_{2,5} = \sum_{n=1}^{\infty} \frac{1}{n^2} \sin(n^5 \cdot)$ est continue mais très irrégulière.

3 Le point fixe des applications contractantes

Nous avons déjà vu des théorèmes de points fixes. Un résultat extrêmement important avec de nombreuses applications pratiques est celui du point fixe des fonctions strictement contractantes.

Théorème 5.13 (point fixe de Picard ou point fixe de Banach)

Soit $(E, \|\cdot\|)$ un espace vectoriel normé complet. Soit D un fermé de E et soit $f : D \rightarrow D$ une fonction strictement contractante, c'est-à-dire K -lipschitzienne avec $K < 1$. Alors f admet exactement un point fixe x_* dans D , c'est-à-dire un point tel que $f(x_*) = x_*$. En outre, pour toute donnée initiale $x_0 \in D$, la suite définie par récurrence par $x_{n+1} = f(x_n)$ converge vers x_* exponentiellement vite.

Démonstration : Si x et x' sont deux points fixes de f dans D , alors on a

$$\|x - x'\| = \|f(x) - f(x')\| \leq K\|x - x'\|$$

avec $K < 1$. Ceci n'est possible que si $\|x - x'\| = 0$. Donc il y a au plus un point fixe pour f .

Prenons maintenant une suite définie par récurrence $x_{n+1} = f(x_n)$ dans D . Remarquons que comme f envoie D dans D , la suite est toujours bien définie et est contenue dans D . Nous allons montrer que cette suite est une suite de Cauchy. Soit $n \geq 1$, on a

$$\|u_n - u_{n+1}\| = \|f(u_{n-1}) - f(u_n)\| \leq K\|u_{n-1} - u_n\|$$

et par une récurrence immédiate

$$\|u_n - u_{n+1}\| \leq K^n \|u_0 - u_1\| .$$

Pour tous $p < q$, comme $K \neq 1$, on a alors

$$\begin{aligned} \|u_p - u_q\| &\leq \|u_p - u_{p+1}\| + \|u_{p+1} - u_{p+2}\| + \dots + \|u_{q-1} - u_q\| \\ &\leq (K^p + K^{p+1} + \dots + K^{q-1}) \|u_0 - u_1\| \\ &\leq \frac{K^p - K^q}{1 - K} \|u_0 - u_1\| \\ &\leq \frac{K^p}{1 - K} \|u_0 - u_1\| \xrightarrow{p \rightarrow +\infty} 0 . \end{aligned}$$

Pour tout $\varepsilon > 0$, si p et q sont suffisamment grands, alors $\|u_p - u_q\| < \varepsilon$ et donc la suite est de Cauchy. Comme E est complet, cette suite a une limite x_* dans E . Mais comme la suite (u_n) est contenue dans D fermé, la limite x_* est aussi dans D . On sait que u_n tend vers x_* et on a aussi par continuité de f que $f(u_n)$ converge vers $f(x_*)$. Mais comme $f(u_n) = u_{n+1}$ qui converge aussi vers x_* , on a par unicité de la limite $x_* = f(x_*)$ qui est bien un point fixe de f .

Finalement, reprenons une suite (u_n) comme ci-dessus. On a par la même astuce

$$\|u_n - x_*\| = \|f(u_{n-1}) - f(x_*)\| \leq K\|u_{n-1} - x_*\| \leq \dots \leq K^n\|u_0 - x_*\|.$$

Comme $K < 1$, cela montre que la convergence de toute suite récurrente vers x_* avec une vitesse exponentielle. \square

Exemples :

- On reprend l'algorithme de Héron pour calculer la racine carrée d'un nombre $a > 0$. On considère la fonction $f : x > 0 \mapsto \frac{1}{2}x + \frac{a}{2x}$. On a $f(\sqrt{a}) = \sqrt{a}$ et $f'(x) = \frac{1}{2} - \frac{a}{2x^2}$ donc $0 \leq f'(x) < \frac{1}{2}$ si $x \in [\sqrt{a}, +\infty[$. Le domaine $D = [\sqrt{a}, +\infty[$ est un fermé de \mathbb{R} qui est complet. La fonction f est croissante sur D et \sqrt{a} est un point fixe de f . Donc f envoie D sur lui-même. Enfin, on a $|f'(x)| \leq \frac{1}{2}$ sur D donc f est $\frac{1}{2}$ -lipschitzienne sur D . L'application du théorème 5.13 montre que toute suite récurrente $u_{n+1} = f(u_n)$ avec $u_0 \geq \sqrt{a}$ converge vers \sqrt{a} .
- De manière plus générale que l'exemple de l'algorithme de Héron, supposons que x_* est un point fixe d'une fonction f telle que $|f'(x_*)| < 1$. Si f est de classe \mathcal{C}^1 , il existe un intervalle $D = [x_* - r, x_* + r]$ sur lequel $|f'(x)| \leq K$ pour un certain $K < 1$. Par contraction, on a forcément $f(D) \subset D$ et l'application du théorème 5.13 montre que toute suite itérative qui commence dans D converge vers x_* . On dit que x_* est un point fixe (localement) (asymptotiquement) stable pour le système dynamique $u_{n+1} = f(u_n)$ et que D est dans son bassin d'attraction.
- Le plan d'une région $D \subset \mathbb{R}^2$ est une représentation d'échelle strictement plus petite que 1 et donc la fonction qui associe à un point géographique concret sa représentation sur la carte est une contraction stricte. Si on pose un plan du département de l'Isère sur une table incluse dans le département, alors il existe un et seul point de la carte qui est exactement à l'endroit réel qui lui correspond.

Application à l'algorithme de Google

Dans l'article *The Anatomy of a Large-Scale Hypertextual Web Search Engine* (Computer Networks and ISDN Systems n°30, 1998, p. 107-117), Sergey Brin et Lawrence Page présentent l'algorithme *Page Rank* qui est derrière le moteur de recherche Google. Il est proposé de mettre un poids PR sur chaque page web et le vecteur de \mathbb{R}^N ainsi formé (avec N égal au nombre de pages web connus!) doit résoudre une équation linéaire.

We assume page A has pages T1...Tn which point to it (i.e., are citations). The parameter d is a damping factor which can be set between 0 and 1. We usually set d to 0.85. There are more details about d in the next section. Also C(A) is defined as the number of links going out of page A. The PageRank of a page A is given as follows:

$$PR(A) = (1-d) + d (PR(T1)/C(T1) + \dots + PR(Tn)/C(Tn))$$

Note that the PageRanks form a probability distribution over web pages, so the sum of all web pages' PageRanks will be one.

Pour être plus précis, appelons x le vecteur dont la coordonnée x_i est le poids de la page i . On considère qu'une page de poids x_j qui admet k_j liens transmet à chacune des pages référencées le poids x_j/k_j (une page est intéressante si plein de pages intéressantes pointent sur elle). Soit c_{ij} le nombre qui vaut 1 si la page j a un lien vers la page i et 0 sinon. Alors on pose $A = (a_{ij})$ avec $a_{ij} = c_{ij}/k_j$ et les poids doivent donc vérifier que $x_i = \sum_j a_{ij}x_j$. Mais ce principe n'est pas satisfaisant : il y a plein de mauvaises solutions, comme $x = 0$, ou bien $x_i \neq 0$ seulement pour une page "impasse" qui ne renvoie sur aucune autre page. L'article de Brin et Page propose donc que chaque page ait un poids $(1 - d)$ minimal avec $d \in]0,1[$ et que l'autre proportion d du poids provienne du mécanisme ci-dessus. Donc x doit résoudre

$$x = (1 - d)\mathbb{1} + dAx$$

avec $d \in]0,1[$ et $\mathbb{1}$ le vecteur rempli de 1. La matrice A associée aux liens dans les pages web a de plus la propriété que pour tout colonne j , $\sum_i a_{ij} = 1$ et que $a_{ij} \geq 0$. On admet que $\max_j \sum_i |a_{ij}|$ est la norme triple de A associée à la norme $\|\cdot\|_1$ sur \mathbb{R}^N (on a vu l'expression de la norme triple pour la norme infinie qui est symétrique). Si on regarde la fonction affine $f : x \in \mathbb{R}^d \mapsto (1 - d)\mathbb{1} + dAx$, on a donc

$$\begin{aligned} \|f(x) - f(x')\|_1 &= \|((1 - d)\mathbb{1} + dAx) - ((1 - d)\mathbb{1} + dAx')\|_1 = \|d(Ax - Ax')\|_1 \\ &\leq d\|A(x - x')\|_1 \leq d\|A\|\|x - x'\|_1 = d\|x - x'\|_1 . \end{aligned}$$

Comme $|d| < 1$, il s'agit d'une fonction strictement contractante et on sait qu'il existe donc un unique point fixe pour l'équation. En outre, comme on sait aussi que A est à coefficients positifs, alors on a que f envoie D sur D où D est le fermé $(\mathbb{R}_+)^d$ de \mathbb{R}^d . Le point fixe a donc des coefficients tous positifs : ce sont les notes de pertinences qui permettent de classer les pages. On voit aussi qu'on a une façon explicite de trouver ces notes par suite itérative, mais ce n'est pas concrètement la méthode utilisée par Google car la matrice est trop grande pour que les calculs soient assez rapides. Google utilise des méthodes probabilistes (le « surfeur aléatoire »).

Application aux équations différentielles

Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ une fonction K -lipschitzienne. On considère l'équation différentielle

$$x(0) = 0 \quad \text{et} \quad x'(t) = f(x(t)) .$$

On va montrer qu'il existe une unique solution à cette équation, au moins pour des temps petits. En fait, la solution est globale dans notre cas et on peut aussi

prendre un cadre bien plus général, mais on ne donne ici qu'un avant-goût de ce qu'on appelle le *théorème de Cauchy-Lipschitz*. Pour $\delta > 0$ assez petit à déterminer plus tard, on considère $E = \mathcal{C}^0([-\delta, \delta], \mathbb{R})$ muni de $\|\cdot\|_\infty$, qui est complet. On pose $\Phi : x \in E \rightarrow y \in E$ la fonction dont l'image est définie par

$$\Phi(x)(t) = y(t) = \int_0^t f(x(s)) \, ds .$$

Pour deux fonctions x_1 et x_2 de E , on a

$$\begin{aligned} \|\Phi(x_1) - \Phi(x_2)\|_\infty &= \max_{t \in [-\delta, \delta]} \left| \int_0^t f(x_1(s)) \, ds - \int_0^t f(x_2(s)) \, ds \right| \\ &\leq \int_{-\delta}^\delta |f(x_1(s)) - f(x_2(s))| \, ds \leq \int_{-\delta}^\delta K |x_1(s) - x_2(s)| \, ds \\ &\leq \int_{-\delta}^\delta K \max_{\tau \in [-\delta, \delta]} |x_1(\tau) - x_2(\tau)| \, ds = 2K\delta \|x_1 - x_2\|_\infty . \end{aligned}$$

Donc si on prend $\delta > 0$ assez petit pour que $2K\delta < 1$, la fonction Φ est strictement contractante sur E et admet un unique point fixe x qui résoud l'équation fonctionnelle

$$\forall t \in [-\delta, \delta] , \quad x(t) = \int_0^t f(x(s)) \, ds .$$

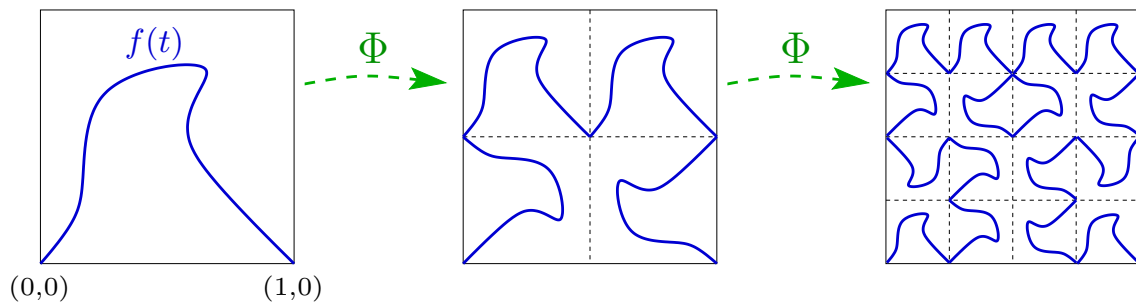
La dérivation de l'équation fonctionnelle ci-dessus nous montre que $x'(t) = f(x(t))$ et on a clairement $x(0) = 0$. Réciproquement, si x vérifie l'équation différentielle, x vérifie l'équation fonctionnelle par intégration. Donc il existe une unique solution à notre équation différentielle sur l'intervalle $[-\delta, \delta]$ où $\delta > 0$ a été choisi assez petit (on parle de solution locale).

Application à la courbe de Peano

On considère l'espace des courbes paramétrées $f(t) = (x(t), y(t))$ dans \mathbb{R}^2 donné par $\mathcal{C}^0([0, 1], \mathbb{R}^2)$ muni de la norme

$$\|f\|_\infty = \max_{t \in [0, 1]} \|f(t)\|_\infty = \max_{t \in [0, 1]} \max(|x(t)|, |y(t)|) .$$

On admet qu'il s'agit d'un espace complet (car il est complet pour chaque coordonnée). On appelle D le fermé de cet espace composé des courbes f telles que $f(0) = (0, 0)$ et $f(1) = (1, 0)$. On considère l'application Φ qui à $f \in D$ associe quatre copies de la courbe recollées comme ci-dessous.



On vérifie que Φ est bien définie car les recollements sont bien continus. On peut aussi vérifier que Φ est $1/2$ -lipschitzienne car chaque copie subit une réduction de facteur $1/2$ et donc la distance entre les courbes $\Phi(f_1)$ et $\Phi(f_2)$ est moitié de celle entre les courbes f_1 et f_2 (même s'il y a 4 copies de chaque courbe). Donc Φ admet un unique point fixe et la courbe f_* correspondante est limite des itérations $\Phi^n(f)$. On peut constater qu'à l'itération n , la courbe $\Phi^n(f)$ passe par tous les points $(i/2^n, j/2^n)$. On peut prouver à la limite que f_* passera alors par tous les points du carré $[0,1]^2$. La courbe f_* est donc une courbe, objet qu'on aurait pu penser de dimension un, mais dont l'image est un carré de dimension 2. Il s'agit d'une courbe similaire à la courbe qu'a introduit Giuseppe Peano (1858-1932, Italie) pour illustrer qu'un intervalle a le même cardinal qu'un carré.

Chapitre 6 : Compacité

1 Les compacts

1.1 Définitions

Dans ce cours, nous allons utiliser la notion suivante de compacité. Celle-ci est ce qu'on appelle la *compacité séquentielle*, c'est-à-dire caractériser par les suites. La définition qui est plus générale est celle par les recouvrements qui apparaît dans la proposition qui suit, qui n'est pas en général la même que la compacité séquentielle. Mais dans les espaces vectoriels normés, les deux définitions sont équivalentes et on peut utiliser l'une comme l'autre.

Définition 6.1

Soit $(E, \|\cdot\|)$ un espace vectoriel normé. Un sous-ensemble $K \subset E$ est dit (séquentiellement) **compact** si pour toute suite (u_n) d'éléments de K , on peut extraire une sous-suite $(u_{\varphi(n)})$ qui converge vers une limite dans K . Autrement dit, toute suite de K a au moins une valeur d'adhérence dans K .

Théorème 6.2 (Propriété de Bolzano-Weierstrass)

Soit $(E, \|\cdot\|)$ un espace vectoriel normé. Un sous-ensemble $K \subset E$ est compact si et seulement si de tout recouvrement de K par des ouverts, on peut en extraire un recouvrement fini, c'est-à-dire que si $(\mathcal{O}_j)_{j \in J}$ est une famille d'ouverts de E telle que $K \subset (\cup_j \mathcal{O}_j)$, alors il existe j_1, \dots, j_p tels que $K \subset (\mathcal{O}_{j_1} \cup \dots \cup \mathcal{O}_{j_p})$.

Démonstration : Soit $K \subset E$ un ensemble dont on peut extraire un recouvrement fini de tout recouvrement par des ouverts. Soit (u_n) une suite dans K . Imaginons que (u_n) n'ait pas de valeur d'adhérence dans K . Alors la proposition 3.49 nous dit que $\cap_N \overline{\cup_{n \geq N} \{u_n\}}$ est disjointe de K . Donc en particulier, son complémentaire $\cup_N (\overline{\cup_{n \geq N} \{u_n\}})^C$ est une union d'ouverts (car complémentaires de fermés) qui recouvre K . On peut donc en extraire un recouvrement fini avec les indices $N_1 < N_2 < \dots < N_p$ et, en passant au complémentaire, on obtient que

$$K \text{ disjoint de } \overline{\cup_{n \geq N_1} \{u_n\}} \cup \dots \cup \overline{\cup_{n \geq N_p} \{u_n\}} = \overline{\cup_{n \geq N_p} \{u_n\}}.$$

Mais c'est absurde car $\overline{\cup_{n \geq N_p} \{u_n\}}$ contient des éléments de la suite qui sont dans K . Donc (u_n) a bien au moins une valeur d'adhérence dans K .

Réciproquement, supposons que K est séquentiellement compact. Soit $(\cup_j \mathcal{O}_j)$ un recouvrement de K par des ouverts. Nous affirmons d'abord qu'il existe $r > 0$ tel que toute boule $B(x, r)$ avec $x \in K$ est entièrement incluse dans un \mathcal{O}_j . En effet, si ce n'est pas le cas, alors on peut trouver des boules $B(x_n, 1/n)$ telles qu'aucune ne soit incluse dans un \mathcal{O}_j . Comme K est compact, on peut extraire de (x_n) une suite $(x_{\varphi(n)})$ qui converge vers $\ell \in K$. Comme ℓ est dans K , il doit exister un ouvert \mathcal{O}_{j^*} tel que $\ell \in \mathcal{O}_{j^*}$. Et comme \mathcal{O}_{j^*} est ouvert, il existe une boule $B(\ell, \rho)$ incluse dans \mathcal{O}_{j^*} . Par inégalité triangulaire, on voit alors que les boules $B(x_{\varphi(n)}, 1/\varphi(n))$ finissent par être incluses dans $B(\ell, \rho)$ et donc dans \mathcal{O}_{j^*} , ce qui est absurde.

Nous considérons maintenant les boules $B(x, r)$ avec $x \in K$ et $r > 0$ comme ci-dessus. Nous prétendons qu'il est possible de recouvrir K avec un nombre fini de telles boules. En effet, si ce n'est pas le cas, on peut trouver une suite de points $(x_n) \subset K$ tels que x_n n'est pas dans $\cup_{k \geq n} B(x_k, r)$. Cette suite vérifie par construction que $\|x_p - x_q\| \geq r > 0$ pour tout p et q . Donc on ne pourra jamais extraire une suite de Cauchy de cette suite et donc aucune sous-suite convergente, ce qui est absurde.

Nous pouvons conclure : nous pouvons recouvrir K par un nombre fini de boules $B(x_k, r)$ avec $k = 1, \dots, p$. Mais chaque $B(x_k, r)$ est entièrement incluse dans un \mathcal{O}_{j_k} . Donc les ouverts \mathcal{O}_{j_k} ($k = 1, \dots, p$) suffisent à recouvrir tout K .
□

L'intérêt du concept de compacité est assez évident. En effet, nous savons qu'il existe des suites bornées qui ne convergent pas. Mais il arrive qu'on construise une suite et qu'on aimerait qu'elle converge, quitte à ne pas prendre tous les termes de la suite. On peut voir cette démarche dans plusieurs preuves de la fin du chapitre 2 par exemple (théorème des valeurs extrêmes, théorème de Heine...). Si on a bien de la compacité, alors on peut toujours faire la démarche d'extraire une sous-suite convergente, ce qui est très pratique et fournit des théorèmes très puissants comme nous allons le voir. La question est maintenant de savoir quels ensembles sont des compacts.

Proposition 6.3

Un compact K est un fermé borné de E .

Démonstration : Pour nous entraîner à manipuler des deux caractérisations, nous allons faire cette preuve de deux façons différentes : une avec les suites, l'autre avec des manipulations d'ouverts.

Supposons que K est séquentiellement compact. Soit (x_n) une suite de K qui converge vers $\ell \in E$. On peut extraire de (x_n) une sous-suite $(x_{\varphi(n)})$ qui a une limite $\ell' \in K$. Mais comme la suite extraite converge aussi vers ℓ , par unicité de la limite, on a $\ell = \ell' \in K$. Ceci montre que la limite ℓ reste dans K et donc

K est fermé. Supposons que K n'est pas borné, alors pour tout $n \in \mathbb{N}$, il existe $x_n \in K$ tel que $\|x_n\| \geq n$. Toute sous-suite de (x_n) aura aussi une norme qui tend vers $+\infty$ et donc aucune sous-suite ne peut converger. Donc K n'est pas compact et par contraposée, si K est compact, il est borné.

Supposons maintenant que K vérifie la caractérisation de la compacité par les recouvrements d'ouverts. Soit $x \notin K$. Pour tout $r > 0$, on note $\mathcal{O}_r = (\overline{B}(x,r))^C$. Comme l'intersection de toutes ces boules se limite au centre x qui n'est pas dans K , on a $K \subset \cup_{r>0} \mathcal{O}_r$. Par hypothèse, on peut trouver un recouvrement fini. Si r_* est le plus petit de tous les rayons du recouvrement fini, alors on a $K \subset \mathcal{O}_{r_*}$. Ceci montre que la boule ouverte $B(x,r_*)$ est dans K^C et donc que K^C est un ouvert. Par complémentaire, K est fermé. Par ailleurs, l'union des boules ouvertes $\cup_n B(0,n)$ étant tout l'espace, c'est un recouvrement de K par des ouverts. Donc on peut le recouvrir par un nombre fini de telles boules et la plus grande de ces boules contient tout K , ce qui montre que K est borné. \square



On fera bien attention à ne pas utiliser le résultat dans le mauvais sens car sa réciproque n'est pas toujours vraie.

1.2 En dimension finie

En dimension finie, comme souvent dans ce cours, les propriétés sont similaires à celles de \mathbb{R} et la réciproque de la proposition 6.3 est vraie.

Théorème 6.4 (Borel-Lebesgue)

Dans un espace de dimension finie, un ensemble K est compact si et seulement s'il est fermé et borné.

Démonstration : Il n'y a que le sens « si » à montrer puisque le « seulement si » est général d'après la proposition 6.3. Tout espace de dimension finie peut se ramener à des coordonnées dans \mathbb{R}^d par le choix d'une base. Par ailleurs, toutes les normes sont équivalentes d'après le théorème 3.54 et donc il suffit de considérer la norme ∞ sur \mathbb{R}^d par exemple. Soit K un fermé borné de \mathbb{R}^d et soit $(x_n)_{n \in \mathbb{N}}$ une suite de K . Comme K est borné, la suite (x_n) est bornée et d'après le théorème 3.50, elle admet une sous-suite $(x_{\varphi_1 \circ \dots \circ \varphi_d(n)})$ qui converge vers un point x_* dans \mathbb{R}^d . Pour finir, il reste à voir que le fait que K soit fermé implique bien que cette limite x_* est dans K . \square

En appliquant ce résultat à une boule fermée de \mathbb{R}^d , on a le corollaire suivant.

Corollaire 6.5

De toute suite bornée d'un espace de dimension finie, on peut extraire une sous-suite convergente.

Exemples :

- Le carré $[0,1]^d$ est un compact de \mathbb{R}^d .
- La suite $x_n = (\cos n, \sin n)$ qui est contenue dans le cercle unité de \mathbb{R}^2 admet des sous-suites convergentes. En fait, on peut montrer que ses valeurs d'adhérence forment exactement le cercle unité.

1.3 En dimension infinie

En dimension infinie, il existe toujours des fermés bornés qui ne sont pas compacts. Voici quelques exemples.

Exemples :

- On considère l'espace ℓ^∞ des suites bornées muni de la norme $\|(u_n)\|_\infty = \sup_{n \in \mathbb{N}} |u_n|$. On regarde la suite de suites $(u^k)_{k \in \mathbb{N}}$ où u^k est la suite où $u_n^k = 0$ sauf si $k = n$ et alors $u_n^k = 1$. On peut la voir comme la suite de vecteurs avec une infinité de composantes

$$\begin{aligned} u^0 &= (1, 0, 0, 0, \dots) \\ u^1 &= (0, 1, 0, 0, \dots) \\ u^2 &= (0, 0, 1, 0, \dots) \\ &\vdots \quad \vdots \end{aligned}$$

Chaque u^k est de norme 1 et donc la suite (u^k) est bornée. Mais pour tout p et q , on a $\|u^p - u^q\|_\infty = 1$. Donc tous les termes sont éloignés les uns des autres et on ne pourra jamais en extraire une suite de Cauchy. Donc on ne peut en extraire une sous-suite convergente. Cela montre que dans ℓ^∞ , la boule unité fermée est un fermé borné qui n'est pas compact.

- On peut reprendre l'exemple ci-dessus dans l'espace des polynômes $\mathbb{R}[X]$ muni de la norme $\|a_d X^d + a_{d-1} X^{d-1} + \dots + a_1 X + a_0\| = \max_j |a_j|$ (cf le corollaire 3.18). La suite correspondante est la suite des polynômes $P_n = X^n$ qui sont tous de norme 1 et tous à distance 1 les uns des autres.
- On se place dans $E = \mathcal{C}^0([0,1], \mathbb{R})$ muni de la norme infini, qui correspond à la convergence uniforme de fonctions. On considère la suite de fonctions $f_n : x \mapsto x^n$, qui sont toutes de norme 1. Les fonctions convergent simplement vers la fonction f_* définie par $f_*(x) = 0$ si $x \neq 1$ et $f_*(1) = 1$. Si on pouvait extraire une sous-suite de fonctions $(f_{\varphi(n)})$ convergeant uniformément, alors celle-ci convergerait aussi simplement et sa limite devrait être f_* . Mais la limite uniforme de fonctions continue est continue et f_* n'est pas continue (on pourra revoir ses cours de L2 pour se rappeler les convergences de suites de fonctions). C'est donc impossible et on ne peut pas extraire une sous-suite convergente de la suite (f_n) . On vient de montrer que la boule unité fermée de $\mathcal{C}^0([0,1], \mathbb{R})$ muni de la norme infini est un fermé borné qui n'est pas compact.

Pour trouver des compacts dans un espace de dimension infinie, on peut regarder des fermés bornés dans des sous-espaces de dimension finie qui donne alors des compacts. Par exemple, dans l'espace ℓ^∞ des suites bornées muni de la norme $\|\cdot\|_\infty$, l'ensemble

$$K = \{(u_n) \in \ell^\infty, \forall n \leq 10, |u_n| \leq 1 \text{ et } \forall n \geq 11, u_n = 0\}$$

est compact. Mais c'est un peu « triché ». Pour donner un exemple de compact de dimension infinie, on va montrer le résultat suivant.

Proposition 6.6

On considère l'espace ℓ^∞ des suites bornées muni de la norme $\|(u_n)\|_\infty = \sup_{n \in \mathbb{N}} |u_n|$. L'ensemble

$$K = \{(u_n) \in \ell^\infty, \forall n \geq 0, |u_n| \leq e^{-n}\}$$

est compact dans ℓ^∞ .

Démonstration : Soit $(u^k)_{k \in \mathbb{N}}$ une suite de K , où chaque u^k est une suite (u_n^k) de ℓ^∞ . On applique un argument d'extraction diagonal. On peut extraire de la suite des premières coordonnées une sous-suite $(u_1^{\phi_1(k)})$ qui converge vers u_1^∞ dans $[-1,1]$. Puis de la sous-suite $(u_2^{\phi_1(k)})$, on peut extraire une sous-suite $(u_2^{\phi_1 \circ \phi_2(k)})$ qui converge vers $u_2^\infty \in [-e^{-1}, e^{-1}]$ et on continue ainsi de suite. On ne peut pas extraire une infinité de fois sous peine de risquer de ne plus avoir de termes restants. Par contre, on peut regarder la suite (v^k) extraire de (u^k) par l'extraction dite « diagonale »

$$v^k = u^{\phi_1 \circ \dots \circ \phi_k(k)}.$$

On vérifie que v_n^k converge vers v_n^∞ pour tout n : dès que $k \geq n$, on obtient une suite extraite de l'extraction convergente $u_n^{\phi_1 \circ \dots \circ \phi_n(k)}$ construite ci-dessus. Par définition de K et la convergence de chaque coordonnée, on a forcément $|v_n^\infty| \leq e^{-n}$ et donc v^∞ appartient à K .

Montrons que v^k converge vers v^∞ . Soit $\varepsilon > 0$, on choisit $n_0 \in \mathbb{N}$ tel que $e^{-n_0} < \varepsilon/2$. Puis, on choisit k_0 tel que pour tout $k \geq k_0$ et $n < n_0$, $|v_n^k - v_n^\infty| < \varepsilon$. Notons que cela est possible car on ne considère qu'un nombre fini de coordonnées n . On a alors, pour tout $k \geq k_0$,

$$\begin{aligned} \|v^k - v^\infty\|_\infty &= \sup_n |v_n^k - v_n^\infty| \\ &= \max \left(\max_{n < n_0} |v_n^k - v_n^\infty|, \sup_{n \geq n_0} |v_n^k - v_n^\infty| \right) \\ &\leq \max \left(\max_{n < n_0} |v_n^k - v_n^\infty|, \sup_{n \geq n_0} |v_n^k| + \sup_{n \geq n_0} |v_n^\infty| \right) \\ &< \max \left(\varepsilon, \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \right) = \varepsilon. \end{aligned}$$

Ceci montre bien que $v^\infty \in K$ est limite de la suite extraire (v^k) et donc K est compact. \square

Un exemple important est dû à Giulio Ascoli (1843-1896, Italie). Pour avoir de la compacité dans les ensembles de fonctions, il faut que la continuité soit en quelque sorte uniforme, par exemple avec une pente ou une dérivée bornée.

Proposition 6.7 (théorème d'Ascoli)

Soient $a < b$ et soit $E = \mathcal{C}^0([a,b],\mathbb{R})$ muni de la norme $\|\cdot\|_\infty$. Un ensemble K de fonctions est compact si et seulement si :

- (i) K est fermé,
- (ii) K est borné
- (iii) et K est *équicontinu* c'est-à-dire que

$$\forall \varepsilon > 0, \exists \delta > 0, \forall f \in K, \forall x, x' \in [a,b], |x-x'| \leq \delta \Rightarrow |f(x)-f(x')| \leq \varepsilon.$$

Corollaire 6.8

Si (f_n) est une suite de fonctions de classe $\mathcal{C}^1([a,b],\mathbb{R})$ telle que $\|f_n\|_\infty$ et $\|f'_n\|_\infty$ sont bornées uniformément, alors on peut trouver une sous-suite $(f_{\varphi(n)})$ et une fonction continue $f \in \mathcal{C}^0([a,b],\mathbb{R})$ telles que $\|f_{\varphi(n)} - f\|_\infty \rightarrow 0$.

Démonstration : Soit $M \in \mathbb{R}$ qui majore $\|f_n\|_\infty$ et $\|f'_n\|_\infty$ pour tout n . On se place dans $\mathcal{C}^0([a,b],\mathbb{R})$ muni de la norme infinie. La suite (f_n) est incluse dans l'ensemble

$$K' = \{f \in \mathcal{C}^0([a,b],\mathbb{R}), f \text{ est dérivable et } \|f\|_\infty \leq M \text{ et } \|f'\|_\infty \leq M\}.$$

On note $K = \overline{K'}$ l'adhérence de cet ensemble. Il nous suffit maintenant d'appliquer le théorème d'Ascoli ci-dessus à K .

Comme K' est borné, K est borné et par ailleurs K est fermé par construction. On va montrer que K est équicontinu. On commence par constater que toute fonction de K' est M -lipschitzienne car sa dérivée est bornée par M (cf proposition 4.20). Il en est de même pour les fonctions de K . En effet, si $(g_n) \subset K'$ converge vers $g \in K$, alors $|g_n(x) - g_n(x')|$ converge vers $|g(x) - g(x')|$ et la majoration $|g_n(x) - g_n(x')| \leq M|x - x'|$ passe à la limite. Soit $\varepsilon > 0$ et soit $\delta = \varepsilon/M$. Comme toute fonction $g \in K$ est M -lipschitzienne, on a, pour tout x, x' avec $|x - x'| \leq \delta$, $|g(x) - g(x')| \leq M|x - x'| \leq \varepsilon$. □

On peut donc extraire des sous-suites convergentes aux fonctions, à condition de « perdre » de la régularité au passage. Il s'agit d'un principe important en analyse fonctionnelle (pour obtenir de la compacité, il faut donc « gagner » de la régularité par un certain argument).

1.4 Quelques propriétés des compacts

Nous finissons cette partie par de petits résultats qui peuvent être utiles pour montrer que des ensembles sont compacts. Leur preuve font aussi de bons entraînements pour manier le concept de compacité.

Proposition 6.9

Soit $(E, \|\cdot\|)$ un espace vectoriel normé. Soit K_1, \dots, K_p un nombre fini de compacts de E . Alors l'union $K_1 \cup \dots \cup K_p$ est un compact de E .

Démonstration : Utilisons la caractérisation par recouvrement. Soit $(\mathcal{O}_j)_{j \in J}$ un recouvrement de $K_1 \cup \dots \cup K_p$ par des ouverts. Comme chaque K_i est compact et est recouvert par $\cup_j \mathcal{O}_j$, on peut trouver un recouvrement fini $K_i \subset \mathcal{O}_{j_1^i} \cup \dots \cup \mathcal{O}_{j_{q_i}^i}$. Mais alors l'union de tous les recouvrements finis fournit un recouvrement fini de l'union des K_i . \square

Il faut évidemment se limiter à un nombre fini de compacts car une union infinie de bornés n'a aucun raison d'être bornée par exemple. Pour l'intersection, on peut être plus général.

Proposition 6.10

Soit $(E, \|\cdot\|)$ un espace vectoriel normé. Soit K un compact de E et F un fermé de E , alors $K \cap F$ est un compact de E .

Démonstration : Soit (x_n) une suite de $K \cap F$. C'est a fortiori une suite du compact K et donc on peut en extraire une sous-suite qui converge vers $\ell \in K$. Il reste à voir que F est fermé et donc que la limite est aussi dans F et donc au final $\ell \in K \cap F$. \square

Corollaire 6.11

Une intersection quelconque de compacts est compact.

Démonstration : Soit $(K_j)_{j \in J}$ une famille de compacts. Soit K_{j_0} l'un deux et soit $F = \cap_{j \neq j_0} K_j$ l'intersection des autres. Comme les K_j sont compacts, ils sont fermés et F est fermé comme intersection de fermés. Donc $\cap_j K_j = K_{j_0} \cap F$ est un compact. \square

On a aussi une généralisation du théorème des segments emboîtés.

Proposition 6.12

Soit (K_n) une suite de compacts non vides de E qui sont emboîtés dans le sens où $K_{n+1} \subset K_n$. Alors $\cap_n K_n$ est un compact non vide.

Démonstration : Pour tout n , on choisit $x_n \in K_n$. Par emboîtement, la suite (x_n) est incluse dans K_0 qui est compact, donc on peut en extraire une sous-suite $(x_{\varphi(n)})$ qui converge vers $\ell \in K_0$. Mais pour tout N , la suite $(x_{\varphi(n)})_{n \geq N}$ est incluse dans K_N qui est fermé et donc $\ell \in K_N$ pour tout N . \square

2 Fonctions et compacité

En dimension $d = 1$, nous avons deux théorèmes importants utilisant la continuité et la compacité des intervalles $[a, b]$. Maintenant que le travail de préparation des bonnes notions est fait, la généralisation est immédiate.

Théorème 6.13 (théorème de Heine)

Soient $(X, \|\cdot\|_X)$ et $(Y, \|\cdot\|_Y)$ deux espaces vectoriels normés et soit $K \subset X$ un ensemble compact. Alors toute fonction f continue de K dans Y est uniformément continue.

Démonstration : Soit $\varepsilon > 0$. Comme f est continue, pour tout $x \in K$, il existe un rayon $\delta(x) > 0$ tel que pour tout $x' \in B(x, \delta(x))$, $\|f(x) - f(x')\|_Y < \varepsilon/2$. Nous considérons maintenant toutes les boules $B(x, \delta(x)/2)$ pour $x \in K$, qui forment une famille d'ouverts qui recouvrent K (puisque chaque x est au moins dans la boule dont il est le centre). Le théorème de Borel-Lebesgue montre que la compacité de K implique qu'on le peut recouvrir par un nombre fini de boules $B(x_i, \delta(x_i)/2)$, $i = 1, \dots, p$. On pose $\delta = \min_i \delta(x_i)/2$. Soient x et x' deux points de K avec $\|x - x'\|_X < \delta$. Par construction, x est dans une des boules $B(x_i, \delta(x_i)/2)$ et par inégalité triangulaire, x' est au pire dans la boule $B(x_i, \delta(x_i))$. Par définition de cette boule, on a donc $\|f(x) - f(x_i)\|_Y < \varepsilon/2$ et $\|f(x') - f(x_i)\|_Y < \varepsilon/2$. Par inégalité triangulaire, on obtient donc que $\|f(x) - f(x')\|_Y < \varepsilon$. On vient de démontrer l'uniforme continuité de f . \square

Le théorème des valeurs extrêmes se généralise ainsi.

Théorème 6.14

Soient $(X, \|\cdot\|_X)$ et $(Y, \|\cdot\|_Y)$ deux espaces vectoriels normés et soit $K \subset X$ un ensemble compact. Alors l'image $f(K)$ de K par une fonction continue $f : K \rightarrow Y$ est un compact de Y .

Démonstration : Nous allons utiliser le critère séquentiel. Soit (y_n) une suite de l'image $f(K)$. Par définition, il existe une suite de points $(x_n) \subset K$ tels que $f(x_n) = y_n$. Comme K est compact, on peut extraire de cette suite une sous-suite $(x_{\varphi(n)})$ qui converge vers un point $x_* \in K$. Par continuité, on a $f(x_{\varphi(n)})$ qui converge vers $y_* := f(x_*)$. La suite $(y_{\varphi(n)})$ est bien une sous-suite de (y_n) qui converge vers un point y_* de $f(K)$. Donc $f(K)$ est compact. \square

Le théorème précédent est important et s'exprime plus littéralement par « l'image d'un compact par une fonction continue est un compact ». On fera attention au fait que le sens est différent du sens transférant les propriétés d'ouverture et de fermeture. En particulier, l'image réciproque d'un compact par une fonction continue est forcément un fermé mais pas forcément un compact : l'image réciproque de $\{0\}$ par la fonction constante $f : x \in \mathbb{R} \mapsto 0$ est tout \mathbb{R} , qui n'est pas compact.

Un corollaire immédiat du théorème 6.14 est le suivant.

Corollaire 6.15

Soient $(X, \|\cdot\|_X)$ un espace vectoriel normé et soit $K \subset X$ un ensemble compact non vide. Soit f une fonction continue de K dans \mathbb{R} . Alors f est bornée et atteint ses bornes.

Démonstration : Le théorème 6.14 montre que $f(K)$ est un compact de \mathbb{R} , qui est forcément fermé et borné. La borne et le fait que K est non vide nous dit que $\sup_{x \in K} f(x)$ est bien défini. Le fait que l'image est fermé oblige le sup à appartenir à l'image et donc il s'agit d'un max. Le raisonnement est le même pour la borne inférieure. \square

Exemples :

- Une courbe continue de \mathbb{R}^2 décrite par un paramétrage continue $f : t \in [0,1] \mapsto f(t) \in \mathbb{R}^2$ est un compact de \mathbb{R}^2 . En particulier, c'est un ensemble borné.
- Si un coût $c(x,y,z)$ dépend continuellement de trois paramètres dans $[0,1]$, alors il existe forcément un point $(x_*, y_*, z_*) \in [0,1]^3$ où ce coût est minimal.
- Soit K un compact de \mathbb{R}^d . On vérifie facilement que K^2 est un compact de \mathbb{R}^{2d} (par exemple parce qu'il est fermé et borné en utilisant la convergence composante par composante). La fonction $f : (x, x') \in K^2 \mapsto \|x - x'\|_2 \in \mathbb{R}$ est continue comme composée de fonctions continues, elle atteint donc son maximum en un couple de points $(x_1, x_2) \in K^2$. Ce couple est le plus éloigné possible dans K : c'est une réalisation de son diamètre (cf figure 4.6.1).
- Soit $f \in C^0([0,1], \mathbb{R})$ une fonction continue et qui est donc bornée par une constante M . On considère l'ensemble K des polynômes $P \in \mathbb{R}_d[X]$ tels que $\|f - P\|_\infty \leq M$. C'est un ensemble non vide (car $P = 0$ est dedans), il est borné (car par inégalité triangulaire $\|P\|_\infty \leq 2M$) et fermé (par continuité de la norme). Comme $\mathbb{R}_d[X]$ est un espace de dimension finie, K est un compact. Il existe donc un polynôme P_* tel que l'erreur $\|f - P_*\|_\infty$ est minimale parmi toutes les approximations de f . Il est par contre délicat de savoir précisément lequel ce qu'il vaut, un exemple simple est donné en figure 6.1. Si on considère la norme $\|\cdot\|_2$ plutôt que la norme infini, alors on a l'outil du produit scalaire et on peut montrer que P_* est la projection orthogonale de f sur les polynômes de degré au plus d .

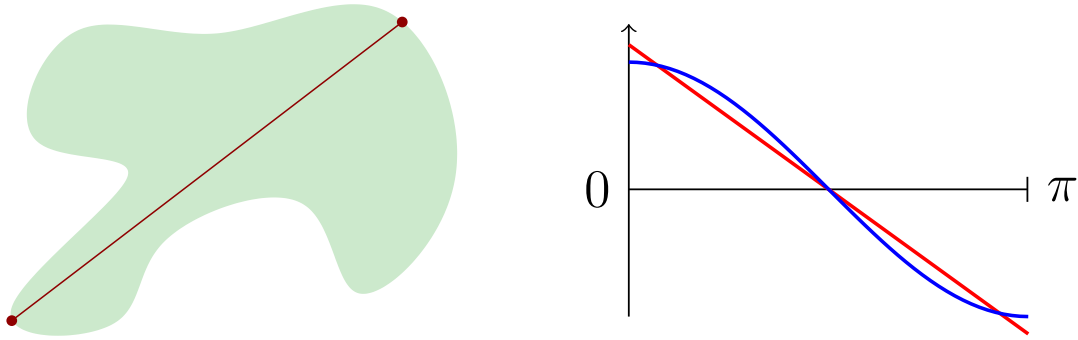


FIGURE 6.1 – À gauche, un exemple de compact du plan et de ses deux points les plus éloignés possibles. À droite, la fonction cosinus et le polynôme de degré 1 qui minimise le maximum de l'erreur sur $[0, \pi]$.

Application aux vecteurs propres des matrices symétriques

Soit $A \in \mathcal{M}_d(\mathbb{R})$ une matrice symétrique, c'est-à-dire que $A^t = A$. On munit \mathbb{R}^d de la norme $\|\cdot\|_2$ qui est associée au produit scalaire canonique. On note S la sphère unité de \mathbb{R}^d , c'est-à-dire $S = \{x \in \mathbb{R}^d, \|x\|_2 = 1\}$. Pour tout $x \in \mathbb{R}^d$, on note $f(x) = \langle Ax|x \rangle$. La fonction f est continue sur S car elle est $2\|A\|$ -lipschitzienne

$$\begin{aligned} |f(x) - f(y)| &= |\langle Ax|x \rangle - \langle Ay|y \rangle| = |\langle Ax|(x - y) \rangle - \langle A(x - y)|y \rangle| \\ &\leq |\langle Ax|(x - y) \rangle| + |\langle A(x - y)|y \rangle| \leq \|Ax\|_2 \|x - y\|_2 + \|A(x - y)\|_2 \|y\|_2 \\ &\leq \|A\| \cdot \|x\|_2 \|x - y\|_2 + \|A\| \cdot \|x - y\|_2 \|y\|_2 \\ &\leq 2\|A\| \cdot \|x - y\|_2 . \end{aligned}$$

Par ailleurs, S est fermée bornée dans \mathbb{R}^d de dimension finie, donc S est compacte. On en déduit que f admet un maximum sur S , atteint en un point x_* . Soit Y l'espace orthogonal à x_* et soit $y \in Y$. Les points $x_t = \cos(t)x_* + \sin(t)y$ sont sur la sphère S par théorème de Pythagore (ou calcul explicite). On a que, pour $t \in \mathbb{R}$ petit, $x_t = x_* + ty + \mathcal{O}(t^2)$ et donc

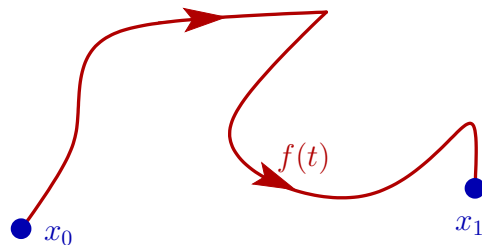
$$\begin{aligned} f(x_t) &= \langle Ax_*|x_* \rangle + t\langle Ax_*|y \rangle + t\langle Ay|x_* \rangle + \mathcal{O}(t^2) \\ &= f(x_*) + t\langle Ax_*|y \rangle + t\langle y|Ax_* \rangle + \mathcal{O}(t^2) \\ &= f(x_*) + 2t\langle Ax_*|y \rangle + \mathcal{O}(t^2) . \end{aligned}$$

Comme $f(x_*)$ est maximum, ce développement implique que $y \perp Ax_*$. Comme c'est vrai pour tout $y \in Y$, c'est que Ax_* est colinéaire à x_* et donc x_* est vecteur propre de A . Puis on considère maintenant le même problème mais dans Y . Comme $\langle Ay|x_* \rangle = \langle y|Ax_* \rangle = 0$, on sait que A envoie Y sur Y . On peut recommencer la méthode dans un espace ayant une dimension de moins. En continuant ainsi, on construit une base orthonormale de vecteurs propres pour A : toute matrice symétrique est diagonalisable dans une base orthonormale.

Chapitre 7 : Compléments

1 Connexité par arc

Soit $(E, \|\cdot\|)$ un espace vectoriel normé. Une *courbe*, un *chemin* ou un *arc* dans E est l'image d'une fonction continue $f : [0,1] \rightarrow E$. Cet arc relie deux points x_0 et x_1 si $f(0) = x_0$ et $f(1) = x_1$. Comme f est continue, la courbe peut être irrégulière, avoir des coudes, être une ligne brisée etc. mais elle ne peut pas faire des « sauts ».



Un chemin reliant deux points du plan

Définition 7.1

Un ensemble C de E est dit *connexe par arcs* si tous points x_0 et x_1 de C peuvent être reliés par un chemin $f : [0,1] \rightarrow E$ continu restant dans C , c'est-à-dire que $f(0) = x_0$, $f(1) = x_1$ et pour tout $t \in [0,1]$, $f(t)$ est dans C .

Proposition 7.2

Un ensemble C convexe est connexe et donc les boules de E sont connexes.

Démonstration : Soit x_0 et x_1 dans C . Par définition de la convexité, tout le segment $\theta x_1 + (1 - \theta)x_0$ avec $\theta \in [0,1]$ est dans C . Il suffit donc de prendre $f(t) = (1 - t)x_0 + tx_1$. On vérifie que f est bien continue puisque

$$\|f(t) - f(t')\| = \|(t' - t)x_0 + (t - t')x_1\| \leq (\|x_0\| + \|x_1\|)|t - t'|$$

ce qui montre que f est lipschitzienne. \square

Proposition 7.3

La propriété d'être reliable par un arc est une relation d'équivalence parmi les points d'un ensemble A .

Démonstration : Il est évident qu'un point x_0 est relié à lui-même par l'arc constant $f : t \mapsto x_0$. Si x_0 est relié à x_1 par le chemin $f : [0,1] \rightarrow A$, alors x_1 est relié à x_0 par le chemin $\tilde{f} : t \mapsto f(1-t)$ (\tilde{f} est continue comme composée de fonction continue). Si f est un chemin reliant x_0 à x_1 et g un chemin reliant x_1 à x_2 , alors le chemin

$$h : t \in [0,1] \mapsto \begin{cases} f(2t) & \text{si } t \leq \frac{1}{2} \\ g(2t-1) & \text{si } t > \frac{1}{2} \end{cases}$$

est un chemin reliant x_0 à x_2 (il suffit de voir que les courbes se recollent en $t = 1/2$ en x_1 , ce qui assure la continuité de h , voir la figure 7.1). \square

On appelle *composantes connexes (par arcs)* d'un ensemble les classes d'équivalences des points reliables par des chemins. Chaque composante connexe est connexe et leur réunion est tout l'ensemble. Par construction, aucun point d'une composante ne peut être relié à un point d'une autre composante.

Proposition 7.4

Si A et B sont connexes par arcs et si $A \cap B \neq \emptyset$, alors $A \cup B$ est connexe par arc.

Démonstration : Soit x dans $A \cap B$. Par hypothèse, tout point de A est relié à x par un arc et idem pour tout point de B . Par la transitivité prouvée ci-dessus, on peut donc aussi relier les points de A à ceux de B . \square

! On ne peut pas dire mieux que la proposition ci-dessus pour les unions et intersections comme on le voit dans la figure 7.1 ci-dessous.

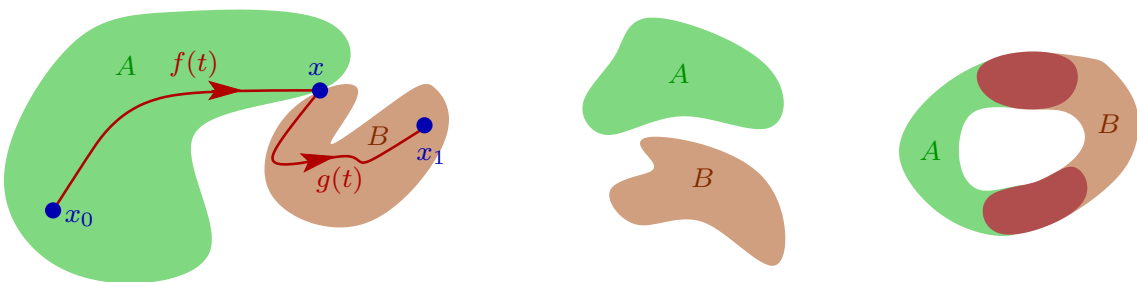


FIGURE 7.1 – À gauche, une illustration de la proposition 7.4. Au centre et à droite, des exemples d'union et d'intersection non-connexes par arcs d'ensembles connexes par arcs.

Proposition 7.5

L'image par une fonction continue d'un ensemble connexe par arcs est connexe par arcs.

Démonstration : Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces vectoriels normés et soit $f : E \rightarrow F$ continue. Soit A un ensemble connexe par arcs de E et soit y et y' deux points de $f(A)$. Par définition, il existe deux points x et x' de A tels que $f(x) = y$ et $f(x') = y'$. Comme A est connexe par arcs, il existe un chemin continu $h : [0,1] \rightarrow A$ tel que $h(0) = x$ et $h(1) = x'$. Par composition, la fonction $f \circ h$ est un chemin continu reliant y à y' . Donc $f(A)$ est bien connexe par arcs. \square

Proposition 7.6

Soit A un ensemble connexe par arcs d'un espace vectoriel normé $(E, \|\cdot\|)$ et soit $f : A \rightarrow \mathbb{R}$ une fonction continue à valeurs réelles. Soient a et b deux points de A tels que $f(a) < f(b)$, alors pour toute valeur $y \in]f(a), f(b)[$, il existe $c \in A$ tel que $f(c) = y$.

Démonstration : Comme A est connexe par arcs, il existe $h : [0,1] \rightarrow A$ continue telle que $h(0) = a$ et $h(1) = b$. La fonction $f \circ h$ est alors une fonction continue de $[0,1]$ dans \mathbb{R} telle que $(f \circ h)(0) < (f \circ h)(1)$ et par le théorème des valeurs intermédiaires classique, il existe un temps $t_* \in]0,1[$ tel que $(f \circ h)(t_*) = y$. Le point $c = h(t_*)$ convient alors. \square

Exemples :

- Dans tout espace vectoriel normé, la norme est une fonction continue. Donc la sphère unité sépare un « intérieur » $A = \{x \in E, \|x\| < 1\}$ d'un « extérieur » $B = \{x \in E, \|x\| > 1\}$ dans le sens où si un chemin h passe de A à B , alors il passe forcément à travers la sphère unité (on applique la proposition 7.6 à $\|h(\cdot)\|$).
- Dans tout espace vectoriel normé de dimension plus grande que 2 (donc tous sauf \mathbb{R}), la sphère unité est connexe par arc. En effet, si x et $x' \neq -x$ sont de norme 1, alors le chemin

$$h : t \in [0,1] \mapsto h(t) = \frac{1}{\|(1-t)x + tx'\|} ((1-t)x + tx')$$

est bien défini et continu sur la sphère (car $(1-t)x + tx' \neq 0$ pour tout t). Si on considère $x' = -x$ et si on est en dimension plus grande que 2, il existe au moins un autre point sur la sphère unité et on peut simplement passer via ce troisième point en collant deux chemins comme ci-dessus.

- Sur Terre, il existe toujours deux points diamétralement opposés où la température est identique. En effet, la surface de la Terre est connexe par arc. On pose $f(x) = T(x) - T(-x)$ où T est la fonction « température » et où $-x$ désigne le point diamétralement opposé à x . Physiquement, il est raisonnable que f soit continue et $f(x) = -f(-x)$. Donc si en un point x_0 , on a $T(x_0) \neq T(-x_0)$, alors $f(x_0)$ et $f(-x_0)$ sont de signes opposés et il existe

un point x_* tel que $f(x_*) = 0$, c'est-à-dire $T(x_*) = T(-x_*)$. En fait, on voit qu'il suffit de travailler sur un cercle. Avec la sphère bidimensionnelle, on peut même dire mieux : on peut avoir la même température et la même pression.

2 Homéomorphismes

2.1 Homéomorphismes

Un concept extrêmement important en topologie est celui d'homéomorphisme.

Définition 7.7

Soient $A \subset X$ et $B \subset Y$ deux sous-ensembles d'espaces vectoriels normés. On dit que f est un **homéomorphisme** entre A et B si f est bijective de A vers B et si f et f^{-1} sont continues.

Si $A \subset X$ et $B \subset Y$ sont deux sous-ensembles d'espaces vectoriels normés pour lesquels il existe un homéomorphisme $f : A \rightarrow B$, on dit qu'ils sont **homéomorphes**.

Proposition 7.8

La relation d'homéomorphie est une relation d'équivalence puisqu'elle est

- (a) réflexive : A est homéomorphe à A .
- (b) symétrique : A est homéomorphe à B si et seulement si B est homéomorphe à A .
- (c) transitive : si A est homéomorphe à B et si B est homéomorphe à C , alors A est homéomorphe à C .

Démonstration : La fonction identité $Id : x \in A \rightarrow x \in A$ est une bijection et est toujours continue car 1-lipschitzienne. Sa réciproque étant elle-même, (a) est clair. Pour (b), il suffit de voir que si $f : A \rightarrow B$ est un homéomorphisme, alors $f^{-1} : B \rightarrow A$ en est un aussi. Enfin, si $f : A \rightarrow B$ et $g : B \rightarrow C$ sont des homéomorphismes, alors $g \circ f$ est un homéomorphisme de A vers C , ce qui montre (c). \square

Pour un topologue, deux espaces homéomorphes sont identiques dans le sens où ils ont les mêmes propriétés fondamentales. En effet, en corollaire des caractérisations de la continuité, un homéomorphisme donne une relation parfaite entre les topologies de A et B (si A et B ne sont pas des espaces entiers, on utilise la topologie induite introduite dans les compléments du chapitre 4).

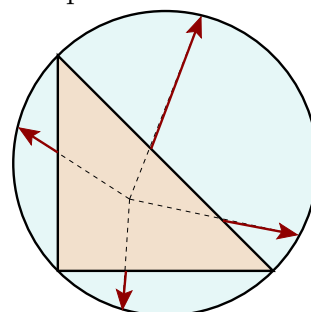
Proposition 7.9

Soient $A \subset X$ et $B \subset Y$ deux sous-ensembles d'espaces vectoriels normés et f un homéomorphisme entre A et B .

- $\mathcal{O} \subset A$ est un ouvert de A si et seulement si $f(\mathcal{O})$ est un ouvert de B
- $\mathcal{F} \subset A$ est un fermé de A si et seulement si $f(\mathcal{F})$ est un fermé de B
- La suite (u_n) converge dans A si et seulement si $f(u_n)$ converge dans B
- $K \subset A$ est un compact de X si et seulement si $f(K)$ est un compact de Y .
- A est connexe par arcs si et seulement si B est connexe par arcs.

Exemples :

- deux intervalles $]a,b[$ et $]a',b'[$ de \mathbb{R} sont homéomorphes par une translation et une homothétie, c'est aussi le cas de deux intervalles du type $[a,b]$, mais aussi de deux intervalles du type $[a,b[$ ou $]a,b]$ (en ajoutant une éventuelle symétrie).
- $[a,b]$ n'est ni homéomorphe à $]a,b[$ ni à $[a,b[$ puisque le premier est compact alors que les deux autres ne le sont pas.
- Les intervalles $]a,b[$ sont homéomorphes à \mathbb{R} et aux intervalles $]a,+\infty[$. En effet, la tangente est un homéomorphisme entre $] -\pi/2,\pi/2[$ et \mathbb{R} et l'exponentielle entre \mathbb{R} et $]0,+\infty[$.
- Les ensembles $[0,1]$ et $[0,1] \cup \{2\}$ ne sont pas homéomorphes : l'un est connexe par arcs et l'autre non.
- Les ensembles $]a,b[$ et $[a,b[$ ne sont pas homéomorphes. En effet, s'il existait un homéomorphisme $h : [a,b[\rightarrow]a,b[$, alors $h|_{]a,b[} :]a,b[\rightarrow]a,b[\setminus \{h(a)\}$ serait aussi un homéomorphisme. Mais $]a,b[$ est connexe par arcs alors que $]a,b[$ privé de n'importe quel point est séparé en deux composantes connexes. La même astuce marche pour montrer que \mathbb{R} et \mathbb{R}^2 ne sont pas homéomorphes : si on retire un point à \mathbb{R}^2 , il reste connexe par arcs, mais ce n'est pas le cas de \mathbb{R} .
- Toutes les boules ouvertes (resp. fermées) d'un espace vectoriel normé sont homéomorphes entre elles par translation et homothéties.
- Tous les triangles du plan sont homéomorphes entre eux par des déformations affines : translations, symétries, compression et dilatations, rotations... Mais les triangles pleins du plan sont aussi homéomorphes aux disques fermés. Pour cela, on peut utiliser des dilatations qui dépendent de la direction comme dans le dessin ci-contre. L'idée fonctionne avec tous les polygones. Donc tous les polygones pleins sont homéomorphes entre eux et homéomorphes aux disques.



2.2 Le point fixe de Brouwer

On a énoncé dans la proposition 4.33 un théorème de point fixe dans les triangles, dont on rappelle ici l'énoncé.

Proposition 7.10 (point fixe de Brouwer dans un triangle)

Soit $T \subset \mathbb{R}^2$ le triangle rectangle isocèle de sommets $(0,0)$, $(1,0)$ et $(0,1)$, bords compris. Toute fonction continue $f : T \rightarrow T$ admet un point fixe.

Grâce aux homéomorphismes, nous pouvons passer à d'autres formes géométriques.

Théorème 7.11 (point fixe de Brouwer en dimension 2)

Si F est un fermé de \mathbb{R}^2 homéomorphe au disque fermé, alors toute fonction continue $f : F \rightarrow F$ admet un point fixe.

Démonstration : On a vu plus haut que le triangle T est homéomorphe au disque fermé. Donc une forme homéomorphe au disque est aussi homéomorphe à T . Soit $h : T \rightarrow F$ une bijection telle que h et h^{-1} sont continues. On pose $g = h^{-1} \circ f \circ h$ qui est bien définie et continue de T dans T . La proposition 5.10 montre que g a un point fixe x dans T . Mais alors $h^{-1}(f(h(x))) = x$ et donc $f(h(x)) = h(x)$, ce qui montre que $h(x)$ est un point fixe de f dans F . \square

On a vu que ce résultat de point fixe ne peut pas se généraliser pour toute forme de \mathbb{R}^2 puisque la rotation dans l'anneau F ci-contre est une fonction continue sans point fixe. Cela montre qu'un anneau et un disque ne sont pas homéomorphes. En fait, on peut classer l'ensemble des surfaces géométriquement raisonnables (bords pas trop irréguliers etc.) en fonction de leur nombre de « trous » : deux surfaces sont homéomorphes si et seulement si elles ont le même nombre de « trous ».

